TWO SIMULATION APPROACHES FOR EVALUATING CATCH CURVE MODELS AS AN ASSESSMENT METHOD FOR RIVER HERRING

by

Mark Billard

Thesis
submitted in partial fulfillment of the requirements for
the Degree of Master of Science (Biology)

Acadia University
Spring Convocation 2020

This thesis by Mark Billard was defended successfully in an oral examination on 26 March 2020.

The examining committee for the thesis was:

____
Dr. Tanya Surette, Chair

____
Dr. Cindy Breau, External Examiner

____
Dr. Trevor Avery, Internal Examiner

____
Dr. A. Jamie F. Gibson, Supervisor

____
Dr. Anna Redden, Co-Supervisor

____
Dr. Rodger Evans, Head (Biology)

This thesis is accepted in its present form by the Division of Research and Graduate Studies as satisfying the thesis requirements for the degree Master of Science (Biology).

....

I, Mark Billard, grant permission to the University Librarian at Acadia University to archive, preserve, reproduce, loan or distribute copies of my thesis in microform, paper, or electronic formats on a non-profit basis. I undertake to submit my thesis, through my University, to Library and Archives Canada and to allow them to archive, preserve, reproduce, convert into any format, and to make available in print or online to the public for non-profit purposes. I, however, retain the copyright in my thesis.

Author

Supervisor

Co-Supervisor

Date

Table of Contents

List of Tables

List of Figures

Abstract

Stock assessment methods for river herring (Alewife, *Alosa pseudoharengus*, and Blueback Herring, *A. aestivalis*) were tested using two simulation approaches as part of a larger assessment program for river herring in DFO's Maritimes Region. Catch curve models, which produce estimates of the instantaneous total mortality rate (Z) from age composition data collected in a single year were the type of stock assessment model evaluated in each approach. In the first set of simulations statistical criteria (bias and precision) were used to evaluate eight versions of catch curve models. All models performed poorly, with highly variable and negatively biased estimates of Z, indicating that catch curves are not useful for river herring. In the second set of simulations management strategy evaluation was used to evaluate the performance of a fishery where catch curves were used to inform management of the fishery. The preferred catch curve model was used in conjunction with a suite of harvest control rules, assessment schedules and management schedules to determine whether spawning escapements, landings and exploitation rates could be kept within pre-defined boundaries if catch curves are used for the stock status assessment. Despite the poor performance of catch curve models from a statistical perspective, they led to reduced exploitation rates in fisheries where over-exploitation was occurring and to increased exploitation rates for under-exploited stocks, in addition to increased landings in the fishery. The results of these simulations are dependent on the assumptions of representative sampling, which can be difficult to achieve. The results from the two simulation approaches suggest that the methods used to evaluate stock assessment models can alter the conclusions that are drawn.

Glossary

ASMFC	Atlantic States Marine Fisheries Commission
BRP	Biological reference point
DFO	Fisheries and Oceans Canada
GLM	Generalized linear model
GLMM	Generalized linear mixed effects model
HCR	Harvest control rule
LRP	Limit stock reference point
MSE	Management strategy evaluation
OLS	Ordinary least-squares
PVA	Population viability analysis
SCA	Statistical catch-at-age
SCE	Statistical criteria evaluation
SSB	Spawning stock biomass
SR	Stock recruitment
TAC	Total allowable catch
USR	Upper stock reference point
VPA	Virtual population analysis
Z	Total instantaneous mortality rate

Acknowledgements

Thank you to my supervisors Dr. Jamie Gibson and Dr. Anna Redden for their support, advice, and thoughtful direction throughout this thesis. I thank Dr. Jamie Gibson for the conception of the ideas which led to this work, and for his model development which provided a starting point for the simulation evaluation herein. I am grateful for the opportunities I had throughout this graduate program that he arranged for me, and for the scientific skills he helped me develop. The opportunity to participate in a DFO stock assessment workshop which exposed me to many new ideas, including MSE, was particularly important to me. I thank Dr. Anna Redden for her compassion and expertise in guiding me through this graduate program. Thanks also goes to Dr. Heather Bowlby for her guidance and affirmation, especially during the later stages of completing this work. As I complete this thesis, I realize how fortunate I was to have these mentors. Thanks also go to thesis examiners Dr. Cindy Breau and Dr. Trevor Avery for their insightful comments on the defence version of the thesis and engaging defense questions.

Financial support for this project was coordinated by Dr. Jamie Gibson and Dr. Anna Redden. The Graham Daborn Scholarship and the Acadia Graduate Teaching Assistantship contributed to the overall funding received for this work. Portions of the simulation framework were completed while working under contract for DFO.

I want to thank Mike Adams for the ideas he contributed to this work and for his camaraderie as we worked through our Master’s projects at the Acadia Centre for Estuarine Research. Thank you to Samantha Fulton and George Nau for their help and advice during the completion of this thesis. I also want to thank Freya Keyser for her advice and encouragement to pursue this graduate program. Finally, I would like to thank my partner Jackie Bradbury and my family and friends for their support in helping me reach my goals.

CHAPTER 1: Prolegomenon

Fisheries are managed globally to preserve stocks at healthy levels, maintain commercial fisheries, and to preserve the social and recreational value of the stock (Hilborn and Walters 1992, Caddy and Mahon 1995, Quinn and Deriso 1999). Stock assessments are carried out to inform management of stock status. A stock assessment is the collection and analysis of various types of data from a stock and its fishery to estimate and predict current and future stock status (Hilborn and Walters 1992). Ideally, a stock is delineated as a single, closed population of a fish species that reproduces together and shares biological characteristics such as reproductive rates, natural mortality rate, and maturity schedule (Gibson et al. 2017). Data collected for a stock assessment can be used to estimate abundance, age structure, productivity, instantaneous natural and fishing mortality rates, and other important characteristics of the stock that can indicate stock status (Wallace and Fletcher 2001). To effectively manage stocks, stock assessments should be carried out specific to the habitat and life history of the stock of interest.

The results of a stock assessment are compared to biological reference points (BRP) to determine stock status. BRPs are metrics based on the biological characteristics of a fish stock, the characteristics of the stock's fishery, and objectives for managing the fishery (Caddy and Mahon 1995, Gibson and Myers 2003a). Within Canada, Fisheries and Oceans Canada (DFO) has developed a precautionary approach framework to managing fisheries (DFO 2006). Within the framework stock status can be in one of three zones: healthy, cautious, or critical (DFO 2006, DFO 2012). The zone will determine what the removal rate should be, which is the rate at which fish are removed from the stock by any human process (DFO 2006). The thresholds between the three stock status zones, and the removal rates themselves, are examples of BRPs.

Many different stock assessment methods exist to determine stock status. Stock assessment models can be complex, using different data types over several years to estimate many stock parameters. A statistical catch-at-age model is an example of a more complex stock assessment model; using time series of landings, escapement, larval abundance indices and age-composition data, estimates of abundance and time-varying natural or fishing mortality rates can be produced (Deriso et al. 1985, Gibson and Myers 2003a). Statistical catch-at-age models are flexible and can be modified to include different data types or estimate different parameters of interest. Stock assessment models can also be simple, using a single data type within a single year to estimate a single parameter. Catch curve models, the focus of this thesis, use a single year of age-composition data to produce an estimate of the instantaneous total mortality rate for that year (Ricker 1975). Complex models typically provide more precise parameter estimates than simpler models, at the cost of more data (

and Deriso 1999). In data-limited situations, there may be no other option than to use simple assessment models (ASMFC 2012, Gibson et al. 2017).

River herring comprise two similar species, Alewife, *Alosa pseudoharengus* (Wilson, 1811) and Blueback Herring, *A. aestivalis* (Mitchill, 1814). River herring are short-lived, anadromous, iteroparous fishes that are indigenous to Eastern North America, and are fished throughout their range for bait and food (Loesch 1987). River herring stocks in DFO's Maritimes Region are not regularly assessed; the last regional assessment was in 2001 (Gibson et al. 2017). There is a short-term assessment need for many of these stocks. The main options for assessing status are estimating spawning stock biomass (SSB) or the fishing mortality rate. Estimation of SSB generally requires some method of estimating the escapement from the fishery, which can be facilitated by installation of enumeration systems in fishways (e.g. Gibson and Myers 2001). Reference points for escapement are available for Alewife, which is largely dependent on the available habitat and thus is population specific (Gibson et al. 2017). These reference points delineate the status zones (critical, cautious, healthy) for the stocks and fisheries. River herring stocks in rivers with hydro facilities are ideal candidates for assessment, but many rivers do not have fish passage or BRPs for escapement of the stock, making assessment based on abundance or biomass difficult. Estimation of the fishing mortality rate can be done by ageing a sample of the catch from a fishery and fitting a catch curve model to the sampled age-structure data (Gibson et al. 2017). However, selective fisheries or unrepresentative sampling will bias the sample of the age structure (Billard 2017). BRPs for the fishing mortality rate, which are based on the reproductive rates of the population, are more comparable across populations within a species, therefore calculation of fishing mortality rate BRPs for each population is not as imperative.

Most river herring populations and fisheries within DFO's Maritimes Region have not been assessed (Gibson et al. 2017). When initiating an assessment program on an unassessed stock, use of catch curves to estimate the fishing mortality rate is perhaps the easiest stock assessment method to implement, leading to the method being recommended for new assessments in the region (Gibson et al. 2017). However, recent evaluations of catch curves for other species (Smith et al. 2012, Millar 2014) have shown they are not particularly good methods for estimating fishing mortality rates. For this reason, when recommending their use, Gibson et al. (2017) also recommended that the method be tested via simulation.

This thesis was motivated by the need to develop a method to assess the status of river herring stocks. I evaluated the utility of catch curve models as the sole assessment method for river herring using two simulation-based approaches. In the first chapter, an evaluation was done to determine the accuracy of estimates of the total instantaneous mortality rate produced by catch curve models using statistical criteria to evaluate how well the models estimate total mortality rates for Alewife. In the second chapter, a management strategy evaluation was done with the preferred catch curve model from the first chapter. Through management strategy evaluation the effects of managing a river herring stock, using estimates of the total mortality rate from the catch curve model, were investigated.

1.1 River Herring Biology

Alewife and Blueback Herring, the two species of fish collectively referred to as gaspereau or river herring (Loesch 1987), occupy much of the same region and share many life-history characteristics. However, they differ slightly in morphology, phenology, spawning substrate preference, and geographic range (Loesch 1987).

1.1.1 Life History

River herring are anadromous and iteroparous (Loesch 1987). Male to female sex ratio tends to be equal within a population, possibly favouring males; males tend to mature at younger ages than females, altering the sex ratios of individual age classes (Loesch 1987). Adult Alewives are typically 10-11 inches (25-28 cm), although adults up to 15 inches (36 cm) long have been observed (Bigelow and Schroeder 1953). Adult Blueback Herring are similar in size to Alewives but weigh slightly less on average (Bigelow and Schroeder 1953).

Sexually mature Alewives travel up their natal rivers to spawn each year. Alewives spawn in still waters such as lakes or head ponds (Loesch 1987). The spawning run begins in the spring as waters warm and can continue for up to two months (Scott and Scott 1988). Alewives will remain in the spawning grounds for several days to several weeks before returning to the sea (Scott and Scott 1988). After about a month in fresh water, juvenile Alewives, 50-100 mm long, migrate downstream (Bigelow and Schroeder 1953). Alewives stay at sea until they reach sexual maturity, which is generally at four to five years of age but can be as young as three or as old as six years of age (Bigelow and Schroeder 1953).

The life cycle of Blueback Herring is largely similar to that of Alewife but differs in a few characteristics. Blueback Herring spawn in swift-flowing water, rather than the still waters Alewives spawn in (Loesch and Lund 1977, Loesch 1987). Blueback Herring spawning runs tend to begin three to four weeks after Alewife spawning runs commence (Loesch 1987). Alewife and Blueback Herring are known to mature at different ages. For example, on the Saint John River, Blueback Herring were observed to mature approximately a year earlier on average than Alewife (Messieh 1977).

1.1.2 Distribution

Alewife spawn in rivers in Eastern North America, from South Carolina (Berry 1964) to Newfoundland (Rulifson 1994). Adults overwinter off the coast of these regions before migrating up rivers in the spring. Stone and Jessop (1992) found Alewives at depths of approximately 50 to 200 m, changing throughout the seasons. While at sea, vertical distribution of Alewife may be partially determined by prey availability (Neves 1981). Migration of the most southern Alewife populations begins in late February, while migration of the most northern begins as late as June (Loesch 1987). The distribution of Blueback Herring overlaps with that of Alewife but extends as far south as Florida (Bigelow and Schroeder 1953), and as far north as Cape Breton, Nova Scotia (Alexander 1984) and northeastern New Brunswick (Alexander and Vromans 1983).

1.1.3 Ecological significance

River herring provide an important ecological service to freshwater ecosystems by acting as a vector for marine derived nutrients (Durbin et al. 1979, Garman 1992, Barber et al. 2018). Nutrient input from adults into freshwater river systems is largely from river herring carcasses that remain in the freshwater system, and secondarily from excretion and release of gametes (Durbin et al. 1979, Garman 1992). Durbin et al. (1979) found that 39-57% of Alewives that reach the spawning grounds die there, depositing carbon, nitrogen and phosphorus. The influx of nitrogen and phosphorus from adult Alewife carcasses stimulates a bloom of mycoflora within the riverbed leaf litter, which in turn stimulates detritivore activity in the leaf litter, resulting in a trophic cascade of increased productivity up the food chain (Durbin et al. 1979). Some nutrients are removed from the system by the juveniles as they grow in the freshwater nursery habitat and then migrate to sea. The magnitude and direction of the net nutrient export is closely tied to the stock recruitment relationship (Barber et al. 2018). Low escapements of Alewives resulted in a net outflow of nutrients from juveniles, due to the high number of recruits produced per spawner; the flow of nutrients changes direction as escapement increases, and the number of recruits produced per spawner decreases (Barber et al. 2018).

River herring can significantly impact the freshwater ecosystems they enter in several different ways. River herring prey on a variety of zooplankton, as well as fish eggs, larvae, insects, and amphipods (Guest and Drenner 1991, Mills et al. 1992, Lent 1999). River herring are known to alter zooplankton communities in freshwater systems by selectively predating larger zooplankton (Guest and Drenner 1991, Wells 1970). River herring are also an important forage species for predators in both the freshwater and marine environments (Loesch 1987).

1.1.4 Fisheries

River herring support fisheries that are of economic importance. They are harvested commercially throughout their range, directly and indirectly, by a variety of fishing methods such as weirs, gillnets, square nets and tip-traps (Jessop and Parker 1988; Rulifson 1994; Chaput et al. 2001). River herring are targeted by in-river fisheries and are caught as by-catch in marine fisheries (ASMFC 2012). Several social benefits to restoring and maintaining healthy river herring populations were identified via a survey, namely diversifying local economies and developing a sense of community pride (McClenachan et al. 2015). Restoring river herring populations also helps put a societal emphasis on re-establishing ecosystems to healthy levels (McClenachan et al. 2015).

River herring are fished throughout DFO's Maritimes Region. From 2000-2007, 521 unique individuals reported fishing river herring in DFO's Maritimes Region, most of which were in the Yarmouth/Shelburne area (Gibson et al. 2017). More recent numbers of individuals fishing for river herring has not been assessed. Landings of river herring in the Maritimes Region from 1980 to 1989 have been almost as high as 11,600 t. However, mean annual landings decreased in the 1990s to an annual average of approximately 6,500 t; landings further decreased to an annual average of approximately 2,000 t from 2000 to 2007 (DFO 2001, Gibson et al. 2017).

1.2 Stock Assessment

On a broad scale, stock assessment models can be grouped into two categories: those that are age-structured, and those that are not (ICES 2012). The models within those categories vary in their data requirements, in the parameters which are estimated, and in how those parameters are estimated.

Three general types of age-structured stock assessment models are catch curves, virtual population analyses (VPA), and statistical catch-at-age (SCA) models (Quinn and Deriso 1999). In contrast with biomass dynamics models, which only model changes in biomass, VPAs and SCA models use the age structure of the population to track the abundance of cohorts through time. Age composition data can be used to estimate the age structure of the stock, which in turn can provide estimates of quantities such as annual year class size and mortality rates by modelling the changes in cohort sizes from year-to-year. Several years of data are required for these models to provide estimates of multiple quantities. Age-structured models have one or more of these three components: a growth component, a mortality component, and a stock recruitment component (Quinn and Deriso 1999).

VPAs use age composition data to estimate mortality rates. Gulland (1965) was the first to use a VPA method. VPAs differ from age-structured models like SCA models in that they lack a stock recruitment component. The term virtual is used because the population size is not directly observed or estimated, but back calculated using the age structure from the catch (Quinn and Deriso 1999). VPAs generally use catch and effort data, requiring them to be tuned with a relative abundance index (Butterworth and Rademeyer 2008). ADAPT-VPA, developed by Gavaris (1988) is a popular generalized version of a VPA (Butterworth and Rademeyer 2008). Chaput et al. (2001) developed a VPA for the Margaree River (Nova Scotia) Alewife population.

Statistical catch-at-age models, originally developed by Fournier and Archibald (1982) and built upon by Deriso et al. (1985), use age composition data and auxiliary data to estimate stock parameters such as mortality rates, spawning stock abundance, or number of first-time spawners in a cohort (Gibson and Myers 2003a). Auxiliary data can include fishing effort, escapement counts, selectivity-at-age estimates, stock recruitment relationships, and other sources of information or assumptions about the stock (Deriso et al. 1985). Time series of data can be used to estimate parameters that vary annually for each year that data are available. Statistical catch-at-age models can be used to analyze multiple forms of raw data simultaneously (Maunder and Punt 2013). Using raw data allows both the error structure to be conserved and prevents unnecessary transformation bias from being introduced. Statistical catch-at-age models are highly scalable. At one end, they can be used to estimate a handful of parameters for single unfished stock for a few years (Gibson and Myers 2003a) and at the other end they can be modified to estimate a wide range of parameters for a stock fished by a fleet of vessels (Methot and Wetzel 2013). Stock Synthesis (Methot and Wetzel 2013) is an example of a widely used, highly scalable modeling framework using statistical age-structured methods. Gibson and Myers (2003a) developed SCA models that are specifically designed for river herring based on their unique life history.

In contrast with VPA and SCA models, catch curve models use age-composition data from a single year to estimate the total instantaneous mortality rate of the stock (Ricker 1975). They are simple models that require the assumptions of constant, non-selective mortality rates for a closed population with constant recruitment (Ricker 1975). It is also assumed that the sampling of the age structure and method for ageing is unbiased. These conditions are not typically met for most species and assessments. Catch curves are simple compared to other age-structured models in that they use a single year of age composition data to estimate a single parameter. Different methods exist for estimating the total mortality rate, including closed-form solutions (Heincke 1913, Chapman and Robson 1960) or regression-based estimators (Ricker 1975). Catch curve models are discussed in greater detail in Chapter 2.

While SCA models and VPAs are not the focus of this thesis, the age-composition data used for fitting a catch curve model are also a component of the data requirements for SCA models and VPAs. Collecting multiple years of age-composition data along with a time series of landings data and other auxiliary data would provide a basis for shifting from catch curves to SCA models or VPAs for assessing river herring stocks. The analyses in this thesis are focused on using catch curves as a short-term assessment method for river herring, but either SCA models or VPAs could be used as a more long-term assessment method (ASMFC 2012, Gibson et al. 2017).

1.3 Simulation Testing

In this thesis, I use simulation models to evaluate the role of catch curves in the assessment and management of river herring stocks and fisheries. Simulation evaluation has played a large role in conservation biology, particularly in the form of population viability analyses (PVA) in which populations are projected forward through time, and the likelihood that the population will go extinct under specified conditions is estimated (Shaffer 1981, Possingham et al. 1993, Morris and Doak 2002). Simulation testing of stock assessment methods has been used to assess the utility and accuracy of stock assessment models and methods for a variety of stocks (NRC 1998, Deroba et al. 2015). Efforts have been made to provide widely available evaluation frameworks and software for evaluation of stock assessment methods and management strategies (Kell et al. 2007). Management strategy evaluation has emerged as the preferential method for testing management strategies as a whole, which includes stock assessment methods (Punt et al. 2014).

Two approaches to simulation testing of catch curve models are used in this thesis. The first approach is deemed a statistical criteria evaluation (SCE) and the second approach is a management strategy evaluation (MSE). There are several key differences between the approaches. Most importantly, the first approach is an evaluation of an assessment method, the catch curve model, while the second approach is an evaluation of a fisheries management system, which consists of an assessment method as well as a harvest control rule (Table 1.1). Second, the SCE is an open loop simulation, in which data are simulated, catch curves are fit to the data, and results are evaluated with respect to the accuracy of the models' estimates. In contrast, the MSE simulations are a closed loop simulation; a feedback mechanism alters the future status of the population based on assessment results (Table 1.1). Because of the feedback loop in the MSE, the exploitation rate varies from year to year and is changed according to management actions, whereas the exploitation rate is held constant in the SCE (Table 1.1).

Table 1.1 The different characteristics present or absent in the two simulation approaches, organized by model component of the simulation testing. The first approach is the statistical criteria evaluation (SCE) and the second approach is the management strategy evaluation (MSE). Differences between the approaches are bolded.

Model Component	Characteristic	SCE	MSE
Population Simulation	A population is projected forward through time with a life-cycle-based population dynamics model	Yes	Yes
Population Simulation	Random variability in maturation schedules	Yes	Yes
Population Simulation	Random variability in annual recruitment	Yes	Yes
Population Simulation	**Random variability in the exploitation rate**	**No**	**Yes**
Population Simulation	**Exploitation rates allowed to vary according to management actions**	**No**	**Yes**
Population Simulation	Different starting conditions are used in the projections (different exploitation rates, levels of variability in recruitment)	Yes	Yes
Monitoring	The age structure is randomly sampled	Yes	Yes
Assessment	**Multiple catch curve models are fit to the age data collected in each year and simulation**	**Yes**	**No**
Assessment	**A single catch curve model is periodically fit to the age data collected in each year and simulation, depending on the assessment schedule**	**No**	**Yes**
Management	**Estimates of Z are compared to biological reference points to determine stock status**	**No**	**Yes**
Management	**The fishing mortality rate for subsequent years in the projection is modified based on stock status and the harvest control rule**	**No**	**Yes**
Evaluation	**Estimates of Z are compared to the true value to gauge accuracy of the estimates produced by the catch curves**	**Yes**	**No**
Evaluation	**Management strategies are evaluated based on population metrics of escapement, landings, and exploitation rate relative to reference points**	**No**	**Yes**

CHAPTER 2: Statistical criteria evaluation of catch curve models for river herring

2.1 Introduction

Stock assessment models have been used by fisheries scientists for decades to estimate stock parameters to help inform management of those stocks. One of the oldest stock assessment models, catch curve analysis, uses age composition data from a single year to estimate the instantaneous total mortality rate of the stock (fishing plus natural mortality; Ricker 1975). When using catch curve models to estimate the total mortality rate, it is typically assumed that (i) fishing and natural mortality rates are constant over time and across age classes, (ii) mortality rates are not selective, (iii) recruitment is constant over time, (iv) the population is closed, (v) and that sampling of the age structure is unbiased. Generally speaking, a catch curve is produced when a log-linear relationship is established between numbers of fish and age, where numbers of fish decline as age increases. The rate of decline in numbers of fish represents the instantaneous total mortality rate. Closed-form solutions (Heincke 1913, Chapman and Robson 1960) or regression-based estimators (Ricker 1975) can be used in catch curve analysis to estimate the mortality rate.

2.1.1 Simulation testing of catch curve models

Simulation testing has been previously used to evaluate the performance of catch curve analysis. Tests generally focused on comparing mortality rate estimators and investigating the effect of relaxing some of the assumptions made for catch curve analyses. Jensen (1996) and Murphy (1997) compared the Chapman Robson estimator to a ratio estimator and a least-squares estimator, respectively. Dunn et al. (2002), Smith et al. (2012) and Millar (2014) compared several different catch curve methods, including regression-based methods. Strategies such as truncating older age classes with low numbers of fish (Dunn et al. 2002) or adding or removing partially recruited age classes (Smith et al. 2012) were tested as well. Millar (2014) proposed using a random intercept Poisson log-linear mixed model as a regression method, which he found to be superior to other catch curve methods. Allen (1997) and Thorson and Prager (2011) investigated the effect that variation in recruitment and selective mortality rates have on estimated mortality rates. The recommendations of previous simulation testing work (Dunn et al. 2002, Smith et al. 2012) have been used to select appropriate catch curve models for Striped Bass (Rachels and Ricks 2018) and herring populations (Mikkelsen et al. 2016).

Quinn and Deriso (1999) recommended avoiding catch curve analysis if a suitable alternative method of stock assessment is available. However, the data requirements to fit a catch curve are relatively small compared to the data required for alternative methods, providing an argument for their use in data poor situations. The Atlantic States Marine Fisheries Commission (ASMFC) recognizes the recommendation of Quinn and Deriso in their 2007 Shad Assessment Report, but also accepts the use of regression-based catch curve analysis in their assessment based on the results' robustness to the catch curve assumptions (ASMFC 2007). The ASMFC used the Chapman Robson method of catch curve analysis in their 2012 assessment of river herring. The ASMFC did recommend the use of more complex stock assessment models for future assessments, such as a VPA (Chaput et al. 2001) or a SCA model (Gibson and Myers 2003a), which have been used to assess Alewife populations in Nova Scotia, Canada. Despite the shortcomings of catch curve models, their few data requirements and widespread use make them an easily implemented and comparable short-term stock assessment model.

2.1.2 Objective of study

In this study I use a simulation model to test catch curve methods. The simulation model is parameterized for Alewife, an anadromous, iteroparous fish species of the genus *Alosa*, indigenous to rivers in eastern North America (Loesch 1987). Ages of *Alosa* can be determined by counting the growth rings on their scales. When ages of *Alosa* are determined by reading scales, additional information pertaining to the number of times the fish has previously spawned can also be obtained (Cating 1953, Marcy 1969).

In this study, I attempt to take advantage of the additional life history information of Alewife (that both age and spawning history can be recorded from scales) to improve regression-based catch curve models. By using number-of-previous-spawnings, rather than age class as the predictor variable in the regressions and factoring the data by age-at-maturity, I can better account for variations in life history. A group of fish that matured at the same age and spawned the same number of times will have experienced more similar annual mortality rates than a group of fish that only have age in common. By more accurately modelling the life history of groups of fish, I hope to better describe the total instantaneous mortality rate. In addition to more accurately describing the life history of the total population, factoring number-of-previous-spawnings by age-at-maturity allows us to use the partially mature age classes that are removed in standard catch curve analysis where age class is the predictor variable. In standard catch curve analysis, age classes in which only part of the age class is mature and available cannot be used, since the decrease in mature adults in the subsequent age class in the following year is confounded with newly mature individuals entering the spawning population. By describing the population in its maturity components and measuring time as number-of-previous-spawnings rather than age classes, numbers of fish from a partially mature age class can still be used. Using all available data should improve model estimates of the total instantaneous mortality rate.

The objective of this study is to use simulated data to evaluate and compare catch curve models and statistical methods for fitting those models for river herring over a range of exploitation rates, levels of variability, and sample sizes. Additionally, each catch curve model will be tested with and without previous spawning history to determine if factoring age data by age-at-maturity improves mortality rate estimates.

2.2 Methods

This analysis consisted of three steps. First, a population dynamics model was used to simulate a time series of age-structured abundance data. Twelve scenarios were simulated, using combinations of three levels of life history variability and four exploitation rates. Second, I subsampled the data at six sample sizes and fit catch curve models to the resulting data using eight methods. Finally, to evaluate how models performed relative to each other and across different simulated scenarios, I assessed the performance of the eight methods using their success rate and the bias and spread of their estimates of Z to evaluate how models performed relative to each other and across different simulated scenarios. All analyses and modeling for this project were completed in the statistical software language R (R Core Team, 2018).

2.2.1 Population simulation

The age-structured abundance data were simulated using a life-cycle based, sex-aggregated, age-structured population model (Gibson and Myers 2003a). The core of the model is a three-dimensional array that tracks the number of fish returning to the river in year t, age a, that have spawned p times previously, termed $N_{t,a,p}$. Life history variability is introduced in two ways: as annual variability in the stock-recruitment (SR) relationship, and as variability in the probability of maturation-at-age within each cohort.

Following Gibson and Myers (2003a), I defined the age of recruitment to be age-3 and modelled the SR relationship using a Beverton-Holt model. In a meta-analysis of Alewife stock-recruitment data, Gibson and Myers (2003b) found that the Beverton-Holt model provides a better fit to Alewife data than the Ricker model, and that data did not support estimation of a third stock recruitment parameter. The relationship between spawning stock biomass (SSB) in year t and number of recruits (R) in year $t+3$ is:

$$R_{t+3} = \frac{\alpha SSB_t}{1+\frac{\alpha SSB_t}{R_{asy}}} \exp\left(\varepsilon_t \sigma - \frac{\sigma^2}{2}\right),$$

where $\varepsilon_t \sim N(0,1)$. Here, the Beverton-Holt SR model is parameterized in terms of the slope of the function at the origin, α, and the asymptotic recruitment level, R_{asy}. Annual variability about the relationship is incorporated using a random deviate, ε_t drawn from a standard normal distribution and the recruitment variance parameter σ.

Recruits produce first-time spawning adults, $N_{t,a,0}$ as a function of the number of recruits t-a+3 years earlier, the probability that a fish in year t at age a matures at that age, $m_{t,a}$, and the juvenile mortality rate, M^{juv}:

$$N_{t,a,0} = R_{t-a+3} m_{t,a} e^{-M^{juv}(a-3)}.$$

Here the second source of life history variability is introduced by applying annual deviates in the calculation of $m_{t,a}$:

$$\text{logit}(m_a) = \log\left(\frac{m_a}{1-m_a}\right), \text{ and}$$

$$m_{t,a} = \frac{\exp(\text{logit}(m_a)+\varepsilon_{t,a})}{1+\exp(\text{logit}(m_a)+\varepsilon_{t,a})},$$

where m_a is the mean probability that an immature fish of given age matures at that age, and $\varepsilon_{t,a} \sim N(0, \text{var}(\text{logit}(m_a)))$.

In this analysis, Alewives are captured as they ascend the river to spawn. Therefore, the fishery only captures mature fish, and, in this analysis, is assumed to harvest the mature fish non-selectively. To project number of adults forward through time, escapement (E) from the fishery is calculated for each year, age and previous spawning class as:

$$E_{t,a,p} = N_{t,a,p}(1-\mu_t),$$

where μ_t is the exploitation rate in year t. Numbers of adults in the subsequent year, age class, and previous spawning class is calculated as:

$$N_{t+1,a+1,p+1} = E_{t,a,p} e^{-M^{adult}},$$

where M^{adult} is the instantaneous natural mortality rate for adult fish.

The life cycle is closed by calculating the SSB in year t in order to convert number of fish into the biomass to be used in the Beverton-Holt SR model:

$$SSB_t = \sum_a \sum_p E_{t,a,p} w_a,$$

where w_a are the weights-at-age.

I used life history parameter values from the Margaree River Alewife population (Table 2.1) as estimated by Gibson and Myers (2003a). The model was initialized by filling in the first three years with numbers of recruits calculated from the SSB at maximum sustainable yield for the Margaree River Alewife population. The population was then projected forward 100 years. The first 25 years of each 100-year projection were omitted from the evaluation of the catch curve models. Only years 25 through 100, after population abundance and age structure had reached its equilibriums, were used.

2.2.2 Simulation scenarios

I evaluated catch curve model performance under 72 scenarios reflective of four different exploitation rates, three different levels of variability in dynamics, and six different numbers of fish sampled to estimate the age structure.

Exploitation rates of 0, 0.25, 0.5, and 0.75 were chosen to provide a range of plausible rates. Low, medium and high levels of variability in dynamics were used. Recruitment and age-at-maturity variability, as estimated by Gibson and Myers (2003a) for the Margaree River population, includes both estimation and process error. Their values were used for the high variability scenario. The low level of variability in dynamics was chosen to be effectively zero. The medium variability level is half that of the high variability level (Table 2.1).

For each of the 12 combinations of the exploitation rates and variability in dynamics scenarios, the population was projected forward 100 years; different random deviates were drawn for each year t. Each 100-year projection was repeated 50 times, using unique random deviates in each repetition. A total of 3,750 age-structured datasets for each of the 12 scenarios were used to evaluate the catch curve models. For each year of age-structure data, six different sampling levels were used for fitting the catch curve models. Models were fit to numbers-at-age of the entire population, and random samples of 1,000, 500, 300, 200, and 100 fish.

2.2.3 Catch curve models

In order to compare methods of fitting catch curve models, I used four statistical methods to fit two different catch curve models. In the most general sense, the first catch curve model is:

$$N_a = N_0 e^{-Za},$$

where N_a is the number of fish at age a, N_0 is the number of age-0 fish, and Z is the total instantaneous mortality rate. Numbers of fish decline exponentially as age increases; e^{-Za} represents the portion of the initial population remaining in each subsequent age class a. The log-transformed version establishes a linear relationship between numbers-of-fish and age, and can be written as:

$$\log(N_a) = \log(N_0) - Za.$$

Here, the slope and intercept of the linear relationship are Z and $\log(N_0)$, respectively (Figure 2.1, left panel).

A couple of issues are evident when applying this model to river herring. First, survival to a specific age will differ depending on whether or not a fish has matured and has been subjected to in-river fisheries, passage at dams, and higher natural mortality associated with spawning. Second, in the majority of cases fish are sampled as they ascend rivers to spawn, so numbers-at-age are reflective of only the mature component of the population. In the simulation model, Alewives can mature anywhere from age-2 to age-6, with the majority of fish maturing at age-3 and age-4. Consistent with standard practice (Chapman and Robson 1960, Ricker 1975), to avoid the bias associated with fitting the model to partially recruited age classes, only fish age-5 and older were used to estimate Z with this model (Figure 2.1, left panel).

To address the issues in the first formulation of the catch curve model, and taking advantage of the information about previous spawning history available from Alewife scales (Cating 1953, Marcy 1969), I used a second catch curve model, in which the data are first factored by the age-at-maturity and then regressed using number-of-previous-spawnings as the independent variable (Gibson et al. 2017):

$N_{\tau,p} = N_{\tau,0}e^{-Zp}$, and

$\log(N_{\tau,p}) = \log(N_{\tau,0}) - Zp.$

Here, τ is the age-at-maturity, p is the number-of-previous-spawnings, and all other variables are as in the first catch curve model. An advantage to incorporating previous spawning history into the catch curve model is that none of the partially recruited age classes need to be removed (Figure 2.1, right panel). Since the numbers of fish are factored into several sub-cohorts by age-at-maturity, new recruits can only enter the first previous spawning class in each age-at-maturity. Any change in numbers of fish over time (where time is measured by number-of-previous-spawnings) in each age-at-maturity cohort is assumed to be caused solely by the instantaneous total mortality rate for this catch curve model.

Four statistical regression methods were used to estimate the instantaneous total mortality rate via the two catch curve models described above: a log-linear least-squares regression model (OLS), a Poisson generalized linear regression model (GLM), a Poisson robust generalized linear regression model (Robust GLM), and a Poisson generalized linear mixed effects model (GLMM).

As previously mentioned, numbers of fish decline exponentially with increasing age (or number-of-previous-spawnings); log-transforming numbers-at-age forces a linear relationship with the age classes, allowing an ordinary least-squares model to be fit to the data (Ricker 1975). The OLS model estimates Z as the negative of the slope that describes the relationship between log-transformed numbers of fish and age classes. A minimum variance unbiased estimate of the slope is produced by minimizing the sum of squared residuals, where residuals are calculated as:

$$r_a = \log(N_a) - \log(N_0) - Za\ ,$$

or

$$r_{\tau,p} = \log(N_{\tau,p}) - \log(N_{\tau,0}) - Zp.$$

This statistical method was first used by Ricker (1975) and is used for assessing a variety of stocks (Huo et al. 2014, Duan et al. 2016, Dadswell et al. 2017, Roux et al. 2019) including shad stocks assessed by the ASMFC (2007). This method has been demonstrated to be more negatively biased than other catch curve methods by Smith et al. (2012).

The Poisson GLM provides estimates of Z identical to those produced by the Chapman-Robson maximum likelihood estimator (Chapman and Robson 1960, Millar 2014). The Chapman-Robson method has been regularly used for assessing river herring stocks by the ASMFC (2012). Rather than log-transforming numbers of fish as in the OLS model, the dependent variable is described with a Poisson distribution and a log-link function is used in modeling the relationship between numbers of fish and age classes or number-of-previous-spawnings, depending on the catch curve model (McCullagh and Nelder 1989). This statistical method removes any bias introduced by log-transforming the data. Model parameters are estimated by finding the set of parameters that results in the maximum likelihood that the data are observed. Residuals for the model are calculated in log-space, identical to the OLS model.

The presence of outliers in a dataset, especially a small dataset, can have a great effect on the estimation of model parameters by regression. In robust regression, outliers are accounted for by down-weighting observations according to a piecewise function of their residuals (Cantoni and Ronchetti 2001). The piecewise function will weight observations differently if their residuals are greater or less than the tuning constant c. Model parameters are estimated using iteratively reweighted least-squares, where model parameters are repeatedly calculated with different weighted observations until the parameters converge. I used a Mallows quasi-likelihood estimator to estimate model parameters with the Robust GLM and modified the tuning constant c after finding the default value (tcc=1.345) down-weighted the majority of the observations. A tcc value of 15 forced the model to incorporate most of the data, while still down-weighting outliers.

The Poisson GLMM specifically accounts for some of the variability in cohort size that results when a synthetic cohort is used, by modeling the initial year class size as a random effect thereby allowing the intercept to vary. This statistical method was first introduced by Millar (2014) for the catch curve model with age as the independent variable. That model can be written as

$N_a = N_a e^{-Za} + e^{\epsilon_{t,a}}$.

Here, $\epsilon_{t,a}$ is the random effect associated with year class. When the number-of-previous-spawnings is used as the independent variable, random effects associated with initial year class size are nested within the age-at-maturity:

$N_{\tau,p} = N_{\tau,0} e^{-Zp} + e^{\epsilon_{t,\tau,p}}$.

Here, $\epsilon_{t,\tau,p}$ is the random effect associated with year class factored by age-at-maturity.

The random effect in the models helps deal with the violation of the assumption that recruitment is constant in each year, which is generally not true. The recruitment deviate and the maturity deviate are estimated as a single variance value for $\epsilon_{t,a}$, and a single variance value for each age-at-maturity for $\epsilon_{t,\tau,p}$.

A correction to the data had to be made for the GLMM with number-of-previous-spawnings as the independent variable. In order to estimate the random effects, each age-at-maturity had to have the same number-of-previous-spawnings, which is not the case in the data generated by the simulation model which specifies a maximum age of age-9. In the simulation model, the two synthetic cohorts where fish matured at age-2 and age-3 have a value for number of fish for each number-of-previous-spawnings 0 through 6; however, a fish that matured at age-4 and spawned 6 times previously would be age-10 and cannot exist in the simulation model. This issue was corrected by assigning a value of 0 to any combinations of age-at-maturity and number-of-previous-spawnings that exceeded the maximum age, so that each age-at-maturity would have a value for number of fish for all previous spawnings 0 through 6.

2.2.4 Performance metrics

To gauge the reliability of the catch curve methods, unsuccessful estimates of Z were defined and quantified. Unsuccessful estimates of Z were considered model fits that failed to produce a slope estimate. Samples of the age structure that produced fewer than three non-zero observations were not used and considered a failed fit in the model evaluations. Unsuccessful estimates were enumerated for each scenario and sample size to assess model reliability. All estimates of Z that were less than 0 were omitted from the evaluation, since they are biologically impossible and would not be used in a stock assessment; however, they were not considered an unsuccessful estimate. Estimates of Z that were greater than 10 were treated similarly.

Bias and spread of each model's estimates of Z were measured for each scenario. Distributions of individual estimates of Z (denoted $\hat{Z}$) from each model tended to be asymmetric, so non-parametric metrics were used to evaluate model performance. Smith et al. (2012) defined percent bias as

$$\%Bias(\hat{Z}) = \left(\frac{E(\hat{Z})-Z}{Z}\right)100,$$

where $E(\hat{Z})$ is the expected value of $\hat{Z}$ found by averaging over model results. To convert the calculation of percent bias to a non-parametric metric, I replaced $E(\hat{Z})$ with the median estimate of of Z. Percent bias is then calculated as

$$\%Bias(\hat{Z}) = \left(\frac{Median(\hat{Z})-Z}{Z}\right)100,$$

where $\%Bias(\hat{Z})$ is calculated for each of the eight models for all scenarios. Spread of estimates of Z was calculated as the range of the central 80% of the estimates of Z, calculated by taking the difference of the 10^{th} and 90^{th} percentiles of the distribution of the estimates of Z for each model, and for all scenarios.

2.3 Results

The simulations undertaken here provide a basis for evaluating the performance of four statistical methods for fitting catch curves to two data types that can be collected for river herring under various scenarios of mortality rates, sample sizes, and life history variability. As shown below, no single model performed well under all scenarios, although models that included age-at-maturity and previous spawning history markedly outperformed models that used age as the independent variable.

Catch curve models could not be fit to all simulated datasets and at times produced estimates that were biologically implausible or not even possible. Models fit to data that incorporate previous spawning history information failed to estimate Z much less frequently than models fit to data without previous spawning history information incorporated. Across all conditions, models fit to data with previous spawning history information failed at a rate of 1-2% while their counterparts failed at rates of 13-17% (Table 2.2). Estimates of Z could not be obtained using the Poisson GLMM fit to data without previous spawning history information, when the age structure of the entire population was used at exploitation rates of 0, 0.25, and 0.5, for the low variability scenario (see footnote Table 2.2), although this scenario is unlikely to be encountered in real-world sampling. The Poisson GLM fit to data without previous spawning history information produced the greatest proportion of extreme estimates (Z>10), although at a low rate (Table 2.2). Models fit to data with previous spawning history information produced fewer biologically impossible estimates (Z<0) than models fit to data without previous spawning history information (Table 2.2).

The proportion of simulated datasets for which plausible estimates of Z could be obtained varied among exploitation rate, sample size, and variability scenarios. As shown for the medium variability scenario, most of the failed estimates occurred at low sample size and high exploitation rate (Figure 2.2). This pattern was exacerbated as the simulated variability increased (results not shown). At high exploitation rates and low sample size, inclusion of information about previous spawning history produced a greater portion of plausible estimates of Z than were obtained when age was used as the independent variable (Figure 2.2), a pattern that also held under the high variability scenario (results not shown).

As shown for the medium variability scenario, when age was used as the independent variable all models performed similarly at low exploitation rates, even at low sample sizes (Figure 2.3). As exploitation rates increased and sample sizes decreased, estimates of Z became more negatively biased and more variable (Figure 2.3). For a given set of conditions, the distributions of estimates of Z were asymmetrical with a greater number of outliers towards a higher Z (Figure 2.3). At low exploitation rates, sample size had less of an effect on model performance, especially for the Poisson GLM and the Poisson GLMM (Figure 2.3). At high exploitation rate and low sample size, the true value of Z was not encompassed by the interquartile spread (Figure 2.3). This pattern was least prevalent for the Poisson GLM and GLMM models. These patterns were exacerbated in the high variability scenarios, and less evident in the low variability scenarios (boxplots not shown).

Models in which age-at-maturity and number-of-previous-spawnings were the independent variables outperformed the models in which age was the independent variable. The variability in the model estimates of Z was markedly reduced, and for the Poisson GLM, Robust GLM, and Poisson GLMM models the true value of Z was encompassed in the interquartile spread for all medium variability scenarios (Figure 2.4). The pattern of increased bias and increased variability in the estimates of Z was also evident for these models, but to a lesser degree than when age was used as the independent variable. When age-at-maturity and number-of-previous-spawnings were the independent variables, this pattern was also exacerbated in the high variability scenarios (boxplots not shown).

With respect to bias, not all models in which age-at-maturity and number-of-previous-spawnings were incorporated outperformed models in which age was the independent variable (Figure 2.5). As the exploitation rate and the recruitment and maturity variability scenario increased, and sample size decreased, estimates of Z became more negatively biased for all models (Figure 2.5). At an exploitation rate of 0, some models' estimates of Z became positively biased as the variability scenario increased, especially the Poisson GLMM fit to data with previous spawning history information incorporated (Figure 2.5, far left column). At high exploitation rates, the OLS and Robust GLM fit to data without spawning history information incorporated were the most negatively biased. Negative bias of estimates of Z produced by the OLS and Robust GLM increased more rapidly than the negative bias of estimates of Z produced by the Poisson GLM and GLMM models (Figure 2.5, right columns). In most scenarios, for a given estimation method, models that incorporated previous spawning information were less biased than models with age as the independent variable. An exception occurred in the Poisson GLMM at low exploitation rates in which incorporation of the previous spawning history increased the bias (Figure 2.5, left column). However, the Poisson GLMM with previous spawning history information was the least biased model at all sample sizes at high exploitation rates in the high variability scenario (Figure 2.5, lower right panel).

All models produced the least variable estimates of Z in the low exploitation rate, low variability and large sample size scenarios (Figure 2.6, panel in left column and top row). For all scenarios, models fit to data incorporating spawning history information produce estimates of Z that were less variable than the same models in which age is the independent variable (Figure 2.6). In the low variability scenario, estimates of Z became more variable with increasing exploitation rate and decreasing sample size (Figure 2.6, top row). In the medium and high variability scenario, estimates of Z became more variable with increasing exploitation rate. In the high exploitation rate high variability scenarios, the Poisson GLM and Robust GLM models with age as the independent variable showed a pattern of decreasing variability in the estimates of Z as sample size decreased (Figure 2.6, lower right panel). This counter-intuitive result likely occurred because a greater proportion of the simulations were discarded because they did not produce plausible estimates of Z. Incorporation of previous spawning history in the model alleviated this pattern (Figure 2.6, lower right panel). In the medium and high variability scenario, the OLS with spawning history information incorporated into the data provided the least variable estimates of Z, while the Poisson GLM and Robust GLM fit to data without spawning history information provided the most variable estimates of Z.

Selection of a method for fitting catch curves involves an evaluation of the trade-offs between the bias and the variability of the resulting estimates, since no single method performed best under all scenarios (Figure 2.7). Even under certain individual scenarios, models performed identically (Figure 2.7, top row) or no model was clearly superior to the others. At an exploitation rate of 0.25, in a high variability scenario, and at a sample size of 500, estimates of Z produced by the GLM and GLMM with numbers-of-previous-spawnings as the independent variable had nearly identical spreads and magnitudes bias; their estimates differed in the direction of their bias, the GLM has a negative bias, while the GLMM has a positive bias (Figure 2.7, second column on the left, bottom row). Additionally, under the same conditions the GLMM with age as the independent variable produced estimates of Z with approximately 50% greater spread but considerably less bias than the estimates of Z produced by the two previously mentioned models. In order to answer the question of which model performed the best, additional metrics that can answer whether positive or negative bias provide better assessment of the stock, and can evaluate the trade-off between spread and bias of estimates of Z, are needed.

2.4 Discussion

The main conclusion of this study is that using a catch curve model that incorporates previous spawning history information provides better estimates of Z than traditional catch curve methods that use age as the independent variable. I also found that no single statistical method of fitting catch curve models provides the best estimates of Z under all simulated levels of variability in life history, exploitation rates, and sample size.

Using the selected methods, the results suggest that using the catch curve model with number-of-previous-spawnings as the independent variable provided superior estimates of Z compared to using age as the independent variable, for each of the four statistical methods for fitting catch curves. Each regression method used to estimate Z produced considerably more precise estimates of Z when previous spawning history information was incorporated into the catch curve model. Incorporating previous spawning history into the regressions uses all available data, by differentiating the synthetic cohort within a year into sub-cohorts based on age-at-maturity. Using the partially mature age classes provided an advantage especially under high exploitation rates, where fewer fish live to the fully mature age classes (age-5 and above). Regression methods without previous spawning history information had high failure rates at low sample size, high exploitation, and high variability in life history, up to 20 or 40% in some cases. Incorporating previous spawning history into the regression methods lowered the failure rate to <1% in those cases.

No single catch curve model performed best under all simulated conditions, but the Poisson GLM and the GLMM did outperform the Robust GLM and the OLS. The OLS was the most negatively biased method for fitting the two catch curve models, although it produced very precise estimates of Z. The negative bias of the OLS was also observed by Smith et al. (2012). The Robust GLM was slightly better than the OLS in some cases but was outperformed by the Poisson GLM and the GLMM. The results I obtained for the Poisson GLM and the GLMM with age as their independent variables were similar to the results obtained by Millar (2014), despite the differences between data simulation methods. In this study, numbers-at-age were simulated from a short lived population with a maximum age of 9, while Millar's population had a maximum age of 200; additionally, I looked at a wider range of Z values (from 0.44 to 1.83) at a lower resolution (four different rates) compared to Millar who looked at Z values ranging from 0 to 1.0, at a greater resolution (10 different rates).

This work builds upon previous simulation testing of catch curve models (Smith et al. 2012, Millar 2014) by introducing and evaluating catch curve models that incorporate previous spawning history information, a method introduced by Gibson et al. (2017). I found the Poisson GLM and the GLMM with number-of-previous-spawnings as the independent variable produced estimates of Z that are the least biased and the least variable. Overall, I found the results to closely match those of Smith et al. and Millar, despite different methods of simulating the data for fitting the catch curve models.

I found the Poisson GLM to be superior to the GLMM at low exploitation rates, but inferior at high exploitation rates. The GLMM tended to be positively biased (over-estimated Z) at low exploitation rates, but to be accurate at high exploitation rates; while the Poisson GLM was accurate at low exploitation rates and negatively biased at high exploitation rates. This dichotomy presents a tradeoff between assigning a stock as fully-exploited when it is overexploited or assigning a stock as fully-exploited when it is under exploited. While neither situation is desirable, assigning a stock as fully-exploited when it is under exploited (more likely with the GLMM) adheres to Fisheries and Oceans Canada's precautionary approach (DFO 2006), whereas the alternative does not.

As stated in the results, the spread of estimates of Z tended to increase with increasing exploitation rate, increasing variability in life history, and decreasing sample size. However, at medium and high levels of variability in life history and at an exploitation rate of 0.75, the spread of estimates of Z tended to decrease as the sample size neared 100. This exception to the previously stated trend was most notable in the OLS and Robust GLM with age as the independent variable. It is possible the decrease in spread occurs since there are fewer possible sample combinations at a high exploitation rate and small sample size; the majority of the samples suitable for fitting models will have a very similar structure (such as 95 age-5 fish, four age-6 fish, and one fish age-7, -8, or -9). The decreased variability in age structure at small sample sizes would translate to a decreased variability in estimates of Z. I did not find any corresponding pattern in the bias of estimates of Z; bias never decreased with decreasing sample size at medium and high levels of variability in life history and at an exploitation rate of 0.75.

The method used to subsample the age structure can affect models' estimates of Z. In this study I used true random sampling, where the probability of including a fish of age-*a* in the sample is equal to the proportion of those fish age-*a* in the whole population. In reality, sampling would be biased based on factors such as the gear used to sample, or timing of sampling relative to the spawning run. Chaput and Atkinson (2001) found that repeat spawners tend to return to the river earlier than first-time spawners, and that fish that ran earlier were larger than those that ran later. A sampling scheme where fish are sampled throughout the entirety of the run, and the number of fish sampled in a day is weighted by the number of fish that passed through the sampling location in that day (as described in Gibson et al. 2017), would be required to remove the bias associated with unweighted, true random sampling. If the results of this analysis are used to inform stock assessment methods, biased sampling will reduce the applicability of the recommendations herein.

I found the Robust GLM performed worse than the GLM and GLMM models regardless of the value I chose for the tuning constant c for weighting of the observations by the Robust GLM. The selection of the value for the tuning constant was arbitrary, and although methods exist for selecting tuning constants for the Robust GLM (Cantonni and Rochetti 2001), the disparity in model performance between the Robust GLM and the GLM and GLMM suggest that no tuning constant will reliably produce improved estimates of Z compared to estimates of Z produced by the GLM or GLMM. Based on the performance of the Robust GLM, using the GLM or GLMM to estimate mortality rates would be preferable to investigating optimal tuning constant values for the variable datasets I used to estimate the instantaneous total mortality rate in this study.

I decided to analyze each model's estimates of Z with non-parametric metrics of bias and spread, rather than the parametric metrics used by Smith et al. (2012) and Millar (2014). The decision to assess model performance with non-parametric metrics was based on the distributions of the estimates of Z for each model and each set of conditions; the distributions tended to be skewed or contain subsets of outliers at high exploitation rates and high variability scenarios. The decision to treat negative mortality rates as fails also affected the distributions of estimates of Z, effectively truncating a small part of the distribution. The non-parametric metrics of bias and spread described in the methods best describe the skewed and truncated distributions while providing similar results to parametric metrics for normally distributed estimates of Z.

The decision to hold the instantaneous mortality rates constant and to not explicitly include autocorrelation in recruitment variability in the simulation model should not affect the catch curve models' estimates of Z. Since each model is fit to only a single year of age data independent of the other years, varying or holding constant mortality rates or recruitment rates will not make a significant difference in the estimates of Z, aside from providing a range of conditions under which to estimate Z, which has been accomplished herein. If further analysis of stock assessment models that involves linking years of age structure data was to be undertaken, variable mortality rates and autocorrelation in recruitment and mortality rate deviates should be included in the simulations.

While I have addressed the relative performance of the catch curve models under simulated conditions, it is important to consider how appropriate these models are for assessing river herring stocks. As previously stated, catch curve models work under the assumptions that there is no variability in mortality and recruitment and that they are best suited for long lived species, while Alewife are short-lived with variable recruitment and mortality rates (Chaput et al. 2001, Gibson et al. 2017). While the life history characteristics of river herring do not match the required assumptions for catch curve analysis, catch curve models can still fulfill a role in River Herring assessment. Since the majority of river herring stocks in DFO's Maritimes region are not regularly assessed (Gibson et al. 2017), any long-term assessment approach would have to begin by collecting a year of age structure data. Catch curve models would offer an initial, if poor, estimate of the total instantaneous mortality rate, rather than no information. At the very least, catch curve models would be useful in providing a rough early assessment of a stock to bridge the gap to more accurate estimates of stock characteristics from a SCA or VPA.

The small amount of data required to fit a catch curve model makes it possible to implement a standardized wide-scale assessment method along the lines of the assessments carried out by the ASMFC. The ASMFC has used catch curve models in their shad and River Herring assessments, specifically the Chapman Robson method and log-transformed regression, which are analogous to the GLM and OLS methods with age as the independent variable (ASMFC 2007, ASMFC 2012, ASMFC 2017). Using the GLM or GLMM can help reduce the under-estimation of Z by the OLS method. Most importantly, making use of the spawning history of river herring or shad by using any of the catch curve methods with number-of-previous-spawnings as the independent variable will improve the precision of estimates of Z, and reduce the negative bias. For stocks where the instantaneous total mortality rate is known to be high (Z >1), the GLMM with number-of-previous-spawnings as the independent variable may provide the best estimates of Z.

I have addressed the relative performance of the catch curve models and established that they can have a role in a stock assessment framework for river herring, but it would be prudent to investigate the utility of the advice derived from the catch curve models in a more real-world context. A management strategy evaluation (MSE) could be used to assess the quality of the advice derived from the catch curve models. Previous work has been done on MSEs for data-limited stock assessment methods (Carruthers et al. 2014, Geromont and Butterworth 2015), including an MSE where Wayte and Klaer's improved catch curve method is used to estimate Z with a five-year time series of the age structure (Wayte and Klaer 2010). If an MSE was to be used to further evaluate the catch curve models in this paper, important questions to consider are:

- What proportion of the time does the estimate of Z and the true value of Z fall in the same exploitation rate zone (under, fully, and over exploited)?
- If harvest control rules are in place and the exploitation rate is lowered, do the catch curve models detect the change in the mortality rate? Is there a lag in detection?
- Does averaging the mortality rate over time (using a multi-year rolling average) improve estimates of Z?
- Does the over-estimation of Z at low exploitation by the GLMM perform better as a management strategy over time compared to the under-estimation of Z at low exploitation by the Poisson GLM?

In summary, I found that catch curve models that use number-of-previous-spawnings as the independent variable produced more precise estimates of Z than catch curve models that used age as the independent variable. Of the statistical methods used, the Poisson GLM and the GLMM performed best; those methods with number-of-previous-spawnings as the independent variable generally provided the most accurate and precise estimates of Z across all simulated scenarios and sample sizes.

2.5 Tables and Figures

Table 2.1. Parameter values for the age-structured, forward-projecting, population dynamics model used to simulate numbers-at-age datasets of Alewife for testing catch curve regression models. Life-history parameter values were derived by Gibson and Myers (2003a) using data from the Margaree River, NS, Alewife population. Values for indices t, a, and p are the range of values used in the model.

Term	Description	Value
	Indices	
t	Time (years)	1-100
a	Age class	2-9
p	Previous spawning class	0-6
	Life history parameters	
α	Slope at the origin of the Beverton-Holt spawner-recruit relationship	73.88
R_{asy}	Asymptotic recruitment level for the Beverton-Holt spawner-recruit relationship	6,915,954
M^{adult}	Instantaneous natural mortality rate for adults	0.44
M^{juv}	Instantaneous natural mortality rate for juveniles	0.40
m_2	Probability of maturation at age-2	<0.01
m_3	Probability of maturation at age-3	0.52
m_4	Probability of maturation at age-4	0.97
m_5	Probability of maturation at age-5	0.94
m_6	Probability of maturation at age-6	1.00
	Exploitation rate	
μ	Portion of the mature population that is harvested in each year t	0, 0.25, 0.5, 0.75
	Variance parameters	
σ	Recruitment variance; low, medium and high variability scenarios	0.001, 0.63, 1.26
$var(logit(m_2))$	Variance of maturation probability at age-2 (logistic scale); low, medium and high variability scenarios	0, 0.853, 1.76
$var(logit(m_3))$	Variance of maturation probability at age-3 (logistic scale); low, medium and high variability scenarios	0, 0.712, 0.910
$var(logit(m_4))$	Variance of maturation probability at age-4 (logistic scale); low, medium and high variability scenarios	0, 0.919, 2.44
$var(logit(m_5))$	Variance of maturation probability at age-5 (logistic scale); low, medium and high variability scenarios	0, 0.995, 5.32
$var(logit(m_6))$	Variance of maturation probability at age-6 (logistic scale); low, medium and high variability scenarios	0, 0, 0

Table 2.2. Percentages of each catch curve regression model's estimates of the instantaneous total mortality rate (Z) that failed to fit, produced extreme values ($\hat{Z}$>10 or $\hat{Z}$<-10), produced biologically impossible values ($\hat{Z}$<0), and produced values that are deemed successful estimates (the remainder).

Data Type	Model	Failed to fit	Extreme values	Biologically impossible	Success
With	OLS	0.94	0	0.63	98.4
Previous	GLM	0.94	0	0.53	98.5
Spawning	Robust GLM	0.94	0	0.84	98.2
History	GLMM	1.75	0	0.34	97.9
Without	OLS	13.5	0	2.65	83.9
Previous	GLM	13.4	0.03	2.06	84.5
Spawning	Robust GLM	13.5	0	2.93	83.6
History	GLMM	13.5*	0	1.77	80.6

*The true failure rate for the GLMM fit to data with previous spawning history information is 17.6%. The model failed to fit at a rate of 100% when data were simulated with low variability at exploitation rates of 0, 0.25, and 0.5, and the model was fit to the true age structure of the population. This subset of scenarios makes up 4.1% of all possible scenarios, the difference between 17.6% and 13.5%.

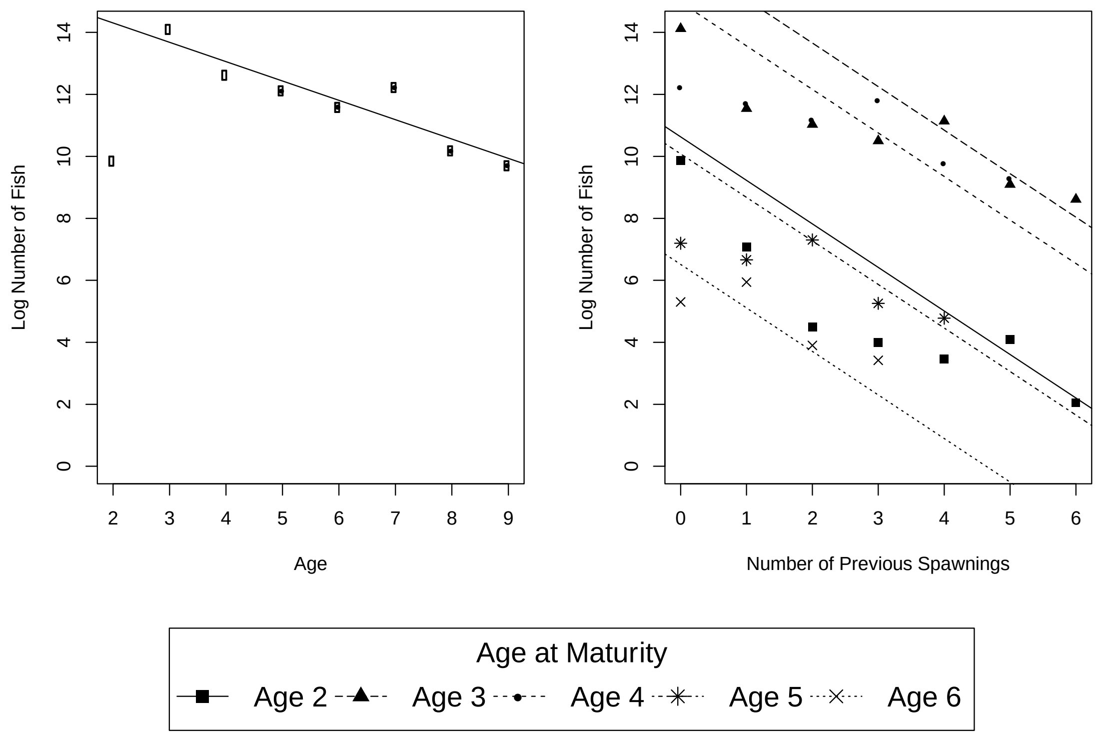

Figure 2.1. A comparison of fitting a catch curve regression model with (right panel) and without (left panel) spawning history. The Y-axes for both panels are log-number of fish. The left panel has age, ranging from age-2 to age-9 on the X-axis. Solid circles represent numbers-at-age that are used in the regression, while hollow circles are not used in the regression. The right panel has number-of-previous-spawnings ranging from 0 to 6 on the X-axis. Data are factored by age-at-maturity on the right panel, represented by different points and lines described in the legend. The same simulated data were used and are represented in each panel.

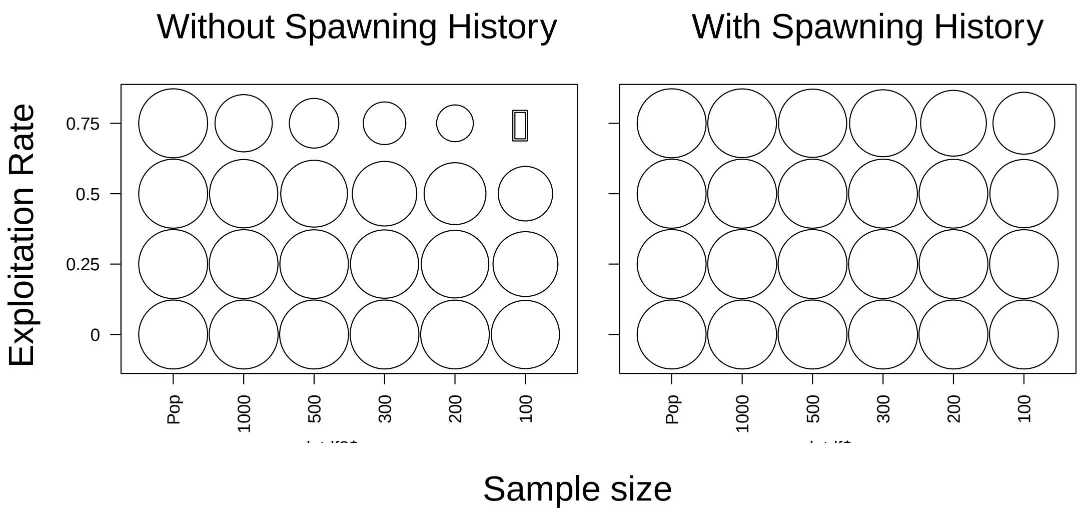

Figure 2.2. The percentage of 3,750 attempted catch curve model fits, using the Poisson generalized linear model at medium recruitment and maturity variability, for which estimates of the instantaneous total mortality rate (Z) were obtained. The area of the bubble is proportional to the percentage of fits for which estimates if Z were obtained, for each combination of exploitation rate (Y-axis) and sample size (X-axis). Bubbles that touch the edges of their neighbouring bubble indicate 100% successful model fits. The Y-axis for both panels shows the four exploitation rates used for fitting the models. The left panel depicts results of model fits using age as the independent variable, and the right panel depicts results of model fits using age-at-maturity and number-of-previous-spawnings as the independent variables.

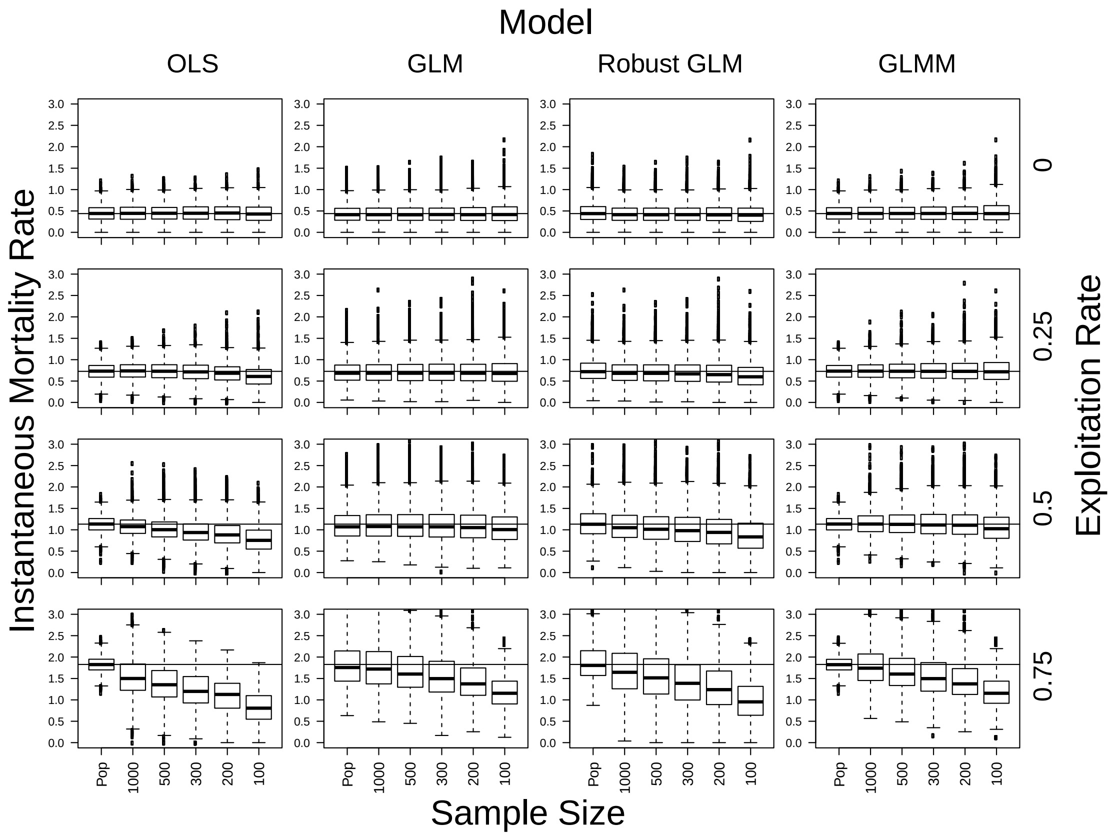

Figure 2.3. Boxplots summarizing the instantaneous total mortality rate (Z) estimates obtained using four different methods of catch curve analysis, using age as the independent variable, fit to data simulated under the medium recruitment and maturity variability scenario. Methods are: log-linear least-squares regression (far left column, OLS), Poisson generalized linear regression (center left column, GLM), robust Poisson generalized linear regression (center right column, Robust GLM), and Poisson generalized mixed effects regression (GLMM). Each row shows estimates of Z using data simulated under four exploitation rates: 0, 0.25, 0.5, 0.75, from top to bottom. Catch curve regressions were fit to simulated data using sample sizes of all fish (Pop), 1,000, 500, 300, 200 and 100 fish, depicted on the X-axis of each panel. Each individual boxplot shows the results of 3,750 attempted estimates of Z for each sample size and exploitation rate combination. The horizontal solid line represents the true instantaneous total mortality rate in each panel.

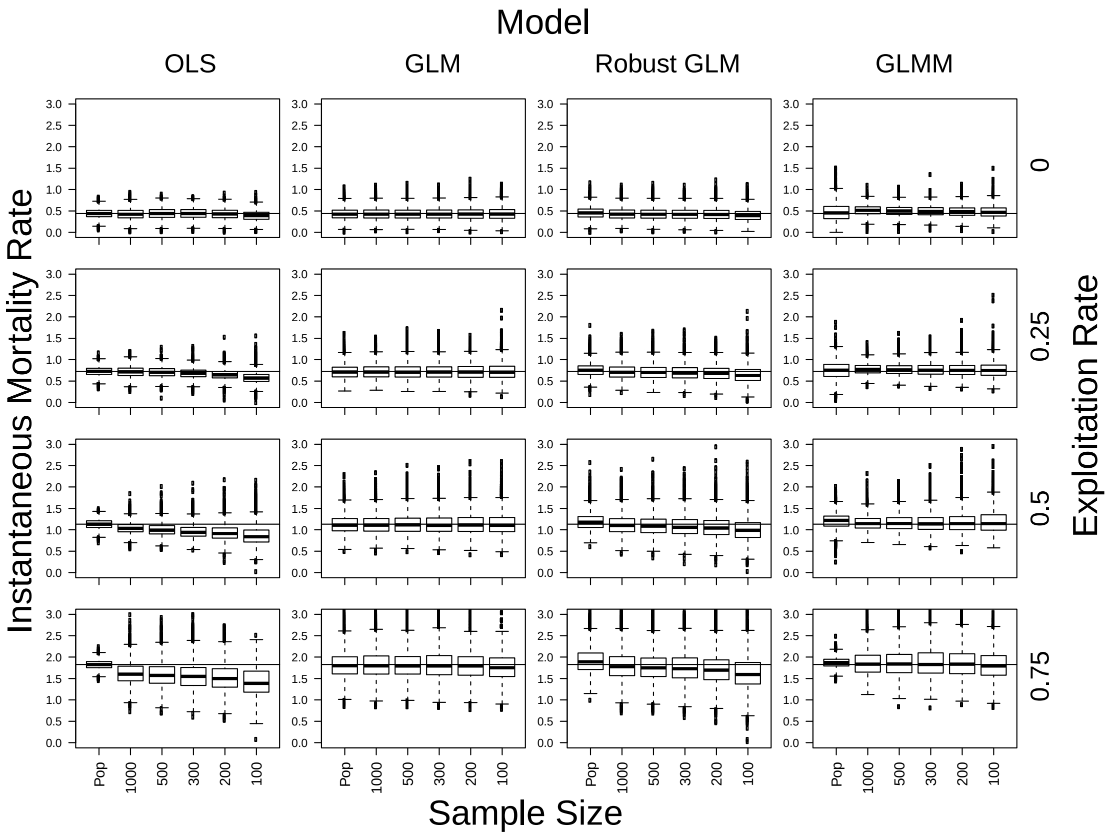

Figure 2.4. Boxplots summarizing the instantaneous total mortality rate (Z) estimates obtained using four different methods of catch curve analysis, using the number-of-previous-spawnings and age-at-maturity as the independent variables, fit to data simulated under the medium recruitment and maturity variability scenario. Methods are: log-linear least-squares regression (far left column, OLS), Poisson generalized linear regression (center left column, GLM), robust Poisson generalized linear regression (center right column, Robust GLM), and Poisson generalized mixed effects regression (GLMM). Each row shows estimates of Z using data simulated under four exploitation rates: 0, 0.25, 0.5, 0.75, from top to bottom. Catch curve regressions were fit to simulated data using sample sizes of all fish (Pop), 1,000, 500, 300, 200 and 100 fish, depicted on the X-axis of each panel. Each individual boxplot shows the results of 3,750 attempted estimates of Z for each sample size and exploitation rate combination. The horizontal solid line represents the true instantaneous total mortality rate in each panel.

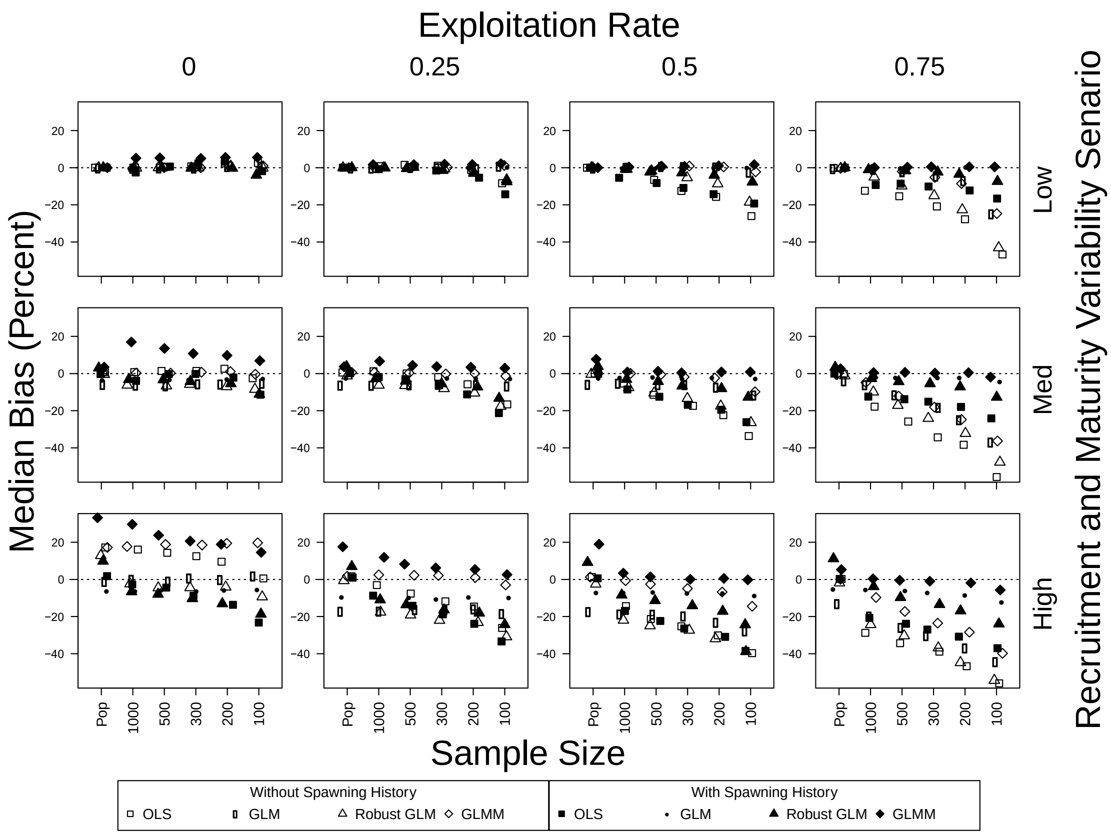

Figure 2.5. Median values of the percent bias for eight catch curve regression methods for estimating the instantaneous total mortality rate (Z) for Alewife. The methods are: log-linear least-squares models (OLS), Poisson generalized linear models (GLM), robust Poisson generalized linear models (Robust GLM), and the Poisson generalized linear mixed effects models (GLMM). Models are fit to data incorporating previous spawning history information (filled symbols), and models without previous spawning history information (open symbols). Values are shown for each combination of exploitation rate (0, 0.25, 0.5, 0.75 columns left to right), recruitment and maturity variability scenarios (low, medium, high, rows top to bottom) and sample size (all fish, 1,000, 500, 300, 200, and 100 fish, left to right across the X-axis). Values are calculated as the median of the signed percent difference, $(Z_{est} - Z_{true})/Z_{true} * 100$, from the successful model fits to 3,750 simulated datasets for each combination of sample size, exploitation rate and variability.

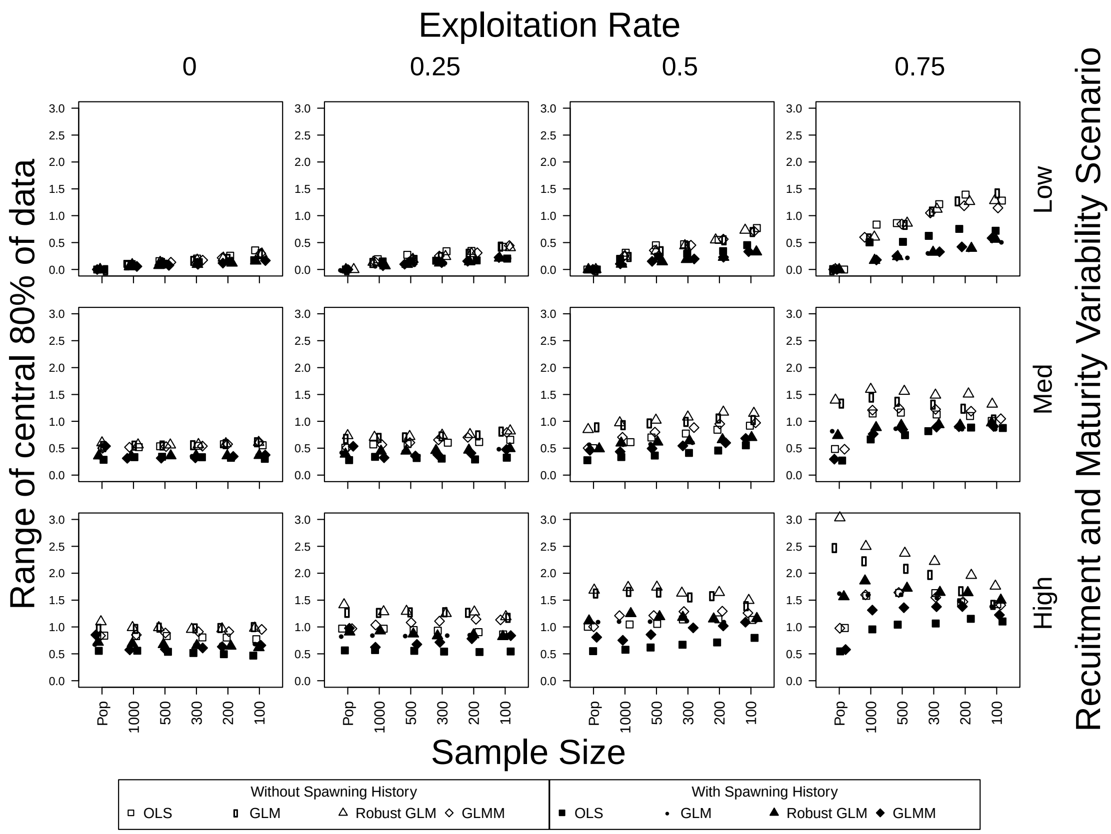

Figure 2.6. Range of the central 80% of eight catch curve regression methods for estimating the instantaneous total mortality rate (Z) for Alewife. The methods are: log-linear least-squares models (OLS), Poisson generalized linear models (GLM), robust Poisson generalized linear models (Robust GLM), and the Poisson generalized linear mixed effects models (GLMM). Models are fit to data incorporating previous spawning history information (filled symbols), and models without previous spawning history information (open symbols). Values are shown for each combination of exploitation rate (0, 0.25, 0.5, 0.75 columns left to right), recruitment and maturity variability scenarios (low, medium, high, rows top to bottom) and sample size (all fish, 1,000, 500, 300, 200, and 100 fish, left to right across the X-axis). Values are calculated as the difference of the 90th and 10th percentiles of the successful model fits to 3,750 simulated datasets for each combination of sample size, exploitation rate and variability.

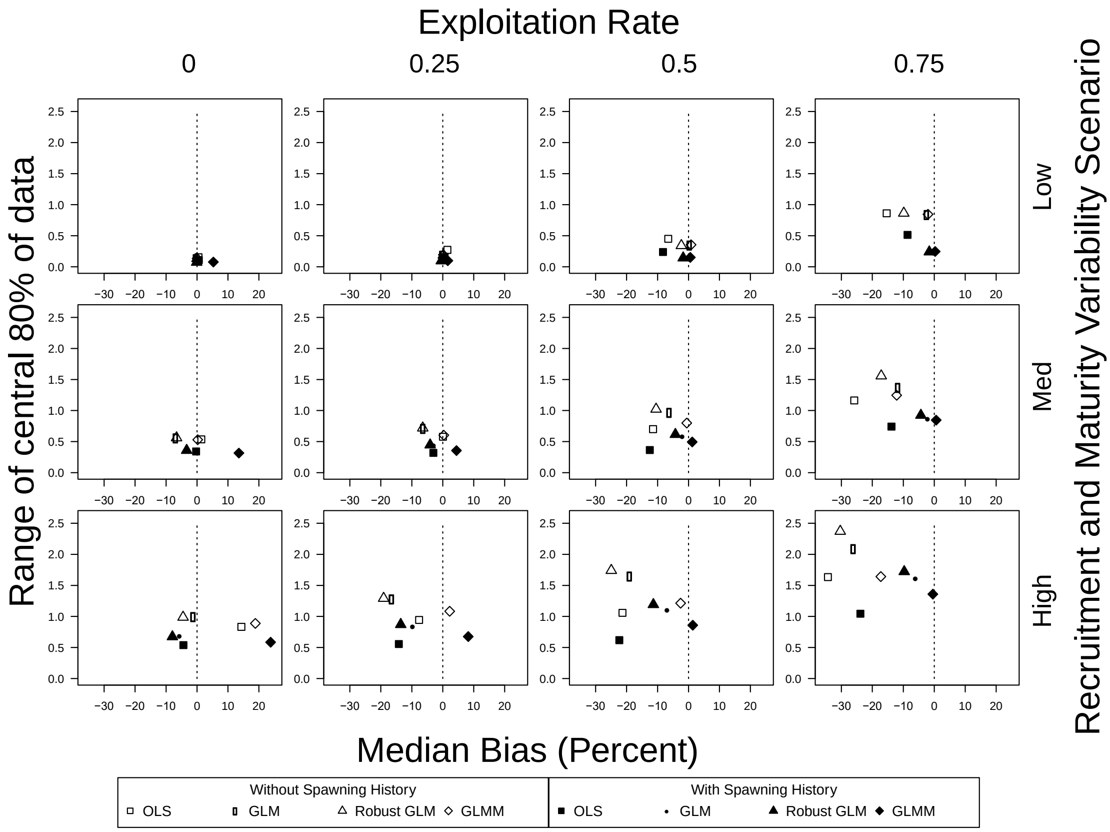

Figure 2.7. Median values of the percent bias and range of the central 80% for eight catch curve regression methods estimating the instantaneous total mortality rate (Z) for Alewife. The methods are: log-linear least-squares models (OLS), Poisson generalized linear models (GLM), robust Poisson generalized linear models (Robust GLM), and the Poisson generalized linear mixed effects models (GLMM). Models are fit to data incorporating previous spawning history information (filled symbols), and models without previous spawning history information (open symbols). Median bias (percent), the X-axis, and range of the central 80% of the data, Y-axis, (metrics described in Figures 2.5 and 2.6) are presented for each model and each combination of exploitation rate (0, 0.25, 0.5, 0.75 columns left to right), recruitment and maturity variability scenarios (low, medium, high, rows top to bottom) all for the sample size of 500 fish.

CHAPTER 3: Management strategy evaluation of catch curve models for river herring

3.1 Introduction

3.1.1 Management strategy evaluation

Management strategy evaluation (MSE) is a process in which the effect of management strategies can be tested using population simulations in which populations are subjected to different management scenarios and projected forward through time (Smith 1994). The goal of using MSE is to evaluate the consequences of management actions on a stock and to examine the trade-offs of different management actions, rather than finding an optimal solution to a management problem (Smith et al. 1999).

A MSE has two primary components: the "operating model" which simulates the stock and fishery being evaluated, and the "management strategy" which includes the assessment method and harvest control rule (HCR) for managing the stock and fishery (Punt et al. 2014). The operating model produces data which are "collected" or sampled and fed into an assessment model to produce an estimate of some stock characteristic, which is compared to a biological reference point (BRP) to determine stock status. The HCR is the method by which the stock is managed; it is the guideline that describes what can be harvested from the stock given the stock status. Management action, based on the HCR, may then be taken depending on the stock status, resulting in a change in total allowable catch (TAC), fishing effort, or some other metric for controlling the fishing mortality rate. This feedback loop between the operating model and the management strategy is a key characteristic of MSE, which separates it from traditional evaluation of assessment methods (Punt et al. 2014), such as the method of evaluation via statistical criteria used in Chapter 2.

MSE can fulfill a variety of purposes: comparing outcomes of management strategies or investigating the effect an alternative management strategy would have on a stock; developing new management strategies; determining whether certain management strategies are appropriate for the target stock and fishery; and evaluating the robustness of management strategies to various operating models based on different biological assumptions (Punt et al. 2014). The earliest use of MSE as per Punt et al. (2014), before MSE was a defined term, was by Bergh and Butterworth in 1987 on the South African anchovy. Bergh and Butterworth compared alternative management strategies and provided a basis for selecting a new harvesting strategy. MSE has been used to evaluate data-limited methods (Carruthers et al. 2014, Fulton et al. 2016) revealing several appropriate and successful management strategies for the target stocks. Wayte and Klaer (2010) presented an MSE where the estimates of recent catch and fishing mortality rate were used to determine the next year's recommended biological catch. They tested this management strategy for a short-lived and a long-lived species and found that a harvest strategy that uses only age composition and catch data can meet their management objectives. MSE has been used on short-lived species with highly variable life histories in a case study of the Bay of Biscay anchovy, where results were used to select a HCR by managers and stakeholders (Sánchez et al. 2019). The use of MSE has expanded beyond investigating the impacts of management on the target stock, to investigating the impacts on entire ecosystems (Weijerman et al. 2016, Surma et al. 2018).

3.1.2 Objective of Study

The successful use of MSE to evaluate data poor methods (Carruthers et al. 2014, Fulton et al. 2016), especially for short-lived species (Wayte and Klaer 2010), suggests that catch curve models may be a useful assessment method for river herring despite their poor performance from a statistical perspective (Chapter 2), and that MSE is a useful tool for determining the utility of catch curve models in a management context.

The objective of this study is to further evaluate the utility of assessing river herring stocks with a catch curve model through MSE, in an effort to determine whether catch curves are an appropriate assessment method for river herring. Specifically, the main objectives of this MSE are determining whether assessments via catch curves can provide advice to correct chronically over-exploited or under-exploited populations to be fully-exploited, and maintaining sustainable exploitation rates for populations. In order to fine tune the use of catch curves for assessment and management, several HCRs, as well the frequency with which assessments and management actions occur, are investigated in this MSE. Overall, the results show that despite their poor performance from a statistical performance, catch curves do provide useful information for managing river herring fisheries. The implications of using the evaluated management strategies are considered in a broader context.

3.1.3 River herring management

River herring are managed throughout their range in North America by governmental organizations in Canada and the United States (ASMFC 2012, ASMFC 2017, Gibson et al. 2017). Management of river herring stocks is typically implemented with some form of effort control, such as limiting the number of licences, limiting the gear used for fishing, limiting season length, limiting the number of hours fished within a time period, or any combination of these methods (ASMFC 2012, ASMFC 2017, Gibson et al. 2017). Some management strategies for river herring stocks in South Carolina and North Carolina have also included TAC as a management tool (ASMFC 2012). A HCR based on a young-of-the-year index was recently implemented for river herring in the Hudson River in New York (ASMFC 2017).

Setting a target escapement is a recommended management strategy for short-lived species, where fishing is only allowed after a pre-determined escapement target is reached, such that landings would be the spawning stock biomass (SSB) in excess of the escapement target (ICES 2013). However, this management approach is not easily implemented for river herring stocks. Most fisheries for adult river herring catch fish prior to spawning, such that escapement is defined as the number of fish that avoid capture by the fishery and continue upriver to the spawning grounds (Gibson et al. 2017). A fishery-independent estimate of escapement would be required, such as a count at a fish way upriver from the fishery. Few river herring stocks have fishery-independent sources of data from which escapement can be estimated (ASMFC 2012). Therefore, a management strategy that does not rely on fishery-independent data is required for most river herring stocks.

Currently, river herring stocks in DFO's Maritimes region are not regularly assessed (Gibson et al. 2017). Ideally, river herring stocks should be assessed and managed in each river, or even in each tributary for larger river systems like the Saint John River in New Brunswick (ASMFC 2012, Gibson et al. 2017). However, management sometimes takes place at a higher level, encompassing stocks from multiple rivers within a single management unit (ASMFC 2012, Gibson et al. 2017). Grouping stocks by arbitrary boundaries can lead to stocks with different life histories having the same management actions applied to them, which can result in varying outcomes across the stocks. Grouping stocks in different rivers according to common population characteristics or genetics and assessing index rivers that represent an assemblage of rivers would help address this issue, but does not rectify it (McBride et al. 2014, Gibson et al. 2017). The challenge of assessing stocks on many rivers is in part due to the lack of data that are readily available; few rivers have fish ways or monitoring facilities installed for conducting escapement counts (ASMFC 2012). The most common source of data that could be used for river-specific assessments are landings data, from which the age structure can be estimated provided regular samples are taken from the catch. An assessment method such as catch curves that can make use of this more common source of data would help overcome the challenge of assessing many river herring stocks.

3.2 Methods

The MSE undertaken here consists of several model components working in conjunction with each other. The first component is the "operating model", which consists of a population dynamics model, parameterized for the Gaspereau River Alewife population. This component projects the dynamics of the population forward in time by producing a time series of age-structured abundance data. The second component is the "management strategy", which consists of an assessment model and a HCR. The assessment model used here is a catch curve model, which produces an estimate of the instantaneous total mortality rate (Z). The estimate of Z is then compared to reference points in the HCRs to determine stock status and the management action to be taken. Management actions consist of adjustments to the exploitation rate for the subsequent years in the projection by applying a correction factor to the instantaneous fishing mortality rate. The two model components work together in a feedback loop as the population is projected through time. Within each time step an assessment occurs, stock status is determined based on the assessment results, and management action is then applied to correct the fishing mortality rate in the next time step of the model. This feedback loop continues for the duration of the population projection.

In addition to harvest control rules, different options for assessment frequency (e.g. the number of years between assessments) were evaluated by conducting assessments at pre-determined intervals (Table 3.1). The same options for management frequency were also conducted (Table 3.1). These assessment and management schedules were included in the MSE to represent how different fisheries can be assessed and managed: some fisheries may be assessed every year, and have changes made to the length of the fishing season or number of licences before the beginning of the next season, while some fisheries may only be assessed and managed every sixth year. Intermediate scenarios were considered, such as assessing every year and managing every sixth year using the average of the previous six assessments results in the HCR. A schedule where the populations are assessed every year but no management changes ever happen was evaluated as well, to represent a *status quo* approach.

All combinations of HCRs and assessment and management schedules were evaluated for three different types of populations: chronically under-exploited, fully-exploited, and over-exploited populations. Relative performance of each management strategy was evaluated by comparing escapements and exploitation rates to BRPs for the Gaspereau River Alewife population, and by comparing trends in other population metrics between management strategies. The analysis for this project was completed in the statistical software language R (R Core Team 2019).

3.2.1 The Operating Model

The operating model is a life-cycle based, sex-aggregated, age-structured population model (Gibson and Myers 2003a). The core of the model is a three-dimensional array that tracks the number of fish returning to the river in year *t*, age *a*, that have spawned *p* times previously. The operating model is identical to the simulation model in Chapter 2, except that, rather than treating the exploitation rate as constant, it is allowed to vary annually with both a systematic and random component. The systematic component reflects the result of management actions that are taken based on the assessment results, and the random component allows for implementation uncertainty (identical fishing efforts may produce different exploitation rates due to differences in water levels, timing of the spawning run, or other factors).

Implementation uncertainty is introduced by drawing a value for the annual exploitation rate, μ_t, from a distribution described as:

$$\text{logit}(\mu_t) \sim N\left(\text{logit}(\mu_{curr}), \sigma_\mu^2\right),$$

where

$$\text{logit}(\mu_{curr}) = \log\left(\frac{\mu_{curr}}{1-\mu_{curr}}\right),$$

and

$$\sigma_\mu^2 = \text{var}\left(\log\left(\frac{\mu_{hist}}{1-\mu_{hist}}\right)\right).$$

The exploitation rate in the current management period is μ_{curr}. The initial value of μ_{curr} is the exploitation rate that the simulation beings with (0.2, 0.53, or 0.8). The value of μ_{curr} then changes to an exploitation rate set by a management action, and will change each time management action is taken. The variance of the logit-transformed historical exploitation rates (μ_{hist}) of the Gaspereau River Alewife fishery is σ_{μ}^{2}: historical exploitation rates are from the current management period established in 2001 (McIntyre et al. 2007). Finally, the logit of μ_t is converted to the actual annual exploitation rate which is then used in the operating model: $\mu_t = \frac{\exp(\text{logit}(\mu_t))}{1+\exp(\text{logit}(\mu_t))}$.

As was done in Chapter 2, life history variability was introduced in two ways: as annual variability in the stock-recruitment (SR) relationship, and as variability in the probability of maturation-at-age within each cohort.

Life history parameter values from the Gaspereau River Alewife population (Table 3.2), as estimated by Gibson and Myers (2003a), were used in the operating model. The model was initialized by filling in the first three years with numbers of recruits calculated from the spawning stock biomass at maximum sustainable yield for the Gaspereau River Alewife population. The population was then projected forward 60 years. Management strategies were only applied after the first 25 years of each 60-year projection, to allow the population time to reach a stochastic equilibrium relative to population vital rates.

3.2.2 Management Strategies

The first component of each management strategy herein is the assessment model. I used a Poisson generalized linear model (GLM) to fit a catch curve where numbers of fish are regressed against number-of-previous-spawnings factored by age-at-maturity. This method is described in more detail in Chapter 2. The Poisson GLM was used to estimate the instantaneous total mortality rate (Z) with age composition data from a random sample of 500 fish. The Poisson GLM provides an estimate of Z identical to that of the Chapman-Robson method (Chapman and Robson 1960, Millar 2014). In this model, I incorporated the previous spawning history of the fish by using number-of-previous-spawnings factored by age-at-maturity as the independent variable, rather than only using age as is traditionally done. Incorporating previous spawning history into the model used all available data, including those from partially recruited age classes. This is the preferred method based on the analyses in Chapter 2.

The catch curve model yields an estimate of the instantaneous total mortality rate ($\widehat{Z_t}$), from which the estimated instantaneous fishing mortality rate ($\widehat{F_t}$) can be calculated, assuming a known and constant natural mortality rate. $\widehat{F_t}$ can then be converted to the estimated exploitation rate, $\widehat{\mu_t}$. Values of $\widehat{\mu_t}$ were compared to BRPs for the population. Exploitation rates of 0.35 and 0.53 mark the lower and upper boundaries of the fully-exploited zone (Figure 3.1). Exploitation rates lesser or greater than these boundaries were considered under- or over-exploited.

Four different HCRs were compared (Table 3.3). Each rule consisted of a target exploitation rate and a correction factor that was applied to the subsequent year's exploitation rate to reach that target. Broadly, the correction factor is the amount that fishing effort needs to be changed to correct for under- or over-fishing. To avoid scenarios where the exploitation rate could exceed 1.0, exploitation rates were converted to instantaneous fishing mortality rates (*F*) prior to calculating correction factors. Corrected *F*s were then converted back to exploitation rates.

HCR 1 had a target exploitation rate of 0.53, the upper boundary of the fully-exploited zone (Figure 3.1). The correction factor δ for the mortality rate in year $t + 1$ was calculated as

$$\delta_{t+1} = \frac{F_{target}}{\widehat{F_t}},$$

where F_{target} is the target instantaneous fishing mortality rate, and $\widehat{F_t}$ is the estimated instantaneous fishing mortality rate for year t. I deemed this correction a "full-correction". In a deterministic model where $\widehat{F_t}$ is known without error, multiplying the instantaneous fishing mortality rate in year $t + 1$ by δ will fully correct the deviated mortality rate to the target. Due to the errors in the estimation of $\widehat{F_t}$, the resulting correction might under- or over-correct the true exploitation rate; the correction may also be in the wrong direction. This is the case with other HCRs as well.

HCR 2 had a target exploitation rate of 0.53 as well but used smaller correction factors to reduce the magnitude of the changes in effort from year to year. I deemed this correction a "half-correction". The correction factor δ for the mortality rate in year $t+1$ was calculated as

$$\delta_{t+1} = \frac{1 - \frac{F_{target}}{\widehat{F_t}}}{2} + \frac{F_{target}}{\widehat{F_t}}.$$

In a deterministic model where $\widehat{F_t}$ is known without error, multiplying the instantaneous fishing mortality rate in year $t+1$ by δ will correct the deviated fishing mortality rate halfway to the target, such that if F_{target} was 0.6 and $\widehat{F_t}$ was estimated as 0.8, F_{t+1} would be 0.7. The "half-correction" was used in an effort to mitigate the effects of implementation error and the errors in estimates of Z produced by catch curve models.

The target for HCR 3 and 4 was the entire fully-exploited zone, rather than the USR. When $\widehat{F_t}$ was within the fully-exploited zone, δ_{t+1} was assigned a value of 1. When $\widehat{F_t}$ was outside of the fully-exploited zone, the correction factor was calculated relative to an exploitation rate of 0.44, the approximate midpoint of the fully-exploited zone (Figure 3.1). HCR 3 and 4 are full- and half-corrections, respectively.

In addition to comparing four HCRs (Table 3.3), I compared six assessment and management schedules (Table 3.1). The first schedule consisted of assessing every year, but no management changes were ever made. This schedule represented a status quo approach. The other schedules consisted of varying combinations of how frequently assessments and management changes were done. A schedule that consisted of assessing every year and managing every year would be analogous to changing the fishing season or number of fishers each year based on the previous years' assessment result; assessment and management happened within the same time step of the model. A schedule that consisted of assessing every year and managing every six years would be analogous to changing fishing effort every six years, based on the average result of the six previous assessments. A wider range of assessment and management schedules were also evaluated, but conclusions did not change and only the initial schedule was included.

3.2.3 Simulation Scenarios

The 60-year projection of the operating model was repeated with different starting exploitation rates to provide a comprehensive evaluation of the management strategies and their capability of recovering under- or over-exploited stocks to the fully-exploited zone. Three starting exploitation rates were used to simulate data, representing an under-exploited (μ=0.2), fully-exploited (μ=0.53), and over-exploited (μ=0.8) stock (Figure 3.1). Each 60-year projection for each starting scenario was repeated 100 times with different random deviates in recruitment, maturation schedule and exploitation rate. The same set of deviates and starting exploitation rates were applied to each management scenario to ensure comparability in the projections. In summary, 100 60-year projections were completed for each combination of the six assessment and management schedules, the four HCRs, and the three starting exploitation rates, for a total of 7200 projections.

3.2.4 Performance Metrics

Performance metrics to evaluate the management strategies were based on comparisons of stock characteristics to existing BRPs for the Gaspereau River Alewife population (Gibson and Myers 2003a). Escapement in number of fish was compared to the upper stock reference point (USR) of 400,000 fish, and the limit stock reference point (LRP) of 235,000 fish. Exploitation rates were compared to the upper and lower boundary of the fully-exploited zone, 0.53 and 0.35 respectively. Trends in the mean and standard deviation of the landings, escapement and exploitation rate (both applied and estimated) were compared between management strategies as well.

Evaluation of the management strategies herein is intended to illustrate the trade-offs between each strategy, rather than to find an optimal solution. Escapements and exploitation rates were compared to BRPs. However, I also qualitatively assessed how landings, variability in escapements, and exploitation rates differed across strategies without reference to a specific quantitative goal or target. Maximizing landings and minimizing variability in each of the metrics were general objectives since they would result in a reliable financial benefit to members of the fishery. Towards achieving this comparison, the mean values of the target metric across the 35 management years in each projection were aggregated across projections and presented as box plots. For example, box plots of mean landings showed which management strategies would be expected to result in high landings; the spread of the boxplot showed how reliable these outcomes were across projections. Similar comparisons were done for the variability between years in each 35-year management period. For example, box plots of the standard deviation in landings showed which management strategies would be expected to result in reliable landings, or which strategies would result in highly variable landings; the spread of these boxplots showed how reliable these outcomes were across projections. Additional presentation of the management strategies included comparison of the proportion of times the correction factor δ moved the mortality rate in the correct direction for each management strategy, and visual depictions of the projected escapement, landings, and exploitation rate over the 60-year projections.

3.3 Results

To demonstrate the nature of the population projections, individual projections with no management changes through time are presented (Figure 3.2). As shown via these individual projections, escapements, landings and the exploitation rates for an individual population were highly variable. With a constant fishing effort that did not change over time (Figure 3.2), a chronically under-exploited population (starting μ=0.20) was found to have an escapement consistently in the healthy zone despite the variability in exploitation rate (Figure 3.2, column 1). Landings in this scenario averaged 273,000 fish and ranged from 92,000 to 671,000, whereas the escapement averaged 1,050,000 fish and ranged from 387,000 up to 2,500,000. In contrast, a fully-exploited population (starting μ=0.53) was found to vary in and out of the target exploitation and escapement zones, including increases in escapement from the critical zone (escapement less than 235,000 fish) to the healthy zone (escapement greater than 400,000 fish) and back in two years (Figure 3.2, row 1, column 2). The landings in this simulation averaged 475,000 and ranged from 170,000 to 1,410,000 (Figure 3.2, row 2, column 2). A chronically over-exploited population was observed to have a mean escapement of 108,000, only ever reaching as many as 351,000 fish (Figure 3.2, row 1, column 3). The landings of an over-exploited and fully-exploited stock were very similar over time (Figure 3.2, row 2, columns 2 and 3), but averaged slightly less in the over-exploited scenario (407,000 versus 475,000). The similarity in landings between the over-exploited and fully-exploited scenarios occurs because their two corresponding exploitation rates are approximately equidistant from the peak of the yield curve for this stock (Figure 3.1). Regardless of starting condition (over-, under-, or fully-exploited), the estimated exploitation rate from the catch curve analysis was highly variable and tracked the true exploitation rate from the population dynamics model relatively poorly (Figure 3.2, row 3). The catch curves tended to underestimate the true exploitation rate for the fully-exploited and over-exploited populations; they rarely overestimated the exploitation rate, and to a much lesser degree (Figure 3.2, row 3, columns 2 and 3). The exploitation rate for the fully-exploited population was underestimated 21 times out of 35, and 20 times out of 35 for the over-exploited population. In the fully-exploited simulation, the exploitation rate was underestimated by 7% on average; this value was 9% for the over-exploited population, whereas the under-exploited population was overestimated by 1% on average.

For illustrative purposes, the single population projection showing the no management change scenario (Figure 3.2) is contrasted with a single population projection using a management strategy of assessing every two years and managing every six years (Table 3.1) with HCR 3 (Table 3.3), where catch curves are used as the basis for the advice (Figure 3.3). When no management changes were made, escapement, landings and the exploitation rate varied greatly from year to year and the status of the population did not change relative to its initialized condition. In contrast, despite the inherent variability, the exploitation rate was typically in the desired zone once assessment and management action commenced (Figure 3.3, row 3). Spawning escapement was rarely in the critical zone (below the LRP) for the population once management began (Figure 3.3, row 1), and the benefits of management on escapement are most clearly seen in the over-exploited population (Figure 3.3, row 1, column 3). Landings increased for the under-exploited population from an average of 311,000 to 426,000 (Figure 3.3, row 2, column 1), and slightly decreased for the fully- and over-exploited populations from averages of 528,000 to 422,000 and 449,000 to 370,000, respectively (Figure 3.3, row 2, columns 2 and 3). The large peak in the landings towards the end of the projections lined up with an escapement well above the average (1,140,000 vs an average of 516,000 for the over-exploited population), which was slightly above the average exploitation rate due to implementation error (Figure 3.3).

The variability in a single projection of the MSE makes it difficult to discern if management actions have a consistent effect on the population; analyzing the distribution of results from many projections can reveal those effects. I used the analyses of 100 projections per scenario to choose among HCRs (Figure 3.4). I found HCR 3, the harvest control rule targeting the fully-exploited zone with a full-correction (Table 3.4), to outperform the other harvest control rules. Exploitation rates for HCR 3 were adjusted much more rapidly than rates for HCR 4, due to the difference between a full and a half-correction (Figure 3.4, row 3, column 4 and 5). HCR 3 resulted in rapid adjustments to appropriate levels for all starting exploitation rates (Figure 3.4, row 3, columns 2-4). Escapement levels were adjusted to appropriate levels in about five years (Figure 3.4, column 4), compared to fifteen years for HCR 4 under similar conditions (Figure 3.4, column 5). The landings under HCR 3, regardless of the starting exploitation rate, were very similar to the landings under a constant exploitation rate at a fully-exploited level (Figure 3.4, row 2, column 1), but an escapement above the target level and an exploitation rate within the target zone were regularly maintained under HCR 3 (Figure 3.4, columns 2-4).

Harvest control rules 1 and 2 performed less well. HCR 1 and 2 tended to be in the over exploited zone most frequently, while HCR 3 and 4 tended towards the fully-exploited zone (Figure 3.5). HCR 4 still had a slight bias towards the over-exploited or under-exploited zone, when the starting exploitation was over- or under-exploited, respectively (Figure 3.5). HCR 3 was the least biased, resulting in an exploitation rate that was more frequently in the fully-exploited zone than the other zones; however, HCR 1 and 2 resulted in an exploitation rate that was in the fully-exploited zone more frequently than HCR 3, but were heavily biased towards the over-exploited zone (Figure 3.5).

None of the management strategies consistently corrected the exploitation rate in the right direction, when the exploitation was in the over- or under-exploited zones (Table 3.4). At best, HCR 3 applied a correction factor that moved the next year's exploitation rate towards the desired zone ~83% of the time; HCR 4 was worse at ~74% of the time, and HCR 1 and 2 applied the correction factor in the desired direction ~70% and 64% of the time, respectively (Table 3.4). Changing the frequency of assessments and management actions had little effect on the performance of all the HCRs. The different assessment and management schedules had little to no effect on how frequently the exploitation rate was in one zone or the other (Figure 3.5). Making management changes less frequently and conducting assessments more often tended to result in less variable landings and escapements for populations with a starting exploitation rate in the under-exploited zone (Figure 3.6, row 2 and 4). This pattern held true for the fully-exploited (Figure 3.7, row 2 and 4) and over-exploited populations (Figure 3.8, row 2 and 4). HCRs 2 and 4 tended to have increased variability in μ when fewer assessments were included in a single management action: assessing and managing every year performed similarly to assessing and managing every six years; variability in μ was lowest when assessments were done every year and management actions were taken every six years (Figure 3.6, row 6). This pattern persisted across all starting exploitation rates and suggests that HCRs using the half-correction resulted in increased variability in μ (Figure 3.7, row 6 and Figure 3.8, row 6). The variability in μ for HCRs

1 and 3 was not affected by changes in assessment and management frequency. HCR 1 resulted in the greatest mean landings across all assessment and management scenarios and all starting exploitation rates, most notably when the starting exploitation rate was in the under-exploited zone (Figure 3.6, row 3). When the starting exploitation rate was fully-exploited or over-exploited, HCR 2 resulted in landings almost identical to those produced under HCR 1 (Figure 3.7, row 3 and Figure 3.8, row 3). HCR 4 had the greatest mean escapement for populations starting in the under-exploited zone (Figure 3.6, row 1). For the fully-exploited populations, HCR 3 and 4 yielded very similar mean escapements; HCR 3 resulted in more consistent mean escapements than HCR 4 (Figure 3.7, row 1). Finally, HCR 3 resulted in the greatest mean escapement for populations that were chronically over-exploited (Figure 3.8, row 1).

The majority of the management strategies tested resulted in successful management of the stocks. Chronically over- or under-exploited stocks recovered to healthier levels. Management maintained exploitation rates within the desired zones, resulting in sustainable escapements and landings throughout the 60-year projections. Of the HCRs evaluated, HCR 3 resulted in escapements above the target level and exploitation rates within the fully-exploited zone more often than the other HCRs. Of the assessment and management schedules evaluated, managing every six years, and assessing at least every three years produced desirable results, although the effect of changing the assessment and management were slight.

3.4 Discussion

My MSE suggests that catch curve models can be a very useful assessment method for managing river herring populations, even though they are poor and biased estimators of the instantaneous total mortality rate (Smith et al. 2012, Millar 2014) for short-lived species with variable recruitment like river herring (Gibson et al. 2017, Chapter 2). Below, I discuss how changes in management tended to increase the median landings and exploitation rate throughout the projections. There were instances where management advice for a single management period of one to several years resulted in decreased landings and escapement, however there were no chronic patterns of these occurrences. Furthermore, given the variable life history of river herring, occasional decreases in landings or escapement are unavoidable. Additionally, I discuss the trade-offs of assessing and managing a stock more or less frequently, and potential avenues to improve upon this MSE.

While catch curves have proven to be a useful assessment tool for river herring, they do have limitations. As discussed by Smith et al. (2012), Millar (2014), Gibson et al. (2017) and in Chapter 2, catch curves are known to produce negatively biased estimates of the instantaneous total mortality rate. Additionally, using a HCR based solely on an estimate of the fishing mortality rate derived from the total mortality rate requires the assumption that the natural mortality rate is both known and constant over time. Any changes in the natural mortality rate would be attributed to changes in the fishing mortality rate, potentially leading to poorly informed management actions. In addition, recruitment failures, due to reproductive failure or mortality occurring before fish first return to the river, would go largely undetected by a HCR based solely on the total mortality rate. If both the landings and spawning escapement decline simultaneously, there would be little change in the annual exploitation rate.

Simpler assessment methods can be even more problematic than catch curves. Many HCRs proposed for other fisheries are catch-based, especially for data-limited stocks, and have performed relatively poorly in some previous MSEs (Carruthers et al. 2014). A HCR based solely on catch would be insensitive to changes in the exploitation rate or fishing effort if escapement is also changing, and may only reliably detect stock declines when numbers of fish dramatically decreased. Conversely, a HCR that is based on multiple stock metrics, such as combinations of landings, escapement, and exploitation rate have been demonstrated to outperform HCRs based on single metrics (Carruthers et al. 2014); a more comprehensive, multi-faceted HCR could better detect stock declines that are only detectable in select metrics. Catch curves are intermediate between these scenarios.

The MSE undertaken here includes the most important features of an MSE, but not all. Punt et al. (2014) list a number of best practices for conducting an MSE. Referring to their Table 1, I selected objectives and performance metrics by stating the goal of the study, to determine if catch curves are a useful assessment method for river herring, answered by comparing the simulated population metrics to BRPs. I selected uncertainties relative to the life history of river herring and to the fishery and incorporated them into the operating model as recruitment variability and implementation uncertainty. A range of candidate management strategies were proposed, each with the catch curves as the assessment component to provide a comprehensive evaluation of that assessment method. I used a previously tested simulation model as the operating model (Gibson and Myers 2003a). Finally, I presented the results of the MSE graphically to clearly depict how well the various management strategies performed, relative to the BRPs for the simulated stock. I did not include the process of developing management strategies with stakeholders in the fishery and fisheries managers. The management strategies in this MSE were intended to demonstrate the utility of catch curve models as an assessment tool, rather than reflect the goals of multiple stakeholder groups.

While the median escapement, landings and exploitation rate of the simulated projections rapidly decreased to appropriate levels when management was applied, there was still a lot of variability between years in individual scenarios. This variability, modelled here as recruitment variability and implementation uncertainty, is inherent to river herring and their fisheries (Gibson and Myers 2003a). In this study, the target species is short-lived and has variable recruitment leading to large differences in cohort size (Loesch 1987), and the fishery’s exploitation rate can vary greatly year-to-year (Gibson et al. 2017). The Gaspereau River Alewife fishery which is used as a model for this MSE has a fixed fishing season, where fishing is only permitted on Mondays, Tuesdays, Thursdays and Fridays (McIntyre et al. 2007). Instances have been observed when favourable temperature, flow and tide conditions align on weekends resulting in a larger portion of the spawning run escaping the fishery without being caught (personal observation). This fishery is primarily managed by controlling fishing effort, wherein the time spent fishing is increased or decreased to adjust the exploitation rate (McIntyre et al. 2007). Variability in the landings is unavoidable with this type of management strategy. It is therefore appropriate to judge the success of a management strategy by comparing median escapement and exploitation rate to BRPs. Furthermore, assessing and managing via catch curves remain useful for managing river herring populations even while accounting for implementation uncertainty due to the year-to-year variability in individual projections. In short, what I have demonstrated through this MSE is that although the exploitation rate will vary while conditions are constant, changes to management will on average result in stable long-term landings and escapements. The superior results from using HCR 3 compared to the other HCRs could in part be attributed to its target exploitation rate. HCR 3 resulted in the population recovering from over- or under-exploitation to the fully-exploited zone more quickly than the other HCRs. HCR 3 had the lowest target exploitation rate, 0.44, and applied a full-correction each time management action was taken. The rapid adjustment to appropriate exploitation rates contributed to greater landings and escapements than projections under HCRs 2 and 4 where the adjustment to the appropriate zones was slower, resulting in more years spent at less productive harvest rates. A target exploitation rate of 0.44, in the middle of the fully-exploited zone, would reduce bias towards being over-exploited compared to a target of 0.53 on the upper border of the fully-exploited zone. Additionally, the typical landings observed under HCR 3 were very similar to the landings observed under HCR 1 or 2, which had target exploitation rates of 0.53. Despite the higher target exploitation rate of HCR 1 and 2 compared to HCR 3, the expected yield from harvesting at 0.53 is not much more than from harvesting at 0.44 (Figure 3.1). The slight decrease in landings from lowering the exploitation rate from 0.53 to 0.44 resulted in a population that is within the desired zones for escapement and exploitation rate more often and more reliably.

Beyond which assessment technique is appropriate, it is important to consider the real-world impacts of assessment frequency, especially in terms of data collection and impact on participants in the fishery. The cost of collecting data relative to the value of the data being collected must be considered. River herring fisheries in DFO's Maritimes Region vary greatly in size, with even small stocks being exploited. Rivers such as the Saint John River in New Brunswick, and the Margaree River in Nova Scotia can support river herring spawning runs of millions of fish (Chaput et al. 2001, Jessop 2001) while numerous smaller rivers throughout Nova Scotia and New Brunswick support spawning runs as small as a several thousands of fish. The reality of coordinating multiple concurrent stock assessments across rivers ranging over hundreds of kilometers is a difficult task, constrained by available experienced personnel and budget. Rotating annual assessments across several index rivers, such that each population is assessed every two or three years, but only two or three populations are assessed in any given year, may provide an acceptable trade-off between frequency of assessment and the number of populations assessed in an area (Gibson et al. 2017). Management frequency also has real world implications to the participants in a fishery. The Gaspereau River Alewife fishery underwent a management change from fishing five days a week to four in 2001 (McIntyre et al. 2007), yet no changes to fishing effort have been made since. Frequent management changes to a fishery that has historically been relatively unchanged may be received poorly by members of the fishery. In an effort-managed fishery, restrictions could be placed on legal gear, fishing time, or season length. Altering these factors each year would affect gear purchases, hiring of short-term seasonal personnel, and coordinating time for fishing with other work. Too many changes would lead to frustration, especially if the changes in landings and escapement from one management period to another do not appear significantly different. Increasing and decreasing fishing effort too frequently may lead to confusion over which set of regulations is the most current. Managing less frequently or reducing the degree of management changes could help this problem. HCRs 2 and 4 each consisted of a half-correction, and although they were not the preferable HCRs by my metric, they would satisfy a need for reduced management changes; however, they would only reduce the magnitude of the change, but not the frequency. Punt et al. (2014) discuss management strategies in Australia that impose constraints on the degree of change, such as limiting changes of the TAC to no more than 50% of its current value and making no changes at all if the prescribed change is less than a 10% change. Trade-offs of accurately managing a fishery and over-managing a fishery should be considered.

In addition to considering how management actions are received, the challenge of implementing new management rules should be considered. The tendency of people to circumvent regulations is not readily simulated, and difficult to incorporate into MSEs. Neither did I consider simulating the efficacy of monitoring and enforcing management changes. These factors are well beyond the scope of this MSE, but they do underline the point that the management advice produced from a MSE needs to be appropriate and realistic for the fishery being simulated.

Frequent, consistent assessments that collect more than just the age composition data for the catch curve models can have additional benefits not simulated in this MSE. Regular and consistent collection of age composition data for the catch curve model evaluated in this MSE can be paired with auxiliary data such as estimates of escapement, landings, or larval abundance which can be used in more sophisticated stock assessment models for river herring such as statistical catch-at-age models (Gibson and Myers 2003a) or virtual population analyses (Chaput et al. 2001). These more sophisticated models can provide more information about the stock with a greater degree of certainty. Rago (2001) demonstrated with Atlantic salmon (*Salmo salar*) that incorporating even a single year of auxiliary data to a catch-index model that estimates spawner abundance and exploitation rate can dramatically improve the accuracy of those estimates. While not directly transferable to this work, the principle of using a small amount of auxiliary data to anchor an estimated time series could be applicable to the management strategies evaluated herein for river herring.

The MSE presented in this chapter could be modified to address additional challenges of managing river herring stocks. Evaluation of HCRs that incorporate auxiliary data such as estimates of the annual landings or escapement would be useful. Wayte and Klaer (2010) found success in using an estimate of the annual landings along with their estimate of the fishing mortality rate in their MSE. Including a landings or escapement estimate would help detect recruitment failures and would be a more robust assessment tool. Modifying the operating model to simulate scenarios in which mortality rates gradually increase throughout the projection could simulate the effect on the population of declining habitat quantity or quality, or an increase in mortality from other non-monitored sources. These modifications would facilitate the development of management strategies that are robust to large-scale change. The ability of various management strategies to combat these kinds of special circumstances is an important question to consider. Another potential avenue of research would be to investigate the effect of bias in sampling. In the simulations, subsamples of the population age structure are random and unbiased. In reality, some degree of selectivity would be present in the subsample of the age structure. I also assume that the fishery is non-selective, which is a standard assumption for riverine river herring fisheries (Chaput et al. 2001, ASMFC 2012, Billard 2017). Despite this assumption, it would be useful to investigate the effect that varying degrees of selectivity would have on assessment and management of a population.

In summary, I demonstrated that catch curve models are a useful assessment tool for managing river herring stocks by providing advice that can help lead to sustainable long-term management of the stock. A management strategy that is solely reliant on an estimate of the exploitation rate derived from catch curves has its shortcomings and could likely be improved with an assessment method that can incorporate more data from different sources. Despite the shortcomings of using catch curves as the sole assessment component of a management strategy for river herring, using catch curves appear to fulfill the need for a short-term assessment method for river herring.

3.5 Tables and Figures

Table 3.1. Description of the assessment and management schedules assessed in the MSE. Each assessment and management schedule was simulated for the combinations of the four harvest control rules and three starting exploitation rates. The two leftmost columns describe the frequency of assessment and management. For example, A1 is assessing every year, M6 is managing every six years. M0 means no management happens.

Assessment (A) and Management (M) Schedule		Description
A1	M0	Assessments are every year, and no management changes are made.
A1	M1	Assessments are every year and management changes happen every year based on a single assessment result. Assessment and management happen with in the same simulation year.
A6	M6	Assessments are every six years and management changes happen every six years based on a single assessment result. Assessment and management happen with in the same simulation year.
A1	M6	Assessments are every year and management changes happen every six years based on an average of the six assessment results.
A2	M6	Assessments are every two years and management changes happen every six years based on an average of the three assessment results.
A3	M6	Assessments are every three years and management changes happen every six years based on an average of the two assessment results.

Table 3.2. Parameter values for the operating model used to simulate numbers-at-age datasets of Alewife for the MSE. Life history parameter values were derived by Gibson and Myers (2003a) using data from the Gaspereau River, NS, Alewife population. Values for indices t, a, and p are the range of values used in the model.

Term	Description	Value
	Indices	
t	Time (years)	1-60
a	Age class	2-9
p	Previous spawning class	0-6
	Life history parameters	
α	Slope at the origin of the Beverton-Holt spawner-recruit relationship	96.10
R_{asy}	Asymptotic recruitment level for the Beverton-Holt spawner-recruit relationship	1,563,665
M^{adult}	Instantaneous natural mortality rate for adults	0.53
M^{juv}	Instantaneous natural mortality rate for juveniles	0.40
m_2	Probability of maturation at age-2	<0.01
m_3	Probability of maturation at age-3	<0.01
m_4	Probability of maturation at age-4	0.53
m_5	Probability of maturation at age-5	0.98
m_6	Probability of maturation at age-6	1.00
	Exploitation rate	
μ	Portion of the mature population that is harvested in each year t	0.2, 0.53, 0.8
	Variance parameters	
σ_μ	Exploitation rate variance	0.50
σ_R	Recruitment variance	0.42
$var(logit(m_2))$	Variance of maturation probability at age-2 (logistic scale); low, medium and high variability scenarios	0, 0.853, 1.76
$var(logit(m_3))$	Variance of maturation probability at age-3 (logistic scale); low, medium and high variability scenarios	0, 0.712, 0.910
$var(logit(m_4))$	Variance of maturation probability at age-4 (logistic scale); low, medium and high variability scenarios	0, 0.919, 2.44
$var(logit(m_5))$	Variance of maturation probability at age-5 (logistic scale); low, medium and high variability scenarios	0, 0.995, 5.32
$var(logit(m_6))$	Variance of maturation probability at age-6 (logistic scale); low, medium and high variability scenarios	0, 0, 0

Table 3.3. Descriptions of the four harvest control rules (HCR) used in the MSE. The target exploitation rate for HCRs 3 and 4 is the entire fully-exploited zone; the midpoint of the zone, 0.44, is used for calculating correction factors. If the estimate of the exploitation rate ($\hat{\mu}$) is within the fully-exploited zone, no correction is made for HCRs 3 and 4.

Name	Target exploitation rate	Correction	Notes
HCR 1	0.53	Full	Corrections made every time
HCR 2	0.53	Half	Corrections made every time
HCR 3	0.35-0.53 (0.44 as the mid-point)	Full	Corrections made only when $\hat{\mu}$ is not in the fully-exploited zone
HCR 4	0.35-0.53 (0.44 as the mid-point)	Half	Corrections made only when $\hat{\mu}$ is not in the fully-exploited zone

Table 3.4. Proportion of times the harvest control rule changed the fishing effort in the correct direction (i.e. reduced exploitation when population was truly over-exploited, or increased exploitation when population was truly under-exploited), when the actual exploitation rate was not in the fully-exploited zone.

Starting Exploitation Rate	HCR1	HCR2	HCR3	HCR4
Under Exploited (u=0.2)	0.695	0.640	0.830	0.768
Fully-exploited (u=0.53)	0.706	0.645	0.823	0.739
Over Exploited (u=0.8)	0.711	0.642	0.819	0.736

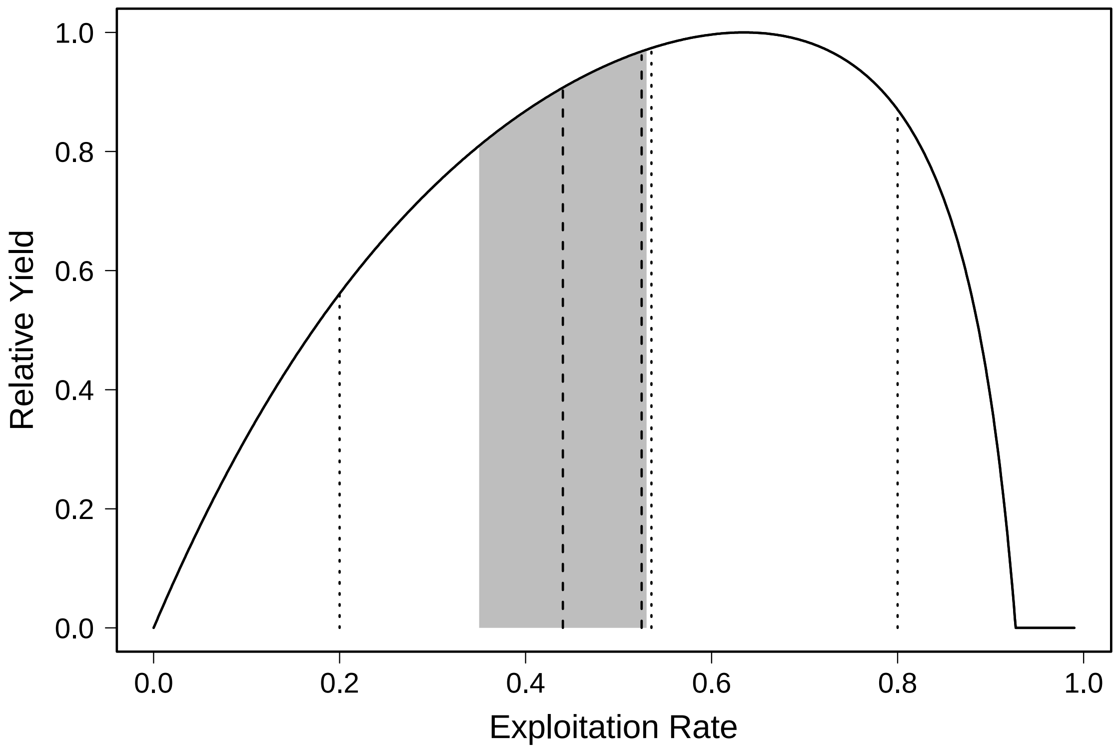

Figure 3.1. A yield curve for the Gaspereau River Alewife population showing the relationship between exploitation rate and yield, where yield is relative to the maximum possible yield. The grey area under the curve represents the fully-exploited zone. The dotted lines represent the three starting exploitation rates used in the MSE. The dashed lines represent the target exploitation rates for harvest control rules 1 and 2 (0.53), and 3 and 4 (0.44). The two lines at 0.53 on the X-axis are staggered for viewing but represent the same values.

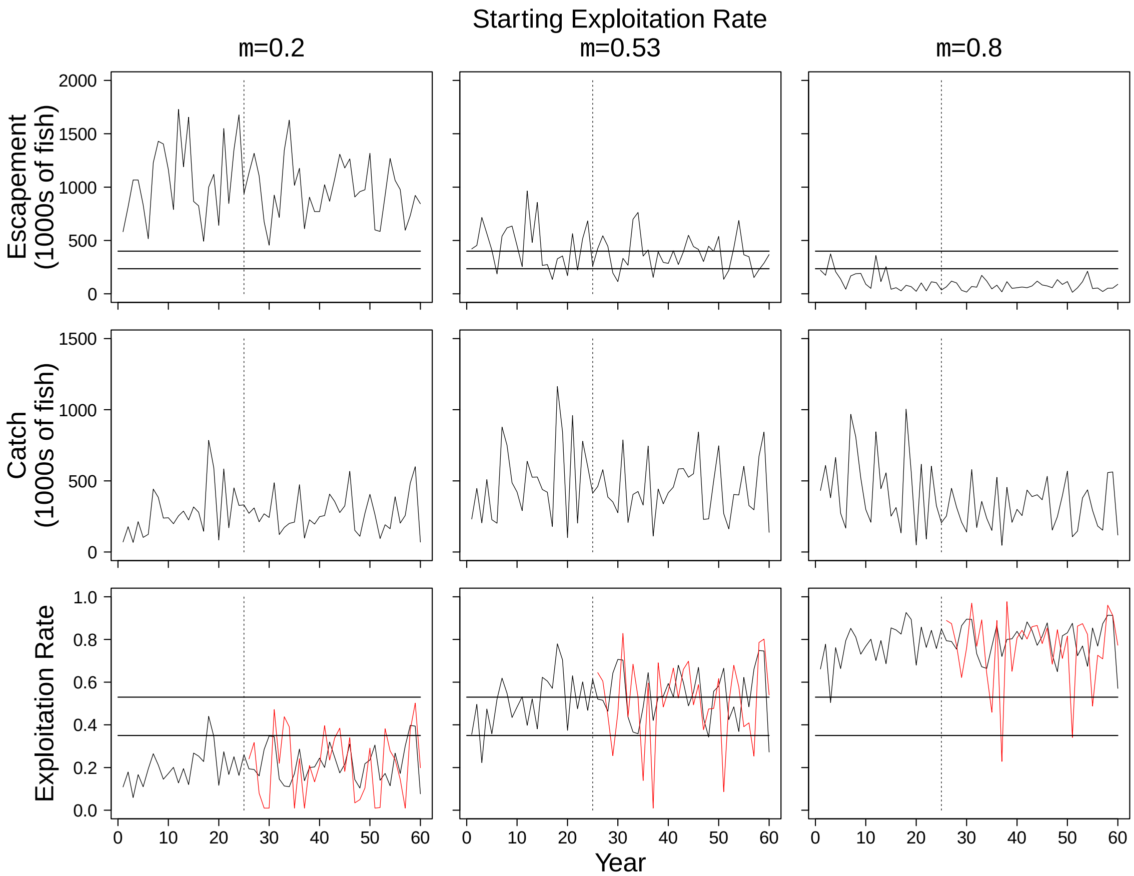

Figure 3.2. Examples of individual 60-year projections of the MSE for populations that are under-exploited (μ=0.20; left column), fully-exploited (μ=0.53; mid column) and over-exploited (μ=0.80; right column) when no management changes are made. For each scenario, the annual escapement (top row), catch (middle row) and exploitation rate (bottom row) are shown. The solid horizontal lines in the top row show the upper stock reference point (400,000 fish) and the limit stock reference point (235,000 fish). The solid horizontal lines in the bottom row show the boundaries of the fully-exploited zone (0.53 and 0.35). The vertical dashed lines show when assessment commenced, at year 25. In the bottom row, the black lines show the true exploitation rate for the population, while the red lines show the exploitation rate estimated with the catch curve model each year. The same random numbers are used in all simulations.

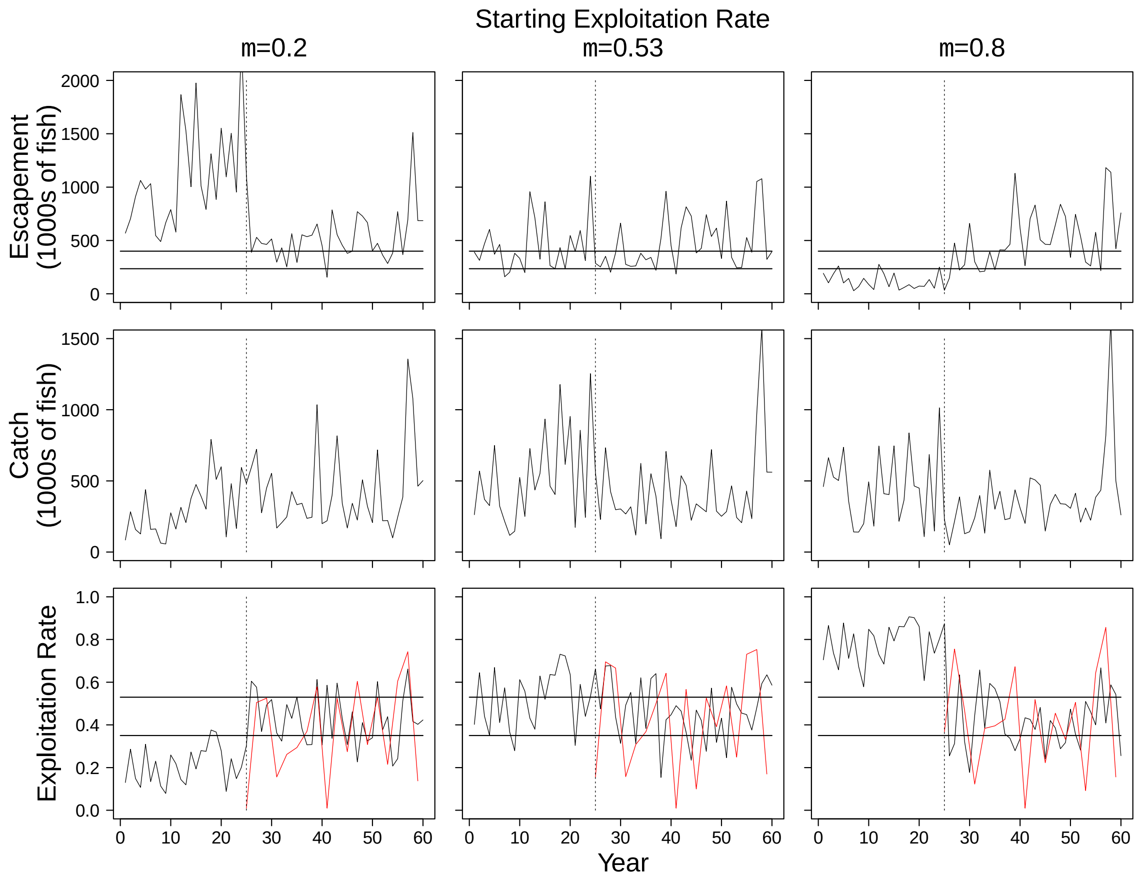

Figure 3.3. Examples of individual 60-year projections of the MSE for populations that are under-exploited (μ=0.20; left column), fully-exploited (μ=0.53; mid column) and over-exploited (μ=0.80; right column) when the stock is assessed every two years and managed every six years under harvest control rule 3. For each scenario, the annual escapement (top row), catch (middle row) and exploitation rate (bottom row) are shown. The solid horizontal lines in the top row show the upper stock reference point (400,000 fish) and the limit stock reference point (235,000 fish). The solid horizontal lines in the bottom row show the boundaries of the fully-exploited zone (0.53 and 0.35). The vertical dashed lines show when assessment and management commenced, at year 25. In the bottom row, the black lines show the true exploitation rate for the population, while the red lines show the exploitation rate estimated with the catch curve model every two years. The same random numbers are used in all simulations.

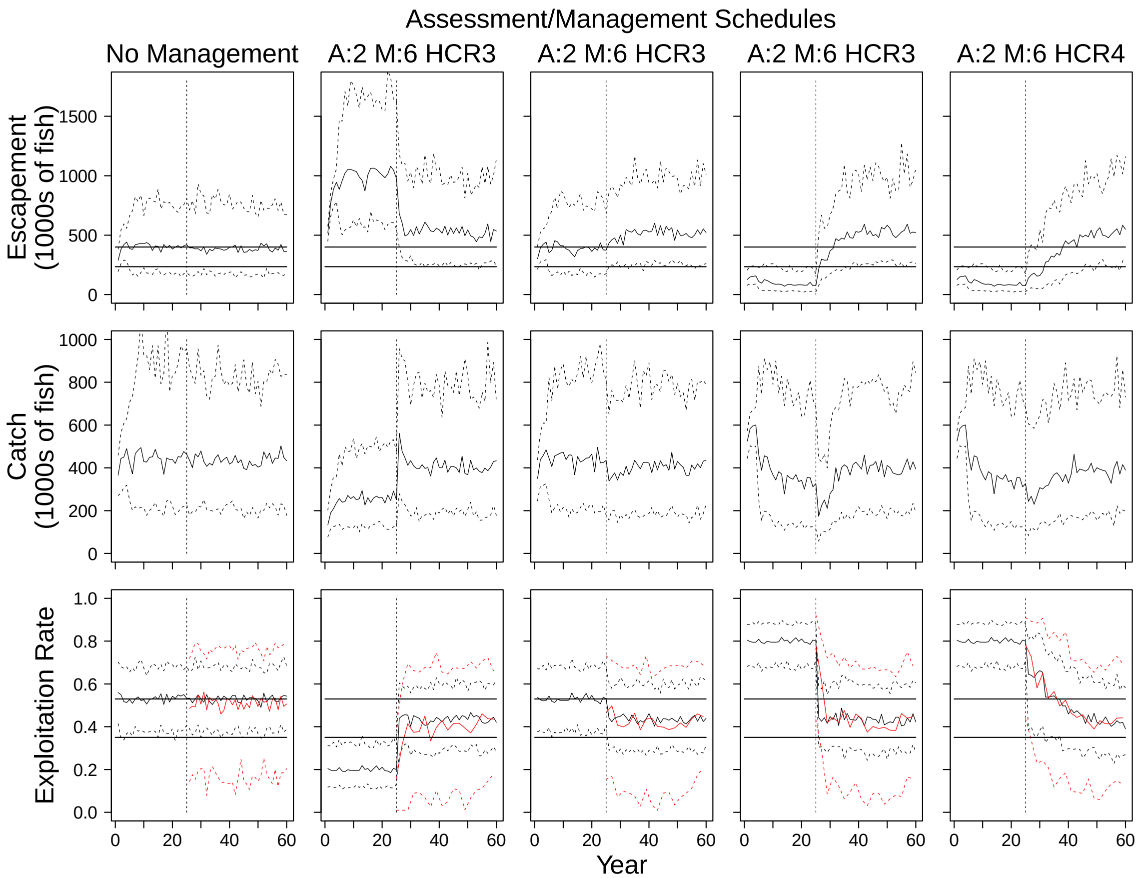

Figure 3.4. The distribution of annual escapement (top row), catch (middle row), and exploitation rate (bottom row) for 100 60-year projections done for five management strategies with different assessment and management schedules (Table 3.1) harvest control rules (HCR: Table 3.3) and starting exploitation rates in the MSE. The first column shows projections of a fully-exploited population (μ=0.53) where no management changes are made. Columns two through four show projections for under-, fully and over-exploited populations (μ=0.2, μ=0.53, μ=0.8) that are assessed every two years and managed every six years under HCR 3. Column five shows projections of over-exploited populations (μ=0.8) assessed every two years and managed every six years under HCR 4. The solid lines in each panel represent the median values, while the dotted lines represent the 10th and 90th percentiles of the 100 projections. The solid horizontal lines in the top row represent the upper stock reference point (400,000 fish) and the limit stock reference point (235,000 fish). The solid horizontal lines in the bottom row represent the boundaries of the fully-exploited zone (0.53 and 0.35). The vertical dashed lines indicate when assessment and management commenced, at year 25. In the bottom row, the black lines show the median and spread of the exploitation rate applied to the population, while the red lines show median and spread of the exploitation rate estimated with the catch curve model.

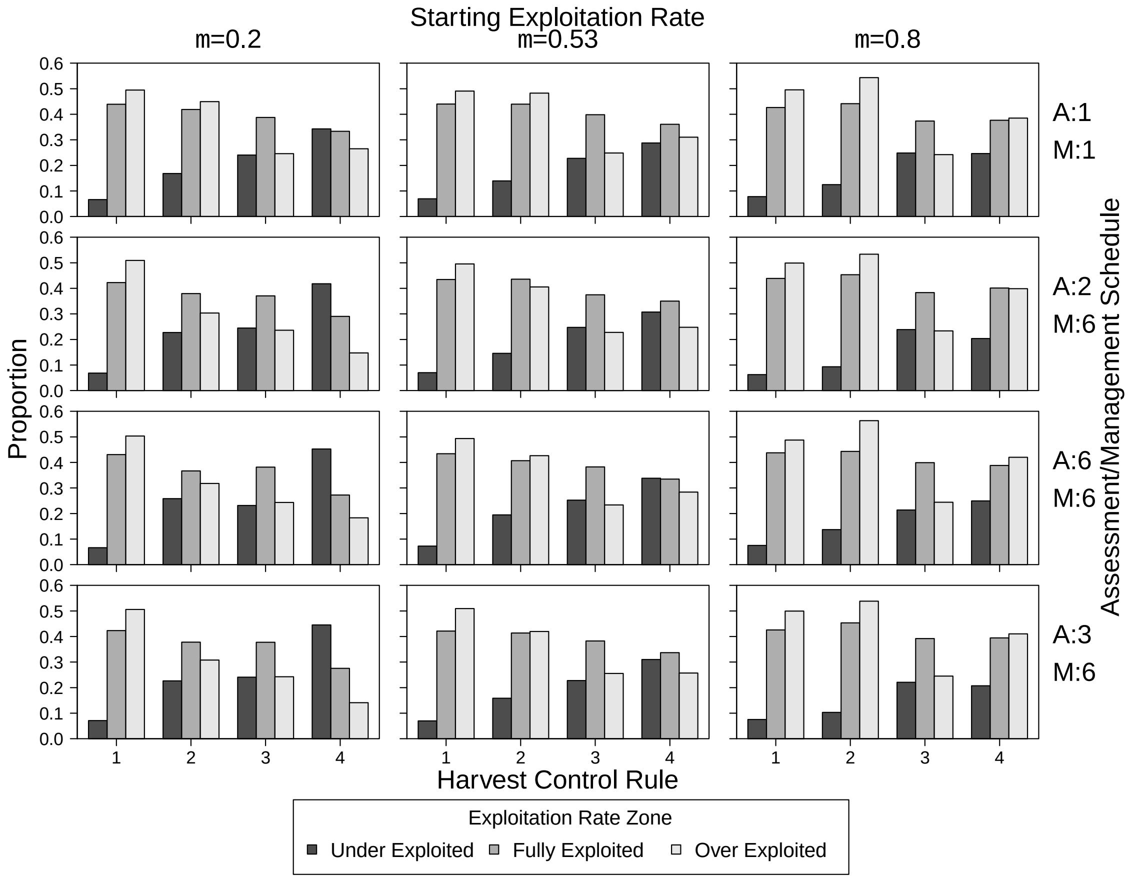

Figure 3.5. A bar plot showing the proportion of times the exploitation rate for any post-management year is in one of the three exploitation rate zones in the MSE. Proportions are shown for the four harvest control rules (HCR: Table 3.3). HCR 1 has a target exploitation rate of 0.53 and has a full-correction. HCR 2 has a target exploitation rate of 0.53 and has a half-correction. HCR 3 has a target exploitation rate of 0.44 and has a full-correction. HCR 4 has a target exploitation rate of 0.44 and has a half-correction. The four rows show four different assessment and management schedules, were the frequency of assessment and management is represented by the value next to the "A" and "M" (Table 3.1). The three columns show the exploitations rates that are applied to the populations prior to management commencing at year 25.

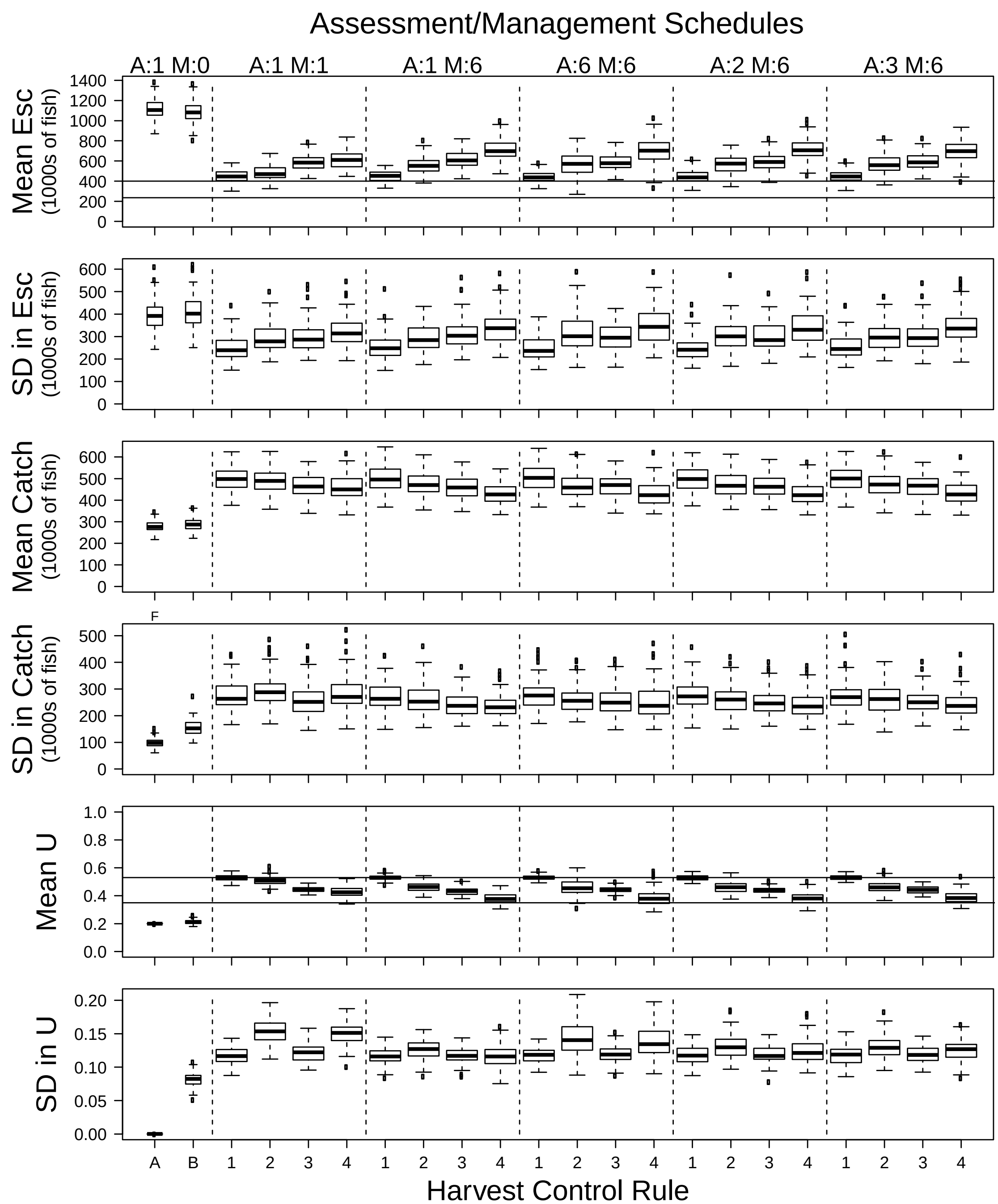

Figure 3.6. Boxplots that summarize 100 simulated management projections of the four harvest control rules (HCRs) tested in the MSE for under-exploited populations (μ=0.2). On the X-axis are the four HCRs described in Table 3.3, and controls A and B. Controls A and B depict annual assessment results when no management actions are taken. Control A was simulated without implementation error, while control B was simulated with implementation error. The vertical dashed lines separate the different assessment and management schedules described in Table 3.1. On the Y-axis are the six metrics for evaluating management strategies. Performance is evaluated with mean escapement (top row), where the boxplots summarize 100 mean escapements; each mean is calculated from the annual escapements in the 35 post-management years in each of the

60-year projections. Similarly, the standard deviation in escapement (second row) is evaluated with boxplots summarizing 100 standard deviation values. Catch and exploitation rate (U) are presented in an identical fashion. The solid horizontal lines in the top row show the upper stock reference point (400,000 fish) and the limit stock reference point (235,000 fish). The solid horizontal lines in the fifth row show the boundaries of the fully-exploited zone (0.53 and 0.35).

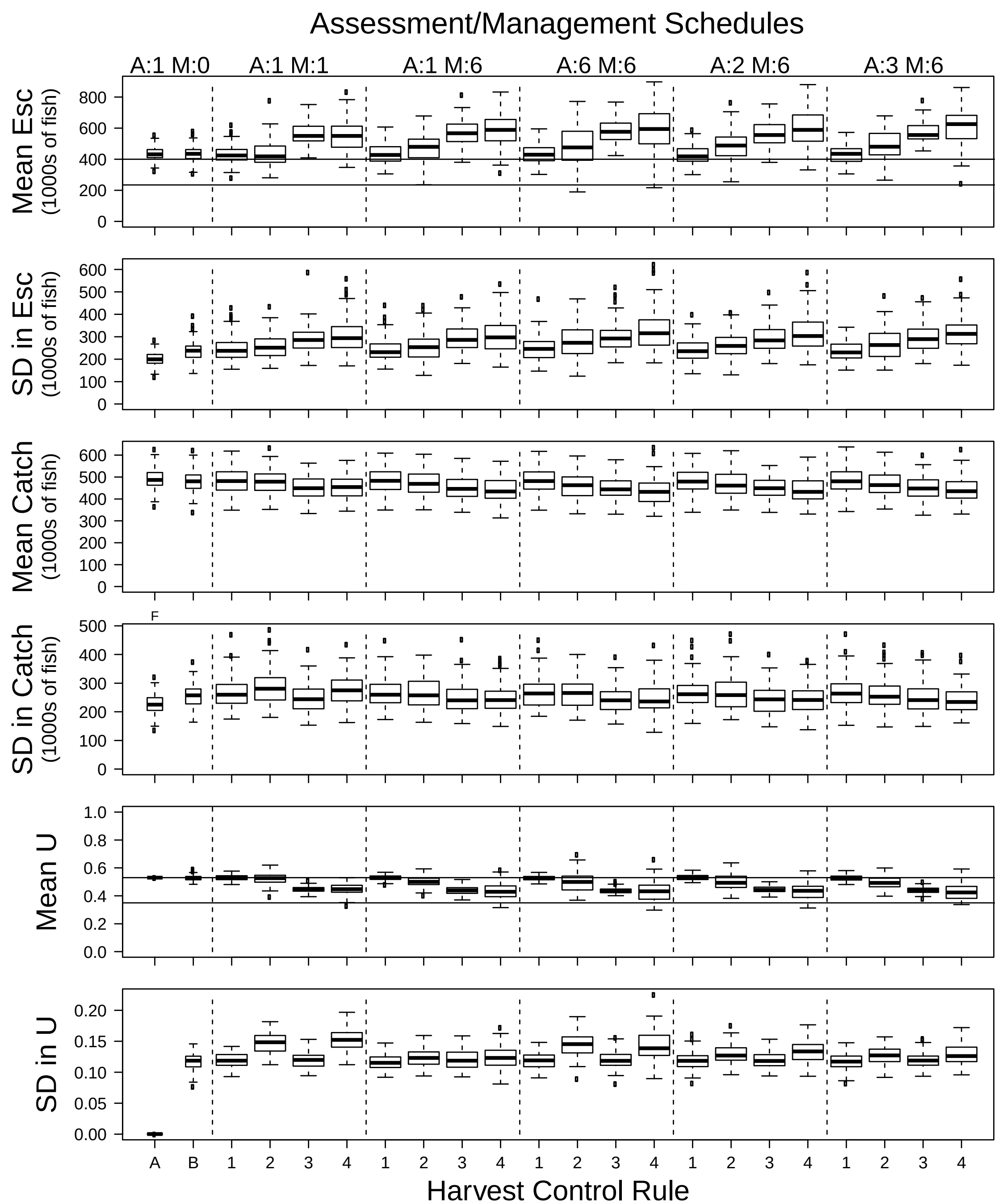

Figure 3.7. Boxplots that summarize 100 simulated management projections of the four harvest control rules tested in the MSE for fully-exploited populations (μ=0.53). Descriptions of the axes can be found in the caption of Figure 3.6.

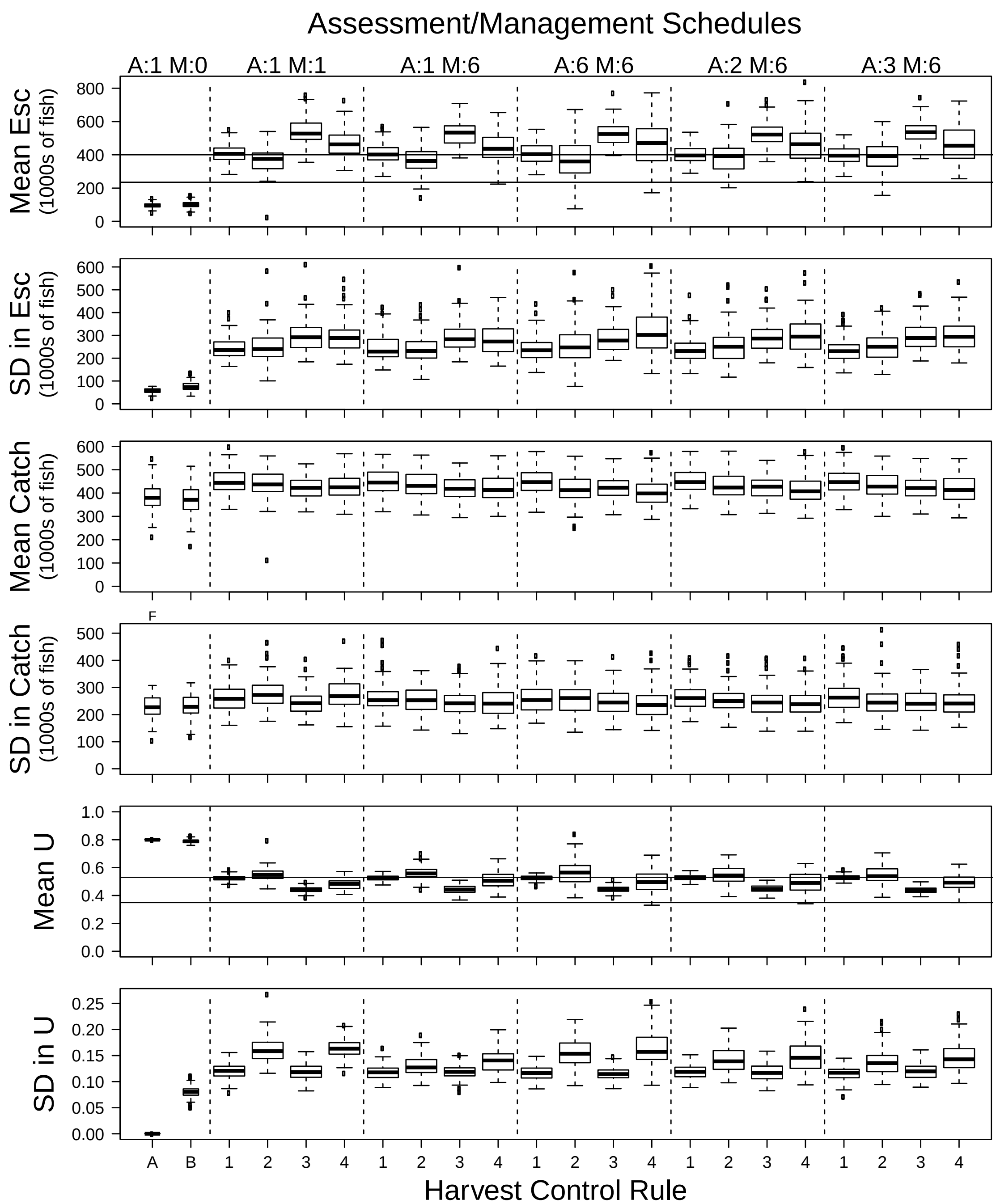

Figure 3.8. Boxplots that summarize 100 simulated management projections of the four harvest control rules tested in the MSE for over-exploited populations (μ=0.8). Descriptions of the axes can be found in the caption of Figure 3.6.

CHAPTER 4: Recapitulation

Two approaches towards evaluating catch curve models as an assessment method for river herring were undertaken in this thesis. The first approach consisted of an evaluation using statistical criteria in which the accuracy of estimates of Z from two catch curve models, fit with four different statistical methods, was compared. Of the models that were evaluated, the use of the number-of-previous-spawnings as the independent variable outperformed using age as the independent variable regardless of the statistical method used. Fitting this model with the Poisson GLM provided the best estimates of Z. However, although incorporating previous spawning history into the catch curve model improved estimates of Z, the models still performed poorly overall; all models produced negatively biased estimates of Z that were highly variable. If catch curves are to be used in an assessment, use of the number-of-previous-spawnings as the independent variable will lead to improved estimates of the instantaneous total mortality rate from a catch curve.

The second approach consisted of a management strategy evaluation, where population projections under different management strategies were compared. Each management strategy shared a common assessment method: a catch curve model incorporating previous spawning history, fit with the Poisson GLM. Four harvest control rules and seven assessment and management schedules were evaluated. All management strategies, regardless of the harvest control rules or assessment and management schedule used, resulted in successful management of the simulated populations; on average, exploitation rates were kept within the target zone and escapements were above target levels. Of the four harvest control rules evaluated, using a full-correction to the center of the fully-exploited zone tended to perform best. Assessing as frequently as possible, but managing less frequently, tended to provide the most reliably sustainable results; assessing every two years and managing every six years was put forward as potentially ideal assessment and management schedule. The overall conclusion of this thesis is that catch curve models, despite being poor estimators of the instantaneous total mortality rate, can still be used as an assessment and management tool for data-limited river herring stocks.

While the management strategy evaluation indicated that Alewife stocks could be managed using catch curves, there are significant caveats associated with this conclusion. First, it was assumed that the sampling for the age structure used to fit the catch curve models was unbiased. Bias could be present in the sample of the age structure, whether due to selectivity of the fishery from which fish are sampled, or from a non-random sampling design. River herring fisheries have been assumed to be non-selective (ASMFC 2012, Gibson et al. 2017), however some fisheries do exhibit some degree of selectivity for older, larger fish when the fishing season does not encompass the entire spawning run (Billard 2017). Using an age sample from a fishery that more readily removes older fish from the population to fit a catch curve will result in under-estimation of Z, by increasing the proportion of older fish observed resulting in a less-steep slope of the fitted model. The magnitude of the bias in the estimation of Z would depend on the degree of selectivity of the fishery. Selectivity could also arise from the type of gear that is used (Millar and Fryer 1999). In this instance, the magnitude and direction of the bias would depend on the type of gear. Producing an unbiased estimate of Z from a stock with a selective fishery would require either an unbiased sampling of fish prior to being caught by the fishery, or an estimate of selectivity to correct the sample of the age structure. The first option would require a great deal of effort, requiring data collection independent of the fishery, and estimates of daily passage to weight daily samples. This would defeat the purpose of using catch curves as a short-term, data-limited approach to assessing river herring stocks. The second option would require additional data collection from the escaped component of the population to estimate selectivity for each year, again defeating the purpose of using data-limited methods, or it would require assuming a selectivity curve for the population. Based on these issues, a strongly selective river herring fishery would not be a good candidate for management based solely on estimates of Z produced by catch curve models.

Non-random sampling of the population's age structure could have an effect on the estimates of Z similar to the effect from selectivity. Little attention has been given to proper sampling of fish to estimate the age structure for river herring. This is a fundamental process that, if done incorrectly, will introduce errors into the estimates of Z produced from the catch curve model. The first issue with obtaining a truly non-random sample pertains to the selectivity discussed above; sampling fish from the fishery introduces bias from gear and availability selectivity, and sampling the spawning escapement upstream from a selective fishery introduces a bias in the opposite direction. Additionally, sampling throughout the duration of the spawning run is imperative, to avoid any bias due to run structure (where older fish run in greater proportions earlier in the run), and some metric of abundance is required in order to scale the sample to be representative of the daily catch or escapement. Simply sampling the same number of fish each day of the run without attributing weights based on the daily total the sample is taken from will result in the samples representing different numbers of fish (Gibson et al. 2017). Samples from the early and later parts of the run will represent fewer fish than those samples taken during the peak of the run. The resulting characteristics of the sample of the population will be skewed towards the fish that run early or late, with less weight being given to the smaller proportion of fish sampled during the peak of the run.

A second major caveat, discussed in Chapter 3, is that catch curves only provide estimates of the total mortality rate. Recruitment failure and changes in the natural mortality rate, both of which could lead to collapse of the stock, cannot be detected using catch curves. Furthermore, basing stock status solely on an estimate of the total mortality rate does not properly address DFO's precautionary approach, which calls for an estimate of the spawning stock biomass or suitable proxy to be compared to a reference point in order to determine stock status (DFO 2006). For these reasons, I recommend using the catch curve methods evaluated herein only for a short-term, data-limited approach for assessing river herring stocks. A longer-term approach requires collecting data for a complete assessment model based on the life history of the species and the characteristics of their fisheries. When designing an assessment program, consideration should be given to the types of data needed for collection and to how much data are required to produce acceptable estimates. The question to focus on then becomes: how much data are required to produce reliable enough estimates of population characteristics to achieve sustainable management? The MSE framework in Chapter 3 can be used to address this question, by comparing management strategies with different assessment methods, ranging from data-limited to data-rich. The statistical catch-at-age model presented by Gibson et al. (2017) as a long-term approach for assessing river herring stocks can be included as a data-rich method in the evaluation of various assessment methods. Violation of the assumptions of unbiased sampling and non-selective fisheries can also be investigated, and management strategies can be tested for their robustness to those violated assumptions. Study of the selectivity of river herring fisheries and sampling bias would further support selection of appropriate management strategies for river herring.

References

Alexander, D. R. and A. H. Vromans. 1983. Status of the Miramichi River gaspereau fishery (1982). Department of Fisheries and Oceans Canada, Canadian Atlantic Fisheries Scientific Advisory Committee Research Document 83/37, Moncton, Canada.

Alexander, D. R. 1984. Status of the Margaree River gaspereau fishery (1983). Department of Fisheries and Oceans Canada, Canadian Atlantic Fisheries Scientific Advisory Committee Research Document 84/23, Moncton, Canada.

Allen, M. S. 1997. Effects of variable recruitment on catch-curve analysis for crappie populations. North American Journal of Fisheries Management 17:202–205.

ASMFC. 2007. Stock Assessment Report No. 07-01 (Supplement) of the Atlantic States Marine Fisheries Commission American Shad Stock Assessment Report for Peer Review Vol. I.

ASMFC. 2012. Stock Assessment Report No. 12-02 of the Atlantic States Marine Fisheries Commission River Herring Benchmark Stock Assessment Vol. I.

ASMFC. 2017. Atlantic States Marine Fisheries Commission River Herring Stock Assessment Update Vol. I: Coastwide Summary.

Barber, B. L., A. J. Gibson, A. J. O'Malley and J. Zydlewski. 2018. Does what goes up also come down? Using a recruitment model to balance alewife nutrient import and export. Marine and Coastal Fisheries 10(2):236-254.

Bergh, M. O. and D. S. Butterworth. 1987. Towards rational harvesting of the South African anchovy considering survey imprecision and recruitment variability. South African Journal of Marine Science 5(1):937-951.

Berry, F. H. 1964. Review and emendation of family Clupeidae. Copeia 1964(4):720-730.

Bigelow, H. B. and W. C. Schroeder. 1953. Fishes of the Gulf of Maine. Vol. 53: Washington D.C.: United States Government Printing Office.

Billard, M. 2017. Age and size selectivity of the Gaspereau River Alewife Fishery: implications for the assessment of this stock. Honours thesis. Acadia University, Wolfville, Nova Scotia.

Butterworth, D. S. and R. A Rademeyer. Statistical catch-at-age analysis vs. ADAPT-VPA: the case of Gulf of Maine cod. ICES Journal of Marine Science 65:1717-1732.

Caddy, J. F. and R. Mahon. 1995. Reference points for fisheries management. FAO Fish. Tech. Pap. No. 347.

Cantoni, E. and E. Ronchetti. 2001. Robust inference for Generalized Linear Models. Journal of the American Statistical Association 96:1022-1030.

Carruthers, T. R., A. E. Punt, C. J. Walters, A. MacCall, M. K. McAllister, E. J. Dick and J. Cope. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research 153:48-68.

Cating, J. P. 1953. Determining the age of Atlantic shad from their scales. U.S. Fishery Bulletin 54:187-199.

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

Chaput, G. and G. Atkinson. 2001. The gaspereau fisheries (*Alosa pseudoharengus* and *A. aestivalis*) of the Miramichi River with updates on the fishery of the Richibucto River of Gulf New Brunswick, 1997 to 2000. Canadian Science Advisory Secretariat Research Document 2001/047.

Chaput, G., P. LeBlanc and R. Crawford. 2001. Assessment of the Margaree River gaspereau fishery, 1997 to 2000. Canadian Science Advisory Secretariat Research Document 2001/046.

Dadswell, M. J., C. Ceapa, A. D. Spares, N. D Stewart, R. A Curry, R. G. Bradford and M. J. W. Stokesbury. 2017. Population characteristics of adult Atlantic Sturgeon captured by the commercial fishery in the Saint John River estuary, New Brunswick. Transactions of the American Fisheries Society 146(2):318-330.

Deriso, R. B., T. J. Quinn II and P. R. Neal. 1985. Catch-Age Analysis with Auxiliary Information. Canadian Journal of Fisheries and Aquatic Sciences 42(4):815-824.

Deroba, J. J., D. S. Butterworth, R. D. Methot, J. A. A. De Oliveira, C. Fernandez, A. Nielsen, S. X. Cadrin, M. Dickey-Collas, C. M. Legault, J. Ianelli, J. L. Valero, C. L. Needle, J. M. O'Malley, Y-J. Chang, G. G. Thompson, C. Canales, D. P. Swain, D. C. M. Miller, N. T. Hintzen, M. Bertignac, L. Ibaibarriaga, A. Silva, A. Murta, L. T. Kell, C. L. de Moor, A. M. Parma, C. M. Dichmont, V. R. Restrepo, Y. Ye, E. Jardim, P. D. Spencer, D. H. Hanselman, J. Blaylock, M. Mood and P.-J. F. Hulson. 2015. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. ICES Journal of Marine Science 72(1):19–30.

DFO. 2001. Gaspereau Maritime Provinces Overview. DFO Science Stock Status Report D3-17.

DFO. 2006. A Harvest Strategy Compliant with the Precautionary Approach. Canadian Science Advisory Secretariat Science Advisory Report. 2006/023.

DFO. 2012. Reference points consistent with the precautionary approach for a variety of stocks in the Maritimes Region. Canadian Science Advisory Secretariat Science Advisory Report. 2012/035.

Duan Y. J., C. P. Madenjian, C. X. Xie, J. S. Diana, T. P. O'Brien, Y. M. Zhao, J. X. He, S. A. Farha and B. Huo. 2016. Age and growth of round gobies in Lake Huron: Implications for food web dynamics. Journal of Great Lakes Research 42(6):1443-1451.

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

Durbin, A. G. S., S. W Nixon and C. A. Oviatt. 1979. Effects of the spawning migration of the alewife, *Alosa pseudoharengus*, on freshwater ecosystems. Ecology 60:8-17.

Fournier, D. and C. P. Archibald. 1982. A general theory for analyzing catch at age data. Canadian Journal of Fisheries and Aquatic Sciences 39:1195-1207.

Fulton, E. A., A. E. Punt, C. M. Dichmont, R. Gorton, M. Sporcic, N. Dowling, L. R. Little, M. Haddon, N. Klaer and D. C. Smith. 2016. Developing risk equivalent data-rich and data-limited harvest strategies. Fisheries Research 183:574-587.

Garman, G. C. 1992. Fate and potential significance of postspawning anadromous fish carcasses in an Atlantic coastal river. Transactions of the American Fisheries Society 121:390-394.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Canadian Atlantic Fisheries Scientific Advisory Committee Research Document 88/29, St. Andrews, Canada.

Geromont, H. F., and D. S. Butterworth. 2015. Generic Management Procedures for Data-Poor Fisheries: Forecasting with Few Data. ICES Journal of Marine Science: Journal Du Conseil 72(1):251-261.

Gibson, A. J. F. and R. A. Myers. 2001. Gaspereau river alewife stock status report. Canadian Science Advisory Secretariat Research Document 2001/061.

Gibson, A. J. F. and R. A. Myers. 2003a. Biological reference points for anadromous alewife (*Alosa pseudoharengus*) fisheries in the Maritime Provinces. Canadian Technical Report of Fisheries and Aquatic Sciences No. 2468.

Gibson, A. J. F. and R. A. Myers. 2003b. A meta-analysis of the habitat carrying capacity and the maximum lifetime reproductive rate of anadromous alewife in eastern North America. In Biodiversity and conservation of shads worldwide. Edited by K. E. Limburg and J. R. Waldman. American Fisheries Society Symposium Series. American Fisheries Society, Bethesda, Maryland. 211–221.

Gibson, A. J. F., H. D. Bowlby and F. M. Keyser. 2017. A Framework for the Assessment of the Status of River Herring Populations and Fisheries in DFO's Maritimes Region. Canadian Science Advisory Secretariat Research Document 2016/105.

Guest, W. C., and R. W. Drenner. 1991. Relationship between feeding of blueback herring and the zooplankton community of a Texas reservoir. Hydrobiologia 209:1-9.

Gulland, J. A. 1965. Estimation of growth and mortality rates. Annex to Arctic Fisheries Working Group Report, meeting in Hamburg, January 1965. ICES Document CM 1965/Gadoid Fish Committee 3:1-9.

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

Hilborn, R. and C. J. Walters. 1992. Quantitative Fisheries Stock Assessment. New York: Chapman and Hall.

Huo, B., C. P. Madenjian, C. X. Xie, Y. Zhao, T. P. O'Brien and S. J. Czesny. 2014. Age and growth of round gobies in Lake Michigan, with preliminary mortality estimation. Journal of Great Lakes Research 40(3):712-720.

ICES. 2012. Report on the Classification of Stock Assessment methods developed by SISAM. ICES document CM 2012/ACOM/SCICOM:01.

ICES. 2013. Report of the Working Group on Guidelines for Management Strategy Evaluation (WKGMSE). ICES HQ, Copenhagen, Denmark. ICES CM 2013/ACOM:39.

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

Jessop, B. M. and H. A. Parker. 1988. The alewife in the Gaspereau River, Kings County, Nova Scotia. 1982-1984. Canadian Manuscript Report of Fisheries and Aquatic Sciences No. 1992.

Jessop, B. 2001. A brief assessment of the commercial fishery on the Lower Saint John River, N.B. DFO Canadian Science Advisory Secretariat Research Document 2001/60.

Kell, L. T., I. Mosqueira, P. Grosjean, J-M. Fromentin, D. Garcia, R. Hillary, E. Jardim, S. Mardle, M. A. Pastoors, J. J. Poos, F. Scott and R. D. Scott. 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science 64(4):640–646.

Loesch, J. G. and W. A. Lund Jr. 1977. A contribution to the life history of blueback herring, *Alosa aestivalis*. Transactions of the American Fisheries Society 106:583-589.

Loesch, J. G. 1987. Overview of the Life History Aspects of Anadromous Alewife and Blueback Herring in Freshwater Habitats; pp. 89-103. In: M. J. Dadswell, R. J. Klauda, C. M. Moffitt, R. L. Saunders, R. A. Rulifson and J. E. Cooper, editors. Common Strategies of Anadromous and Catadromous Fishes. Proceedings of an International Symposium held in Boston, Massachusetts, USA. March 9-13, 1986. American Fisheries Society, Bethesda, MD.

Lent, J. N. 1999. Selective predation by young-of-the-year alewives (*Alosa pseudoharengus*) on zooplankton in Gaspereau Lake, Nova Scotia. B.Sc. (Hon.) Thesis. Acadia University, Canada.

Marcy Jr., B. C. 1969. Age determinations from scales of *Alosa pseudoharengus* (Wilson) and *Alosa aestivalis* (Mitchill) in Connecticut waters. Transactions of the American Fisheries Society 98:622-630.

Maunder, M. N. and G. M. Watters. 2003. A-SCALA: an age-structured statistical catch-at-length analysis for assessing tuna stocks in the eastern tropical Pacific Ocean. Inter-American Tropical Tuna Commission Bulletin 22(5):435-582.

Maunder, M. N. and A. E. Punt. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research 142:61-74.

McBride, M. C., T. V. Willis, R. G. Bradford and P. Bentzen. 2014. Genetic diversity and structure of two hybridizing anadromous fishes (*Alosa pseudoharengus*, *Alosa aestivalis*) across the northern portion of their ranges. Conservation Genetics 15:1281-1298.

McClenachan, L., S. Lovell and C. Keaveney. 2015. Social benefits of restoring historical ecosystems and fisheries: alewives in Maine. Ecology and Society 20(2):31.

McCullagh, P. and J. A. Nelder. 1989. Generalized Linear Models, Second Edition. Chapman and Hall, New York.

McIntyre, T. M., R. G. Bradford, T. D. Davies and A. J. F. Gibson. 2007. Gaspereau River Alewife Stock Status Report. DFO Canadian Science Advisory Secretariat Research Document 2007/032.

Messieh, S. N. 1977. Population structures and biology of alewives (*Alosa pseudoharengus*) and blueback herring (*Alosa aestivalis*) in the Saint John River, New Brunswick. Environmental Biology of Fishes 2:195-210.

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

Mikkelsen, N., T. Pedersen, T. C. dos Santos Schmidt, I-B. Falk-Petersen and A. Slotte. 2016. Are life histories of Norwegian fjord herring populations of Pacific ancestry similar to those of Atlantic or pacific herring? Journal of Marine Systems 180:237-245.

Millar, R. B. and R. J. Fryer. 1999. Estimating the size-selection curves of towed gears, traps, nets and hooks. Reviews in Fish Biology and Fisheries 9:89-116.

Millar, R. B. 2014. A better estimator of mortality rate from age-frequency data. Canadian Journal of Fisheries and Aquatic Sciences 72:364-375.

Mills, E. L., R. O'Gorman, J. DeGisi, R. F. Heberger and R. A. House. 1992. Food of the alewife (*Alosa pseudoharengus*) in Lake Ontario before and after the establishment of *Bythotrephes cederstroemi*. Canadian Journal of Fisheries and Aquatic Sciences 49: 2009-2019.

Morris, W. F. and D. F. Doak. 2002. Quantitative conservation biology. Sinauer, Sunderland, Massachusetts.

Murphy, M. D. 1997. Bias in Chapman-Robson and least-squares estimators of mortality rates for steady-state populations. Fisheries Bulletin 95(4):863-868.

National Research Council. 1998. Improving Fish Stock Assessments. National Academy Press, Washington, DC.

Neves, R. J. 1981. Offshore distribution of alewife, *Alosa pseudoharengus* and blueback herring, *Alosa aestivalis*, along the Atlantic coast. Fishery Bulletin 79:473-485.

Possingham, H. P., D. B. Lindenmayer and T. W. Norton. 1993. A framework for the improved management of threatened species based on population viability analysis (PVA). Pacific Conservation Biology 1(1):39.

Punt, A. E., D. S. Butterworth, C. L. de Moor, J. A. A. De Oliveira and M. Haddon. 2014. Management strategy evaluation: best practices. Fish and Fisheries 17(2):303-334.

Quinn, T. J. and R. B. Deriso. 1999. Quantitative Fish Dynamics. New York: Oxford University Press.

Rachels, K. T. and B. R. Ricks. 2018. Exploring causal factors of spawning stock mortality in a riverine Striped Bass population. Marine and Coastal Fisheries 10(4):424-434.

Rago, P. 2001. Index measures and stock assessment in Atlantic salmon. Stock, recruitment and reference points: assessment and management of Atlantic salmon. INRA Editions. Paris. 137-176.

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

Roux, M-J., R. F. Tallman and Z. A. Martin. 2019. Small-scale fisheries in Canada's Artic: Combining science and fishers knowledge towards sustainable management. Marine Policy 101:177-186.

Rulifson, R. A. 1994. Status of anadromous *Alosa* along the east coast of North America. Pages 108-112 In J. E. Cooper, R. T. Eades, R. J. Klauda and J. G. Loesch, editors. Anadromous Alosa Symposium, Tidewater Chapter, American Fisheries Society, Bethesda, Maryland.

Sánchez S., L. Ibaibarriaga, A. Uriarte, R. Prellezo, M. Andrés, P. Abaunza, E. Jardim, S. Lehuta, L. Pawlowski and B. Roel. 2019. Challenges of management strategy evaluation for small pelagic fish: the Bay of Biscay anchovy case study. Marine Ecology Progress Series 617-618:245-263.

Scott, W. B. and Scott, M. G. 1988. Atlantic Fishes of Canada. Canadian Bulletin of Fisheries and Aquatic Sciences 219:731.

Shaffer, M. L. 1981. Minimum population sizes for species conservation. BioScience 31(2):131-134.

Smith, A. D. M. 1994. Management strategy evaluation: the light on the hill. In: Hancock, D A ed, editor/s. Population Dynamics for Fisheries Management, Australian Society for Fish Biology Workshop Proceedings; August 24-25, 1993; Perth, WA. Perth, WA: Australian Society for Fish Biology; 1994. 249-253.

Smith, A. D. M., K. J. Sainsbury and R. A. Stevens. 1999. Implementing effective fisheries-management systems – management strategy evaluation and the Australian partnership approach, ICES Journal of Marine Science 56(6):967–979.

Smith, M. W., A. Y. Then, C. Wor, G. Ralph, K. H. Pollock and J. M. Hoenig. 2012. Recommendations for catch-curve analysis. North American Journal of Fisheries Management 32(5):956-967.

Stone, H. H. and B. M Jessop. 1992. Seasonal distribution of river herring *Alosa pseudoharengus* and *A. aestivalis* off the Atlantic coast of Nova Scotia. Fisheries Bulletin 90:376-389.

Surma S., T. J. Pitcher, R. Kumar, D. Varkey, E. A. Pakhomov and E. L. Mimi. 2018. Herring supports Northeast Pacific predators and fisheries: Insights from ecosystem modeling and management strategy evaluation. PLoS ONE 13(7):e0196307.

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

Wallace, R. K. and K. M. Fletcher. 2001. Understanding fisheries management, Second Edition. Mississippi-Alabama Sea Grant Consortium. 00-005.

Wayte, S. E. and N. L. Klaer. 2010. An effective harvest strategy using improved catch-curves. Fisheries Research 106(3):310-320.

Weijerman, M., E. A. Fulton and R. E. Brainard. 2016. Management strategy evaluation applied to coral reef ecosystems in support of ecosystem-based management. PLoS ONE 11(3):e0152577.

Wells, L. 1970. Effects of alewife predation on zooplankton populations in Lake Michigan. Limnology and Oceanography 15(4):556-565.