Issues and Importance of Age and Length Data in Fisheries Models

Paul N. Frater

Faculty of Physical Sciences
University of Iceland
2020

Issues and Importance of Age and Length Data in Fisheries Models

Paul N. Frater

Dissertation submitted in partial fulfillment of a
Philosophiae Doctor degree in Ecological Modelling

Advisor
Dr. Gunnar Stefánsson

PhD Committee
Dr. Gunnar Stefánsson (advisor)
University of Iceland
Dr. Bjarki Elvarsson
Icelandic Marine and Freshwater Research Institute
Dr. Rick Methot
NOAA Fisheries

Opponents
Dr. Sarah Gaichas – NOAA Fisheries
Dr. Coilin Minto – Galway-Mayo Institute of Technology

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, June 2020

Issues and Importance of Age and Length Data in Fisheries Models

Dissertation submitted in partial fulfillment of a *Philosophiae Doctor* degree in Ecological Modelling

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavík, Iceland
101, Reykjavik
Iceland

Telephone: 525-4000

Bibliographic information:
Paul N. Frater, 2020, *Issues and Importance of Age and Length Data in Fisheries Models*, PhD dissertation, Faculty of Physical Sciences, University of Iceland, 197 pp.

ISBN 978-9935-9452-6-6

Printing: Háskólaprent
Reykjavik, Iceland, June 2020

Abstract

Age and length (i.e. composition) data are, along with abundance indices, some of the most basic yet informative data collected in fisheries science. Measurements on age and length have been collected since the study of fish and fish stocks began in the 19th century. Composition data are crucial for estimating growth, mortality, selectivity, and, in some cases, recruitment in fisheries stocks. The importance of these data to many types of fisheries models cannot be understated, and yet some of the issues associated with composition data still continue to vex models employed in fisheries.

This thesis focuses on some of the issues associated with composition data in fisheries models as well as the importance of including these data in integrated assessment models. There are five chapters in this thesis; the first two focus on issues in using age and length data to estimate growth in fisheries, the third chapter describes in a novel manner the drivers for spatiotemporal growth for an important fish stock in Icelandic waters, Atlantic cod (*Gadus morhua* L.). The last two chapters focus on the importance of age and length data in integrated assessment models for how they impact both point estimates as well as uncertainty estimates.

Overall, this thesis has advanced the understanding of composition data in fisheries models by: a) revealing new aspects to selectivity bias through aspects of growth that can enhance or increase bias, b) comparing methods to correct for selectivity bias, c) describing in a novel manner the drivers for spatiotemporal growth for an important fish stock in Icelandic waters, d) and systematically testing for the importance of age and length data in both point estimates for parameters of integrated assessment models as well as associated uncertainty estimates.

Ágrip

Gögn um aldur og lengd fiska, ásamt vísitölum úr stofnmælingum, eru undirstöðuupplýsingar í fiskifræði. Mælingum á aldri og lengd hefur verið safnað síðan rannsóknir hófust á fiskum og fiskistofnum á nítjándu öld. Þessi gögn um aldurs- og lengdarsamsetningu eru nauðsynleg til að meta vöxt, afföll, valmynstur og nýliðum fiskistofna. Þrátt fyrir mikilvægi þessara gagna fyrir líkön í fiskifræði eru enn til staðar ýmis vandamál sem fylgt hafa gögnum um samsetningu alveg frá upphafi.

Þessi ritgerð fjallar um helstu mál sem tengjast notkun gagna um samsetningu í fiskilíkönum, auk mikilvægis þeirra í samþættum stofnlíkönum. Í ritgerðinni eru fimm kaflar: Fyrstu tveir fjalla um vandamál sem tengjast aldurs- og lengdargögnum við mat á vexti, sá þriðji fjallar um breytileika vaxtar þorsks (*Gadus morhua* L.) í tíma og rúmi umhverfis Ísland. Síðustu tveir kaflarnir fjalla um mikilvægi aldurs- og lengdargagna í samþættum stofnlíkönum vegna áhrifa gagnanna á bæði punktmat og óvissumat.

Á heildin litið hefur sú vinna sem hér er lýst þróað nokkuð skilning á mikilvægi samsetningargagna með því að a) sýna nýjar hliðar á valmynstursbjaga sem eykur bjaga sem var þó áður þekktur b) samanburði aðferða til að leiðrétta þennan bjaga c) lýsa á nýjan hátt breytileika vaxtar þorsks í tíma og rúmi d) sýna á kerfisbundinn hátt mikilvægi aldurs- og lengdargagna bæði fyrir punktmat og óvissu metla í samþættum stofnmatslíkönum.

Dedication

This dissertation is dedicated to Theodore and Ellis Frater. It is my hope that my current work creates many future opportunities for you both. Remember that you can accomplish anything that you put your mind to.

I also dedicate this work to the memory of my late father, Daniel A. Frater, who always thought I was crazy for wanting to go on for a Ph.D., especially one in Iceland, but was supportive nonetheless.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to sincerely thank my advisor, Gunnar Stefánsson, for his helpful guidance, wisdom, support, and kindness while undertaking this work. Gunnar, you were a great advisor and words cannot express my gratitude for both the educational and cultural opportunities to which you have introduced me. I would also like to thank the members of my doctoral committee, Bjarki Elvarsson and Rick Methot, for their invaluable insight, guidance, and previous work that has paved a path for myself and many others.

I would also like to thank both Dr. Sarah Gaichas of NOAA Fisheries and Dr. Cóilin Minto of the Galway-Mayo Institute of Technology for volunteering their considerable time, knowledge, and efforts to serve as opponents at my defense.

This project was made possible with financial support from the Icelandic Research Fund (Rannís), as well as two European Union projects: MINOUW and MareFrame. This support is greatly appreciated, and the work would not have been possible without it.

I am grateful to Birgir Hrankfelsson for crucial and integral aid in Bayesian statistics that are part of this dissertation. I'm not sure what the probability is that I would have learned Bayesian statistics given that I am an ecologist, but I sincerely thank you for your teaching, assistance, and patience.

In addition, I would like to thank my Icelandic “family" for all of their support while living in a foreign country. In no particular order: Erla Sturludóttir, Joana Ribeiro, Oxana Storozhenko, Narahaim Gonzalez, Pam Woods and family, Giri Gopalan, and Maartje Oostdijk.

Last, but certainly not least, I would like to thank my family. Haley, for all of your loving support and sacrifice that allowed me to follow my dream. I am forever grateful. Theo and Ellis, for being a part of my life. You make everything I do worthwhile.

Part I

Thesis

1 Introduction

Fishing has a major economic impact at local (Béné et al., 2007), regional (Seung and Waters, 2006), and global scales (Dyck and Sumaila, 2010). For example, in 2017 in the United States alone commercial fish landings was valued at $5.4 billion (NMFS, 2018). In Iceland, commercial fishing has a long history (Perdikaris and McGovern, 2009), and has historically been the largest export for the country. Commercial fishing in Iceland has been estimated at $> 25\%$ of the gross domestic product (GDP) at a value of > 150 billion Icelandic krónur (ISK) (Sigfússon et al., 2012). These huge economic benefits are not without cost, though, and the ecological impacts of commercial fishing are also great (Myers and Worm, 2003; Crowder et al., 2008; Kraus and Diekmann, 2018). About a third of the world's fisheries are still considered overfished (FAO, 2018), and increases in technology have made fishing (and overfishing) much easier than historically possible (Finkbeiner et al., 2017). Overfishing, along with other human impacts, also have the potential to reduce biodiversity in the world's oceans, which in turn can decrease vital ecosystem services (Worm et al., 2006). It goes without saying that careful management of such economically salient and ecologically important resources is both crucial and highly beneficial (Fulton et al., 2019).

Nonetheless, great advances have been made in the last century with respect to estimating populations of fishes, the status of stocks, and sustainable methods for maximizing yields. Mathematical descriptions of fish populations began with the derivation of the Baranov catch equation in 1918 (Baranov, 1918). Substantial contributions were made by Beverton and Holt (1957) and Ricker (1954) in the mid-20th century to describe population growth in terms of individual growth rates, fishing intensity, and reproduction curves. The last two decades especially have seen detailed and powerful formulations for modeling both fish populations (Fournier et al., 2012; Methot and Wetzel, 2013) as well as entire ecosystems (Christensen and Walters, 2004; Fulton et al., 2011).

Stock assessment models have traditionally been some of the most advanced and complex models for their time in all of ecology (Bolker et al., 2013; Maunder and Piner, 2015), and are often some of the first to utilize or develop advancements in statistics (McAllister and Ianelli, 1997; Punt and Hilborn, 1997) or modeling software (Fournier et al., 2012). Basic stock assessments started with virtual population analysis (VPA; Fry, 1949) and simple models such as surplus production models (Pella and Tomlinson, 1969), which used minimal data such as catch-at-age and often inferred age outside the model using age-length keys. In recent decades a greater push has been made to using integrated assessment models, which incorporate the processes of growth, reproduction, natural mortality, and both fishery and survey selectivity into the model and can include many different data types (Maunder and Punt, 2013). These models often combine age, length, abundance indices, and catch data (among potentially other types) and use fairly sophisticated likelihood and often weighted likelihoods to estimate the parameters of the various sub-components of the model. Gadget is but one example of these types of models (see section 3.2 below, Begley, 2012). Stock Synthesis 3 (SS3) is another example (Methot and Wetzel, 2013). While the complexity of these approaches has led to sophisticated models used to assess stocks and guide management of fishes it can also lead to problems. Stock assessment models have been criticized as overly complicated (Cotter et al., 2004), and the use of multiple data sets can lead to conflict among sources of data (Francis, 2017). In addition, these models tend to be “data hungry" (Geromont and Butterworth, 2015) often requiring large amounts of data across lengthy time series. Of the data that are used in stock assessment models, composition data (i.e. age data and/or individual (or mean) size such as length or weight) are arguably the most important and often some of the most expensive and intensive to collect (He et al., 2016). Size data, which represent some aspect of body size in fish (i.e. length or weight), are data that are relatively easy and inexpensive to collect. Regardless, for collection of fishery-independent size data, standardized surveys must still be performed (e.g. standardized trawl surveys), which can be quite intensive and costly. Once fish are caught on trawl surveys there are usually many that can be relatively easily weighed or measured for length. Age data, on the other hand, generally require preparation and analysis of hard anatomical parts (i.e. otolith, scale, spine, etc.) to estimate age of fish, which is a much more expensive and intensive endeavor relative to collecting size data. Therefore, age data are generally lacking relative to size data. Paired length-at-age data are important for determining growth of fish, so age data are generally taken as a subset of length data from fish in a survey.

1.1 Definition of size as used in fisheries models

The term size-at-age is often used in context of fisheries models, but this can mean different things. Size can be used to refer to either the length or weight of a fish. Length is a measure of how long a fish is from anterior to posterior (measured in cm, mm, etc.). There are generally three different types of length measurements (Froese and Pauly, 2017a): standard length (anterior tip to base of the caudal fin), total length (anterior tip to the posterior end of the caudal fin), and fork length (anterior tip to the posterior end of the middle caudal rays). Weight is a measure of the biomass of fish (in g, kg, tons, etc.). In the case of composition data the terms size, weight, and length are used to refer to the respective measurements on individual fish or as a group mean (e.g. mean length at age 1). This is in contrast to, say, landings data from commercial catches which are often reported as total landings in metric tons or hundreds of thousands of metric tons. In this dissertation length and weight will be used to refer to the respective definitions above. Size will be used to describe either length or weight ambiguously.

1.2 Importance of Composition Data in Stock Assessment Modeling

As mentioned above, stock assessment models rely on a number of disparate datasets that work together to predict numbers, age, and growth of fish. These models are often referred to as integrated assessment models due to the variety of data types that are used (Maunder and Punt, 2013). Integrated assessment models typically are age- and/or size-based and often rely on three types of data: commercial catch (or harvest), indices of abundance, and either or both of size and age data (collectively referred to as composition data). While catch and abundance data are useful for determining overall trends of a stock, composition data is crucial for the estimation of growth (He et al., 2016; Lorenzen, 2016), selectivity (Crone and Valero, 2014), and recruitment (Maunder and Punt, 2013; Maunder and Piner, 2015). Growth is often estimated using length-at-age in integrated assessment models, and the inclusion of paired age-length data is essential for parameterization of this component of a model. The conversion to weight can then be done either outside or inside the model using a length-weight relationship, which requires paired length-weight data. Selectivity can be implemented as either length- or age-based with age-based perhaps being more common. Gaining an understanding of length-based selectivity requires use of length data for both harvested fish and any fishery-independent surveys that are performed. Length distributions from empirical data can be compared to those from the output of models and used as a likelihood to inform selectivity in integrated assessments. Lastly, including composition data is important for estimating recruitment especially in the case of time-varying recruitment. Strong year classes of fish can often be detected in length distributions as the proportion of lengths in that cohort can be seen moving across years.

Integrated assessment models typically produce numbers of fish at certain ages and length categories (Methot and Wetzel, 2013), but data on commercial catch are often recorded as biomass of landings. For models to determine how many fish were harvested in a particular year or timestep a translation has to be made from weight to numbers at age and size. This is accomplished via one of two approaches: either 1) a length-weight relationship and an age-length key or conditional-age-at-length (CAAL) data (i.e. estimating a growth curve via paired age-length data Francis, 2016), or 2) empirically by using mean size across either age or both age and year (Francis, 2016; Lorenzen, 2016). While the empirical approach is still commonly used especially in data-rich scenarios (Lorenzen, 2016), the growth curve approach is increasingly employed as it more closely follows principles of the integrated assessment approach, which encourages the use of raw data whenever possible (Maunder and Punt, 2013). Either of these approaches require composition data; the empirical approach to estimate mean size-at-age outside of the model, or the growth curve approach to allow the model to estimate growth curve parameters internally. Some sort of data that is representative of size-at-age are essential for a model to appropriately determine growth if it is to be estimated in the model, which is recommended (Maunder and Punt, 2013; Francis et al., 2016). There are three potential data types that are used to accomplish this: 1) CAAL data, 2) tagging data, and 3) length data (Francis et al., 2016). CAAL data are probably the most common (Lorenzen, 2016), but some fisheries organisms, such as invertebrates (i.e. crustaceans, cetaceans, and mollusks), lack the hard parts required for aging (i.e. otoliths or scales Punt et al., 2013); therefore, growth estimation methods that use tagging data or simply just length distributions have been developed (Basson et al., 1988; Froese et al., 2018). In some instances growth can be determined from length data alone, but this is dependent on life-history as year classes must be distinguishable using length distributions. The presence of at least some CAAL data generally aids in the estimation of parameters such as growth (Ono et al., 2015; He et al., 2016). CAAL data provides a direct relationship between age and length, which are almost always the arguments and outputs, respectively, of growth functions when they are estimated within stock assessments.

Composition data are also important for accurately estimating parameters of the selectivity function used in stock assessment models (Crone and Valero, 2014; Maunder et al., 2014). Both CAAL data and length frequency data can aid in the estimation of selectivity, and an understanding of selectivity is crucial to fitting correct stock assessment models as it directly connects process to data in these models (Frater and Stefansson, 2019, 2020). For example, length data that are collected, either by fishery-independent surveys or commercial fleet samples, have been sampled under selectivity. This imposes a bias in the data themselves as shown in papers I and II in part II of this dissertation. In other words, data that represents growth (i.e. length-at-age data) is altered by the process of selectivity. Indeed, selectivity is the process that directly imposes bias on length-at-age data (Sampson, 2014; Francis, 2016; Frater and Stefansson, 2019), and in a similar way it imposes a direct distortion on model output which is used to compare model to data. This is different, although tangentially related, to Rosa Lee's phenomenon (Taylor and Methot Jr, 2013). The integrated assessment model Gadget assigns growth via one of several functions (Begley, 2012), but length proportions that are compared to sampled data are biased by the selectivity functions that are placed on the sampling fleets. Therefore, it is important to have a resolution of composition data to specify correct growth curves, but also to properly specify selectivity in integrated assessments as well.

Lastly, composition data can also indicate the strength of year classes across time, and hence aid in the estimation of recruitment within models. Strong year classes can be seen across length distributions as the lengths associated with a particularly strong year class tend to dwarf the age classes around that class. These patterns may be detectable by length proportion likelihoods when fitting an integrated assessment model. Additionally, age proportions will appear different across years that have particularly high recruitment as there will be proportionally more individuals for a particular cohort that had high recruitment. The signals in age proportions, similar to strong year class signals in length distributions, should also aid in the fitting of integrated assessment models.

Given the above reasons for the importance of composition data in stock assessment models it is generally agreed upon that at least some fishery-independent composition data is beneficial for fitting these models accurately (Wetzel and Punt, 2011; Ono et al., 2015).

1.3 Issues with Composition Data

While composition data are important for accurately fitting stock assessment models they are not without issues (Francis, 2011; Maunder and Punt, 2013; Francis, 2017). There exist four main problems with composition data: 1) selectivity bias, 2) data conflict, 3) time-varying processes, and 4) cost of collection.

1.3.1 Selectivity Bias

All length data (including the length portion of CAAL data) are inherently biased by the selectivity of the sampling process. This issue makes up a large portion of this dissertation, and will be discussed in more detail in part II of this dissertation (papers I and II). The central tenet of the premise of selectivity bias is that sampling inherently skews the length distributions of a sampled age-class by unintentionally preferentially selecting larger or smaller sizes than what is representative of the true length distribution for that particular age (Sampson, 2014). Selectivity is the probability of sampling fish of a certain size or age (either by a fishery or a scientific survey), and is a product of three processes: contact selection, availability, and population selection (population selection is the net effect of contact selection and availability two Sampson, 2014). Contact selection is relatively straightforward – fish of a certain size are either caught or not in fishing or sampling gear due to either size limitations (e.g., smaller fish swim through nets or trawls of a certain mesh size, Somerton et al., 2011) or behavioral traits (e.g., trawl avoidance, Handegard and Tjøstheim, 2005; Ryer, 2008). For example, flatfish have a life-history predation strategy of burying themselves in sandy substrate to ambush prey, which might make them less available to entry into a trawl net (Ryer, 2008). Availability is simply how available fish of a certain size are to fishing gear. As an example, certain fish occupy different spatial areas throughout their life rendering them unavailable to certain gear types at portions of their existence. Population selection is a slightly more complex concept that can combine contact selection and availability and often emerges at larger spatial scales as the result of fish of different sizes or ages using different spatial areas or habitat types (Sampson and Scott, 2011). Sampson (2014) showed that even in a fishery where all fleets imposed asymptotic selectivity on fish the overall population selection was dome-shaped under certain fishing scenarios. This phenomenon was due to different size or age classes utilizing different spatial areas in the model. The entire effect of selectivity can be as simple as small fish passing through the mesh sizes of a net or as complex as spatial disjointedness among sampling fish of different ages interacting with the behavioral changes that come as fish age (Sampson and Scott, 2011). Estimating both growth and selectivity, as is done in Gadget, can potentially alleviate this to some degree, but the issue still persists in the data used by the model and may cause misspecification in growth and/or selectivity as the two may be confounded.

1.3.2 Data Conflict

Another potential problem that can arise with the use of composition data is that of conflicting information among the different types of data sources (Maunder et al., 2017). Part of the issue with data-conflict has to do with model misspecification, which can arise due to a lack of knowledge about certain processes (e.g. using asymptotic selectivity when the true selectivity is dome-shaped or vice-versa Ichinokawa et al., 2014). Since all models are just simplifications of reality and are inherently misspecified there can exist conflicting information among different types of data sources (Maunder and Piner, 2017) notably abundance indices and composition data (Francis, 2011; Lee et al., 2014). For example, since both composition data and abundance data are used in integrated assessment models, if one data type suggests a population trajectory that is different than the other, then the two sources of data are conflicting. This can be seen when one data type is best fitted with using a value of some important parameter (for example, the parameter that estimates virgin recruitment), and another data set is best fitted under a different value of that same important parameter (Francis, 2017). The result of this is that different types of data can inform a model about the trajectory of a stock in opposing directions, and one must then decide which data set to believe (Francis, 2017).

1.3.3 Time-varying processes

Misspecification can also occur due to the difference between process and model setup across time (Martell and Stewart, 2014). For example, growth, selectivity, and recruitment all tend to vary across time (Sampson and Scott, 2012; Lorenzen, 2016; Webber and Thorson, 2016), but many stock assessment models use constant functions for the entire duration of the model. In paper III I show that growth of Atlantic cod (*Gadus morhua* L.) in Icelandic waters does, in fact, vary across time as well as space.

This is simply further evidence that time-varying processes are the rule rather than the exception, and ignoring this variation could lead to model output that may or may not be consistent with the true processes in nature. The issue of time-varying processes and how composition data interacts with them has been at the forefront of many studies in the literature as of late (Maunder et al., 2014, 2016), and many solutions to these issues are being implemented in stock assessment models (Methot and Wetzel, 2013). Most of these solutions incorporate either estimating separate values for annual parameters (as in the case with time-varying recruitment in Gadget), or, as is becoming more common, implementing semi-parametric methods that fit a main function with values that are allowed to stray from the curve produced by the main function (Thorson and Taylor, 2014).

1.3.4 Collection of Composition Data

Lastly, composition data, especially those that are fishery-independent, can be time-consuming and expensive to collect (Wallis and Flaaten, 2001), and fisheries-independent data have not often been collected since the start of a fishery or, in many cases, began after a fishery was already collapsed. There are many instances of either data-poor stock assessments (Dowling et al., 2019; Prince and Hordyk, 2019), or even where fisheries remain un-assessed altogether (Costello et al., 2012). The lack of composition data in data-poor scenarios can lead to bias in fisheries models (Chen et al., 2003; Wetzel and Punt, 2011; Ono et al., 2015), and while including some fisheries-independent data is necessary, an appropriate amount of either age or length data for sufficiently reducing bias is largely unknown. Given the cost of collecting such data it would be useful to know how much composition data is needed for accurate and reliable biomass estimates of a stock. Papers IV and V in part II of this dissertation attempts to gain an understanding of what amount of both age and length data are necessary integrated assessment models by performing simulations for three life history types.

2 Objectives

The objectives of this Ph.D. project were to study causes and consequences of certain issues relating to both age and length data used in fisheries models. Specifically, issues relating to selectivity bias, spatiotemporally-varying growth, and the importance of age and length data sample sizes in stock assessment models were assessed. The specific objectives for individual papers, presented in part II of this dissertation, are as follows:

Paper I: Paul N. Frater and Gunnar Stefansson. 2019. Aspects of both growth and selectivity affect growth parameter estimation bias. *Fisheries Research*, 212:154–161.

The aim of this paper was to test which aspects of both the process of growth in fishes and the effect of selectivity in fisheries (or sampling) are most important in causing selectivity bias. This paper expands upon the concept of selectivity bias to show that aspects of growth can feedback to induce or enhance selectivity bias in length-at-age data.

Paper II: Paul N. Frater and Gunnar Stefansson. 2020. Comparison and evaluation of approaches aimed at correcting or reducing selectivity bias in growth parameter estimates for fishes. *Fisheries Research*, 225:105464

This paper presents a comparison of six selectivity bias correction methods for estimating growth parameters. This comparison uses a simulation study approach on three different life-history types with both asymptotic and dome-shaped selectivity to see what methods are most effective at reducing bias in growth parameter estimates using length-at-age data.

Paper III: Paul N. Frater, Birgir Hrafnkelsson, Bjarki Elvarsson, and Gunnar Stefansson. 2019. Drivers of growth for Atlantic cod (*Gadus morhua* L.) in Icelandic waters – a Bayesian approach to determine spatiotemporal variation and its causes. *Journal of Fish Biology*, 95(2):401–410.

The aim of this study was to estimate growth parameters of Atlantic cod (*Gadus morhua*) in Icelandic waters using a Bayesian approach to determine the appropriate growth model. This study looks at the effects of cohort, year, area, and gender on growth of cod and seeks the most parsimonious model in order to determine if growth is time- and/or space-varying as well as if differences in growth exist between genders. Additionally, parameters from the best fitting models are correlated with potential drivers of growth in Atlantic cod.

Paper IV: Paul N. Frater and Gunnar Stefansson. Systematically testing the relative importance of age and length composition data in an integrated assessment model.

The objective of this paper was to determine what level of age and length data are appropriate in order to accurately estimate the biomass of a stock using a basic stock assessment model. This study presents a systematic test using a simulation study approach to determine how many samples of either age or length data are required for different life-history strategies to sufficiently reduce bias in stock assessment models.

Paper V: Paul N. Frater and Gunnar Stefansson. Effects of reduced age data on uncertainty estimates in an integrated assessment model.

This paper differs from paper IV above in that paper IV seeks to determine sufficient composition data sample sizes for accurate point estimates of stock assessment model parameters. The objective of this study is to discover the effects of reduced composition data, specifically reduced age data, on the uncertainty of parameter and biomass estimates in stock assessment models.

3 Methods

This section details the methods of study used throughout this dissertation. Most of the papers listed in section 2 utilized a simulation study of some sort, and those simulations were created contextually for the papers listed in section 2. These simulations generally fall into two categories: 1) age-at-length simulations and 2) stock assessment simulations, and are useful in the sense that true values for parameters or reference points are known allowing a comparison of models to known values. Additionally, there are a few examples of simulations that utilized actual data in this dissertation, which will be referred to as data simulations. In these simulations true parameter values or reference points are unknown, but examples can be illustrated with real data allowing for a more realistic proof-of-concept. Lastly, one example of strict data analysis is included in paper III.

3.1 Simulation Studies

A majority of the projects included in this dissertation relied on the use of simulation studies. A simulation study is a method whereby a scientist or modeller uses the output from one model–known as the operating model–as “data" for other models–known as the estimation models (Rochet and Rice, 2009). This allows a flexible setup where an endless variety of scenarios can be tested. Simulation studies are a useful abstract for testing ideas or concepts in models as they allow direct comparison of model results to known values from operating models. While the use of simulation studies is widespread in fisheries their results must be interpreted cautiously as they are abstractions of abstractions (i.e. model of a model) as opposed to abstractions of reality. Simulation studies have been criticized for being unrealistic and overly simplistic, which are both true (Rochet and Rice, 2009).

3.1.1 Age-length simulation studies

Papers I and II simulation studies were used to explore aspects of selectivity bias on length-at-age data and possible corrections to this bias. The first of those papers assessed the importance of both the process of selectivity and the process of growth on the phenomenon of selectivity bias. Selectivity bias is a process that skews sampled length-at-age data and is detailed more in papers I and II.

To study the effect of both growth and selectivity on selectivity bias a simple age structured population was simulated as $N_a = N_0 e^{-0.2a}$ where N_a is the number at age a and N_0 is the number at age 0. Each individual at age a was assigned length l as: $l_a \sim \mathrm{N}(L(a), \sigma_a)$, where $L(a)$ is the simple von Bertalanffy growth function (VBGF) $L_a = L_\infty(1 - exp(-k(a - t_0)))$ and σ_a is a hockey stick function of age where:

$$\sigma_a = \begin{cases} a & a \leq 5 \\ 5 & a > 5 \end{cases}$$

These lengths-at-age were then sampled with three different types of selectivity curves: asymptotic, dome-shaped, and constant. The various parameters for the selectivity functions and the VBGF parameters were altered both independently and simultaneously and a VBGF fit to those samples. The asymptotic selectivity curve used was the logistic:

$$S(l) = \frac{1}{1 + \exp(-\alpha(l - l_{50}))}$$

where $S(l)$ is the selectivity at length l and ranges between 0 and 1, α is the parameter that controls steepness of the curve, and l_{50} is a horizontal scaling parameter that controls the inflection point of the curve.

To sample lengths using dome-shaped selectivity the following function was used:

$$S(l) = \left(\frac{l}{(\eta - 1)\beta\gamma}\right)^{\eta - 1} \exp\left(\eta - 1 - \frac{l}{\beta\gamma}\right)$$

where $S(l)$ and l are the same as in the logistic function and η, β, and γ are parameters controlling the steepness and positioning of the domed curve. The constant selectivity function was just a constant value of 1 (i.e. all data were used to fit VBGF parameters).

In this basic age-length simulation study the age structured population with assigned length is the operating model, and the fitted VBGF is the estimation model. Both the OM and the EM were generated, sampled, and fitted in the R environment for statistical computing (R Core Team, 2017) using the above equations. Further details are given in papers I and II.

This approach was used in both papers I and II. In the first of these papers the various parameters in the operating model were altered both independently and simultaneously in an effort to understand which aspects of both growth and selectivity contribute to selectivity bias in length-at-age data and the resulting VBGF parameter estimates in the estimation model. In the latter of these two papers the parameters of the operating model were altered to test how well a derived bias correction method was able to reverse selectivity bias when fitting VBGF parameters.

In both cases the operating models were the same, but the output from them were used in different ways to test different concepts illustrating the flexibility of using simulation studies to gain an understanding about a process. More detail about the methods is given in papers I and II.

3.1.2 Stock assessment simulation studies

Papers IV and V both explored aspects of the importance of and issues with age and length data within statistical catch-at-age models. The modeling software used for both operating and estimation models in these studies was Gadget (see section 3.2 below).

In all the papers listed in this category of simulations studies three different life-history types were used each with three different fishery development scenarios. The life-history types were based on Atlantic cod (*Gadus morhua*), European plaice (*Pleuronectes platessa*), and capelin (*Mallotus villosus*). These species exhibit different life history strategies and represent a range of parameter values for the functions used to model those processes. Additionally, they are all important species both commercially and ecologically in Iceland.

The fishery development scenarios that were used in these stock assessment simulation studies are three commonly used scenarios – fish down, flat MSY, and two way trip. The fish down scenario represents a situation where fishing pressure on a stock increases linearly throughout duration of time in the operating model. The flat MSY scenario is one in which the fishing level that achieves maximum sustainable yield (F_{MSY}) for a stock is employed throughout the duration of fishing in the model. The two way trip scenario begins with increasing fishing pressure that is relieved at some point in the model and decreases to F_{MSY}.

Data for these stock assessment scenarios was taken from operating models generated in Gadget, and the parameters used for operating models as well as sampling output is detailed in the methods of papers IV and V.

3.2 Gadget

Gadget stands for the **G**lobally **A**pplicable area-**D**isaggregated **G**eneral **E**cosystem **T**oolbox, and is a Unix-based modeling software geared towards statistically fitting catch-at-age models. While it is set up well to perform single species stock assessments, it is also flexible enough to fit multi-species models via trophic interactions. Gadget is a forward simulation model that uses standard fisheries equations (e.g. Baranov catch equation, various growth functions, Beverton-holt or Ricker recruitment, *et al.*) to simulate a stock (or stocks). Output from the stock(s) is then compared to data through the inclusion of likelihood components, and a minimization of the sums-of-squares is achieved via one of several algorithms. The total likelihood score is a weighted sum of the individual likelihood components each of which compare output from the model to a specific data type.

A Gadget model can be as simple as a single stock residing on a single area with one commercial fleet, or as complex as multiple interacting stocks residing in any number of areas where migration may or may not occur and with many different fleets.

3.3 gadgetSim–an R package for Gadget simulation studies

While Gadget is useful for modeling fisheries stocks it provides no capability for visualizing or analyzing data, model output, or model fit. To this end Elvarsson and Lentin (2018) created Rgadget, which is an R package for setting up Gadget models, reading output, and visualizing and analyzing model results. This package is incredibly useful for fitting models (including capability to implement the iterative re-weighting approach of Stefansson (2003)) as well as obtaining uncertainty estimates using a spatial resampling bootstrap approach (Elvarsson et al., 2014). While this package is designed for setting up and fitting models to data it is not well-suited for setting up operating models or a large number of estimation models for use in a simulation study framework. To this end I created an R package called gadgetSim (Frater, 2018b). This package is an efficient suite of R functions that can easily set up stocks and fleets for an operating model, sample output from operating models to use as data in estimation models, and quickly create a large number of replicates across many different scenarios to which testing is desired. This package complements Rgadget, which is constructed to fit models and obtain uncertainty estimates, by creating a large number of Gadget models across scenarios and replicates and using Rgadget to fit each replicate. gadgetSim also contains functions to easily extract model output from across simulations to allow for the analysis of a wide variety of models with multiple replicates and across scenarios. gadgetSim would not be useful for setting up and fitting a Gadget model to real data as would be done in a stock assessment, for example. Its niche is in the simulation study framework. This is pertinent as there has recently been a call for more simulation studies using the Gadget framework (Elvarsson et al., 2018). These simulation studies will be important to understand the strengths and weaknesses of the model if it is to be more widely used.

3.4 Data Simulations

Data simulations are wholly different than simulation studies in that real data are used to illustrate or examine a concept or idea. For example, in paper V real data on Atlantic cod (*Gadus morhua*) are used in a bootstrap approach to get uncertainty estimates for model parameters, but in each scenario the amount of age data are reduced and the distribution of parameter estimates assessed. The true values are not known, but an illustration of how uncertainty estimates change with decreasing age data can be made.

All else being equal the expectation would be that model parameters become less certain (i.e. wider confidence intervals) with decreasing age data. This approach is somewhat similar to a sensitivity analysis where model variables or assumptions are changed. In this particular case, though, instead of changing model parameters it is the quantity of data that is changing. More information about the specifics of this particular method can be found in the Methods section of paper V.

The other example of a data simulation in this dissertation comes from paper II. Here the selectivity bias correction method developed is employed on an actual stock as an example of this method used on real data.

3.5 Data Analysis

There is one example in this dissertation of a paper that relies on traditional analysis of actual data collected from real specimens. Paper III presents a Bayesian analysis of spatio-temporal differences in growth of Atlantic cod (*Gadus morhua*) captured in Icelandic waters. This study uses Markov chain Monte Carlo (MCMC) to fit a suite of growth models to cod and determine if growth in cod varies by time, space, and gender. Models are compared using deviance information criterion (DIC) to assess the best fitting growth model to paired length-at-age data. These data were from 1960–2007 and were collected by the Icelandic Marine and Freshwater Research Institute as both bi-annual standardized bottom trawl surveys or port samples from commercial landings.

The Bayesian models in this study fit a VBGF to paired age-length data, and models that incorporated year, cohort, area, gender, and their combinations were compared using DIC. For each model, VBGF parameters were fit to data specific to the time, area, and gender subsets appropriate to each model. For example, a model that incorporated year and area returned a posterior distribution for each VBGF parameter within each area in each year. This approach allows for fitting a number of different models that look at various aspects of spatio-temporal growth, but still allow for easy comparison among whole models to determine the one that obtains the best fit to the data.

A more detailed description of the data, models and model fitting that are used in this study, including a description of the Bayesian posterior distribution, are given in the Methods section of paper III.

4 Results

The overall results from papers I–V are summarized and presented in this chapter. The results in this section are brief and much abridged. Each respective paper should be consulted for more detailed results.

4.1 Issues with composition data

4.1.1 Interactions of growth and selectivity on sampling bias of length data

The results from paper I indicates that composition data are rife with issues. Composition data are inherently biased to begin with from the effects of length-based sampling selectivity, and the results from paper I indicate that this effect can be induced or enhanced based on not only the shape of the selectivity curve but also the life history of the species being sampled. Life-history types with slower growth and greater overlap among age distributions tended to have more biased growth parameter estimates (Fig. 4.1–4.2), and selectivity curves that ascended or descended through a distribution of lengths at a particular age tended to bias length samples to a greater degree. However, the greatest bias was shown in samples that had a large overlap in the distrbution of lengths at different age groups along with a shallowly sloping curve (Fig. 4.3).

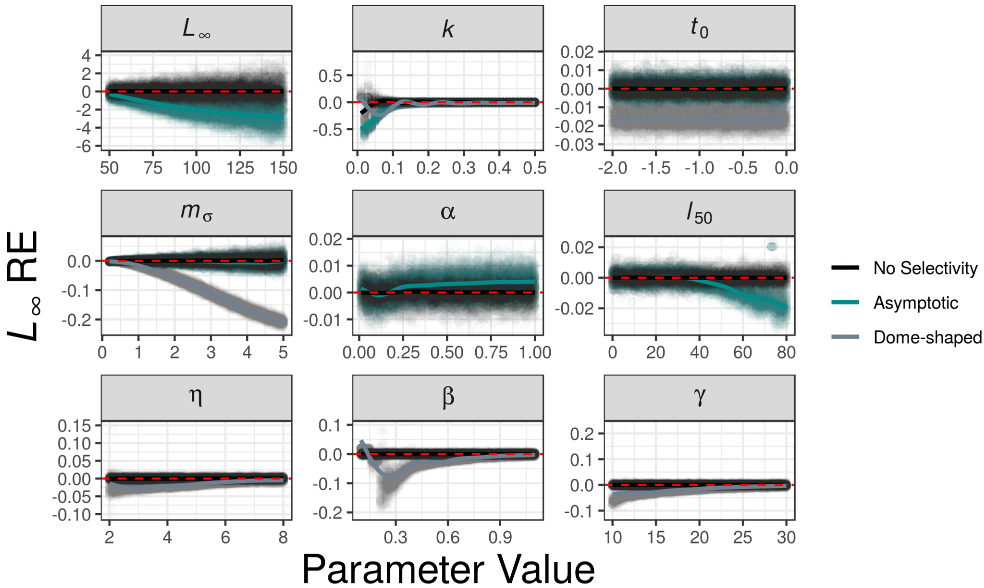


Figure 4.1. Relative error (RE) of L_∞ parameter estimates for the VBGF across a range of values for different parameters. Error is compared to the known, true value. Parameters are described in greater detail in table 6.1.

4.1.2 A comparison of bias correction methods

Paper II compared likelihood-based correction methods for reducing selectivity bias when estimating growth curve parameters. Six methods are compared, and the one that performed the best at correcting selectivity bias when estimating growth curves of length-at-age data was that of Troynikov (1999, Fig. 4.4), which uses Bayes' rule to derive a likelihood equation to correct selectivity bias. The method of Schueller et al. (2014) was a fairly close second to reducing the effect of selectivity on correct estimation of growth curve parameters.

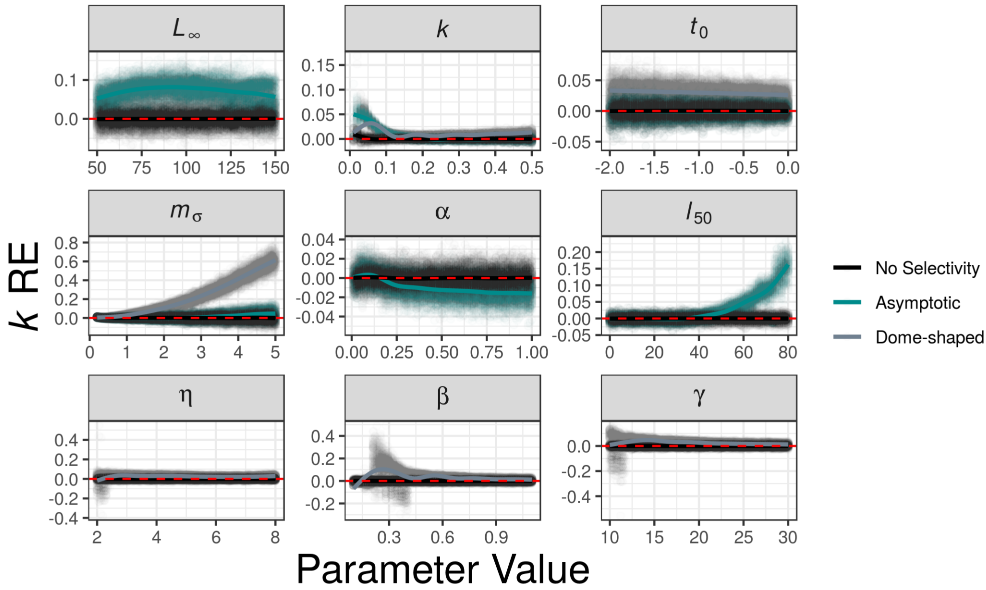


Figure 4.2. This figure is the same as 4.1, but for estimates of the k parameter of the VBGF.

4.1.3 Estimating spatiotemporal growth of Atlantic cod in Icelandic waters: Time-varying growth using age and length data

In paper III, a Bayesian approach was used to estimate growth parameters for Atlantic cod *Gadus morhua*. The best fitting model was that which incorporated both time and space into the growth parameter estimates (Fig. 4.5), which showed that not only does growth in Atlantic cod vary across time but that growth is different in spatial units of the waters surrounding Iceland as well. Although entirely different than the issues presented in subsections 4.1.1 and 4.1.2, the model results presented here reveal another issue with composition data that must be dealt with; the problem of spatiotemporally-varying growth.

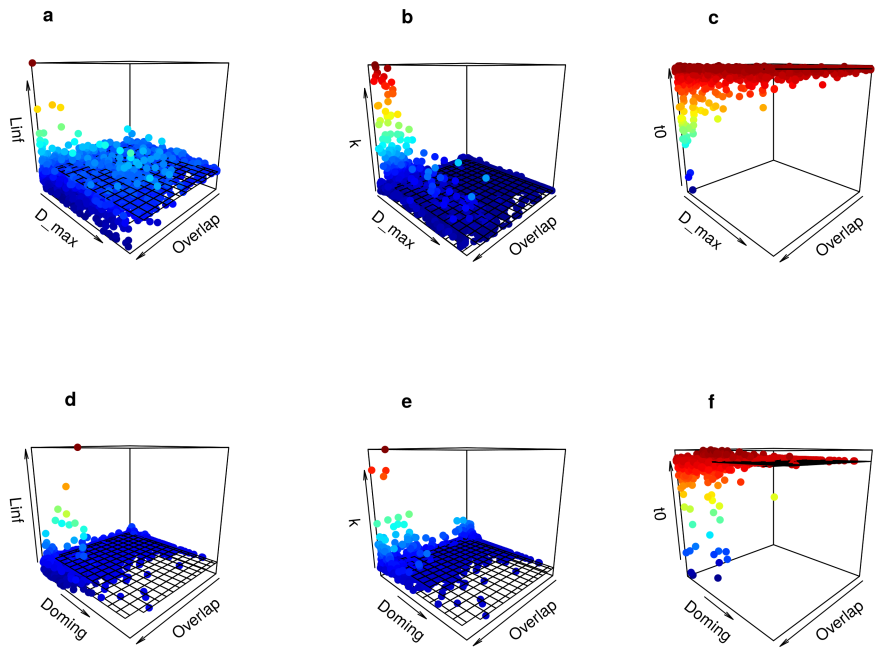


Figure 4.3. Error in growth parameters L_∞ (a,d), k (b,e), and t_0 (c,f) from simulations with randomly chosen growth and selectivity variables for asymptotic selectivity function (a-c) and dome-shaped selectivity function (d-f). Red dots are greater in value while blue dots are lesser in value. "Doming" refers to the metric used to calculate the level of doming in the gamma selectivity function. "Overlap" refers to the combined amount of overlap among l ength distributions from one age to the next. Data in the figures are subset to exclude extreme outliers to better visualize general trends.

4.2 Importance of composition data

While section 4.1 illustrated some of the problems associated with composition data, this section shows some of the results for how important composition data are for use in integrated assessment fisheries models.

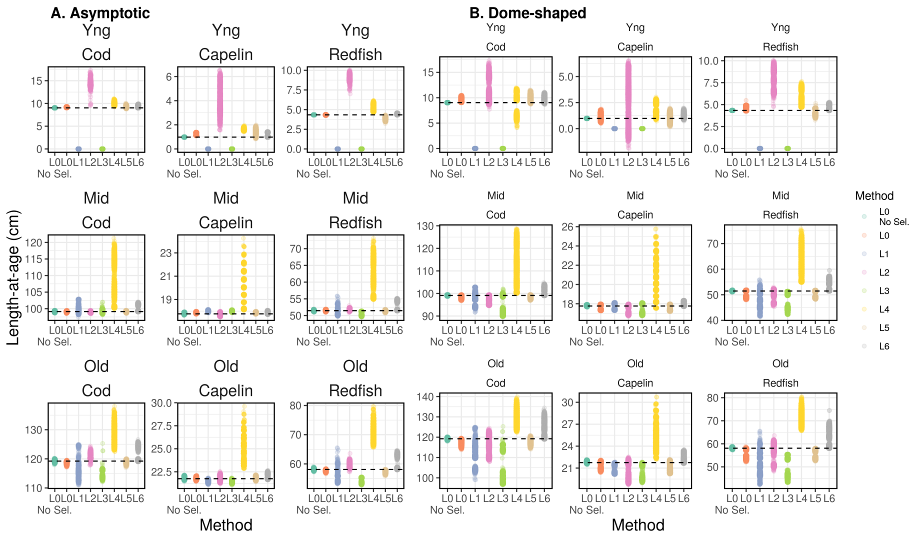


Figure 4.4. von Bertalanffy values for young, mid-aged, and old fish (defined separately for each life history) estimated using 7 different likelihood types for both A) asymptotic and B) dome-shaped selectivity. The methods are described in more detail in paper II, but are as follows: 1) L0 - no correction; L0 No sel. – no correction, no selectivity; 2) L1 – Fix t_0, 3) L2 – remove small sample sizes; 4) L3 – remove small sample sizes and fix t_0; 5) L4 – fix L_∞; 6) L5 – likelihood truncation method of (Schueller et al., 2014); 7) L6 – selectivity inclusion method (Troynikov, 1999).

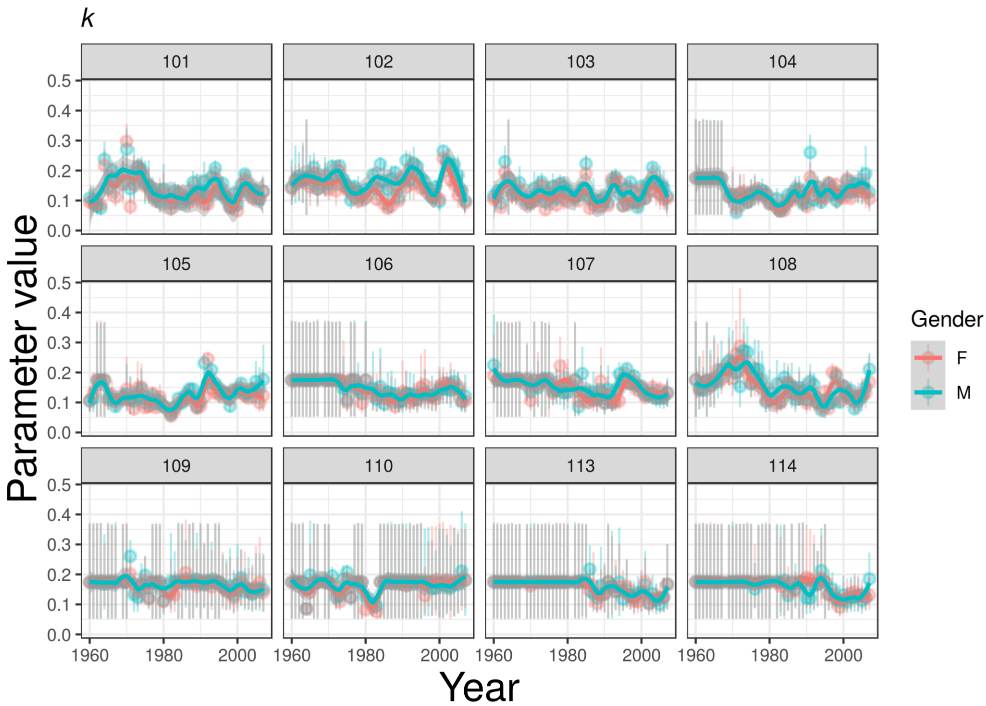


Figure 4.5. Posterior distributions for the values of the k parameter in M10 - gender x year x area where parameter values vary across time. Each panel represents a different region around Iceland corresponding to the divisions shown in the map in Fig. 8.23. This figure does not explain what is causing spatiotemporal variation and differences in growth between gender; merely reflecting that it exists. The periods with no change in parameter estimates across time and wide credible intervals are areas and years where little to no data were available and posterior distributions took on the values of the priors (e.g. Subdivision 113, 1960 – 1980).

4.2.1 Systematically testing the importance of age and length data

The results from paper IV showed that it is important to have at least some age and length data present in integrated assessment models (Fig. 4.6); however, there is a point when adding more composition data samples is no longer beneficial at reducing bias. Bias was reduced within a single age data scenario when the sample size of length data was increased, but the presence of at least 25-50 samples for both age and length data seemed to reduce bias to an extent where further reduction in bias was not realized with

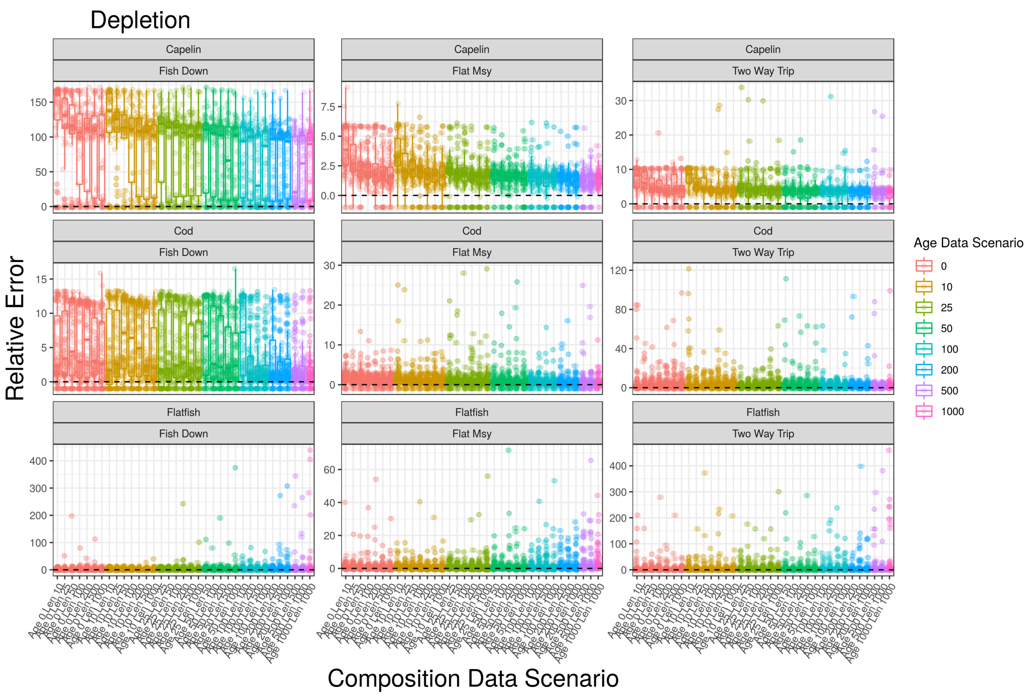


Figure 4.6. Relative error for estimates of depletion across composition data sample size scenarios and for various life-history and fishery development scenarios from paper IV. Composition data scenarios are listed along the x-axis, and each level of age sample size data are plotted as different colors.

.

increased sample sizes.

4.2.2 Importance of age data on uncertainty estimates in integrated assessment

While paper IV indicated that adding composition data decreased bias in models up to a certain point, the aim of this paper was to see how confidence intervals changed with higher sample sizes of age data. Generally speaking confidence intervals narrowed with higher sample sizes of age data (Fig. 4.7); however, this was not the case for some life history types. In addition, the example of Atlantic cod in this study showed that confidence intervals remained static for estimated number of fish across age data sample size, and that confidence intervals actually widened for biomass estimates in these same models (Fig. 4.8).

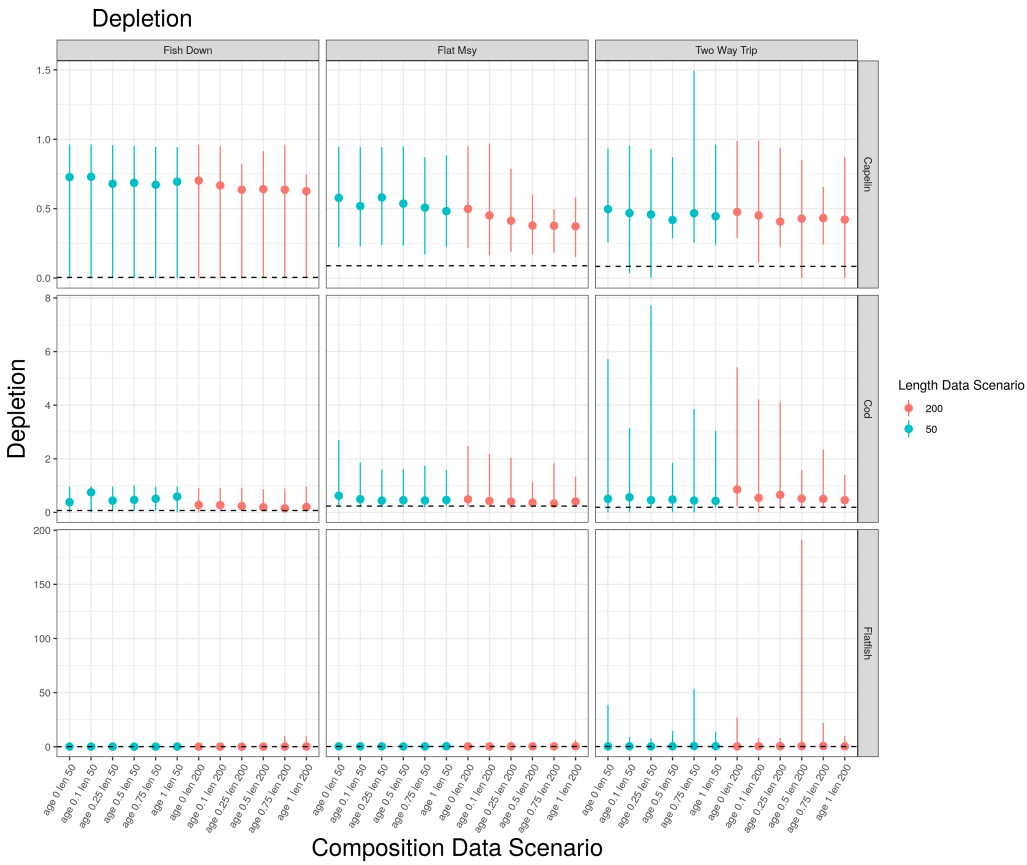


Figure 4.7. Point estimates and 95% confidence intervals (CI) of stock depletion in bootstrapped models from paper V. Point estimates decrease with increased composition data only up to a certain quantity of age data, but CI continues to decrease with increasing age data for cod.

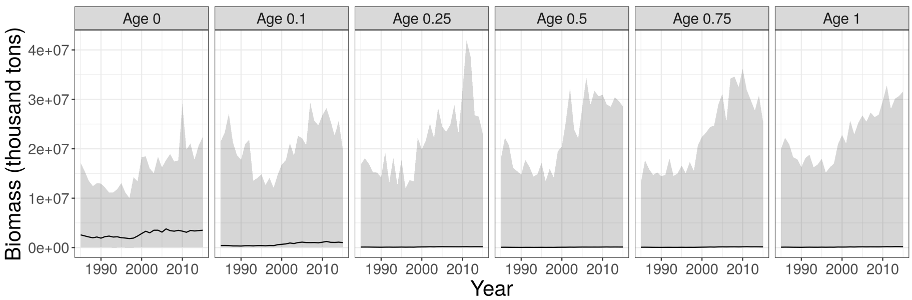


Figure 4.8. Biomass point estimates and 95% confidence interval of bootstrapped cod assessment models across time and for each age data sample size. The title at the top of each panel represents the proportion by which the full age data sample size was reduced.

5 Conclusions

The studies that comprise this thesis show that there are many and variegated issues with using composition data in fisheries models. As is shown in paper I, fish length data are biased to begin with based on common sampling techniques alone. Additionally, as is shown in that paper, the natural history of a fish species growth can potentially feedback to induce or enhance this effect. Selectivity bias can cause a real and potentially drastic bias on the fitting of growth curves to length-at-age data, which in turn can impact calculations that use growth such as basic stock assessments or maximum sustainable yield (MSY).

In addition to length data being inherently biased by sampling, the process of growth itself can vary throughout time for various reasons. Paper III shows that the growth of Atlantic cod growing in Iceland varies throughout both time and space. This is mainly due to abundance and recruitment of cod, which impacted the L_∞ parameter of the VBGF. In years where greater cod biomass was landed by commercial fishing a greater L_∞ parameter was estimated. Similarly, cohorts that had higher estimates of recruitment had greater estimated values of L_∞ in the VBGF. The k parameter of the VBGF was largely affected by abundance and recruitment of capelin (a main source of prey for cod). Years that had a high estimated SSB for capelin estimated a greater k value for cod, which indicates faster growth at younger ages. Similarly, cohorts in years where capelin SSB and recruitment were high had greater estimates of k. This makes perfect biological sense as cod likely grow faster in years where prey forage is abundant and highly available. Lastly, the t_0 parameter of the VBGF was generally negatively correlated with sea surface temperature (SST) of waters surrounding Iceland, which indicates that at warmer temperatures the length of age 0 fish will be slightly greater than in years where water temperature is colder. It is likely that cod is not unique in this regard and therefore reasonable to assume that other fish species exhibit spatiotemporal variation in growth as well. This is of concern as many stock assessment models used to estimate the status of stocks and determine harvest quotas are created without spatial differences among fish stocks accounted for as well as using a single growth model across time. It is advisable that time-varying growth should be considered when performing stock assessment on this species.

Despite the issues surrounding composition data in fisheries models they are incredibly important for accurate estimation of fish stock status and confidence about those estimates. The results from paper IV indicate that having at least some age and length data are essential for accurate estimation of biomass of fish in a stock. This is true up to a point, beyond which, more data does not necessarily lead to more accurate point estimates of model parameters. Additionally, the level of composition data at which more data does not improve estimates seems to be different across life history types. Capelin had greater bias in biomass, depletion, and parameter estimates at low levels of composition data whereas flatfish had relatively little bias in those metrics at those same levels. The cod life-history type fell somewhere in between. It seems that composition data are more important for accurately fitting parameter and biomass estimates for fish that are faster growing and have a higher overall natural mortality.

Even though more data does not seem to aid in accurately fitting point estimates past a certain point, it does seem to have an influence on uncertainty estimates in some scenarios. 95% confidence intervals (CIs) for both biomass and parameter estimates became wider as age data was successively stripped in certain age data sample size scenarios. It seems that including more age data may still be effective at tightening up uncertainty estimates even though there may be negligent returns on the point estimation of parameters.

Even though issues still remain techniques have been and are being developed to remedy at least some of the effects. The results in paper II reveal methods that perform well at correcting growth parameter estimates using length data biased by selectivity. The method presented by Troynikov (1999), which removes the effect of selectivity within the likelihood itself, performed well at reducing bias in growth parameter estimates, and a method similar to this is used to aid in correcting selectivity in the integrated assessment model SS3 (A. Punt *pers. comm*). In addition, integrated assessment models are being developed to incorporate temporal variation in growth. Advances and continued research in this regard are important as future management of stocks will likely depend ever-increasingly on composition data used in integrated assessment models.

Part II

Papers

6 Paper I - Aspects of both growth and selectivity affect growth parameter estimation bias

Paul N. Frater[1] and Gunnar Stefánsson[1]

[1]University of Iceland, Dunhaga 5, 107 Reykjavik, Iceland

Abstract

Selectivity curves for sampling size-at-age data are somewhat infamous for biasing growth curve parameter estimates. However, aspects of growth itself can also feed into this phenomenon and further increase bias into parameter estimates. We performed a simulation study by producing size-at-age data under a variety of growth curve parameter values and length-based selectivity curve parameter values for two selectivity functions by altering parameters both independently and in conjunction. In tests where parameters were altered one at a time the greatest amount of bias in growth curve parameter estimates came from sampling the data using dome-shaped selectivity. In addition, certain aspects of growth (namely variance in size-at-age) enhanced the biasing effect of both dome-shaped and asymptotic selectivity. When all parameters were altered together the greatest bias occurred in instances where length distributions overlapped to a greater degree and when asymptotic selectivity curves were more mildly sloped. We conclude that bias in growth curve parameter estimates is a result of the statistical weighting of both the probability of being a certain size at a given age and the probability of being

captured at that size.

Keywords: growth; selectivity; selectivity bias; von Bertalanffy

6.1 Introduction

Growth is a fundamental process in fish biology, and having an understanding of growth is essential to fisheries management (Lorenzen, 2016). Growth models are used to convert numbers-at-age and numbers-at-length into weight and vice versa within stock assessments (Begley, 2012; Maunder et al., 2016), to compute various life-history invariants such as *M* (Pauly, 1980; Roff, 1984; Jensen, 1996), and even to determine reference points for species management (Beverton and Holt, 1957; Deriso, 1987). Growth and body size has recently been shown to greatly influence reproductive output for fishes generally (Barneche et al., 2018) and food web trophic dynamics (Barneche and Allen, 2018). While it may seem a relatively straightforward task to gain a solid understanding of growth in fish, estimating true parameters to arguably the most common growth model, the von Bertalanffy growth function (VBGF), has proven to be rather difficult (Pardo et al., 2013; Wilson et al., 2015). This difficulty can arise from a number of factors such as aging error (Cope and Punt, 2007; Dortel et al., 2013) or the often-observed correlation of parameters in the model (Ratkowsky, 1986), but a large difficulty in obtaining true VBGF parameter estimates is due to size-based selectivity in collecting the data, which imposes a bias on size-at-age data and hence parameter estimates (Ricker, 1969; Goodyear, 1995; Gwinn et al., 2010; Sampson, 2014; Schueller et al., 2014; Wilson et al., 2015).

Bias in growth data due to size-based selectivity occurs because of the selective sampling of fish based on size rather than age[1] (Sampson, 2014). Since fish size is distributed

[1] Note that only sampling-induced bias is discussed here and not the impact of fishing mortality at length or fishery-induced evolution

around a mean for each age, and since selectivity curves are typically not constant, this combination imposes a greater sampling of certain sized fish that depends on both the shape of the selectivity curve, and the mean and variance of size-at-age. For example, a selectivity curve that rises steeply through a length distribution of fish at a given age will sample relatively more fish of a greater length within that particular age. In this way, the frequency distribution that describes the sampled size of fish for a given age can be thought of as the product of two functions (selectivity and the true size-at-age frequency distribution) where selectivity weights the expected value and variance of the true distribution hence biasing the data. Since the resulting data come from a distribution that is the product of two processes (growth and selectivity), either of these processes might be able to influence data collected from sampling gear. Given that either of these processes could potentially bias size-at-age data it seems important to understand how each of them impacts size-at-age data and, in turn, parameter estimates of the chosen growth model.

This study is an assessment of the individual impacts of parameters from: 1) the VBGF used to produce size-at-age, 2) an asymptotic selectivity function, and 3) a dome-shaped selectivity function, on the estimation bias of VBGF parameters. For each parameter in each of the above functions (*i.e.* growth, asymptotic selectivity, dome-shaped selectivity) we tested the bias that is induced by that parameter across a range of values by simulating length-at-age data, sampling it with the appropriate selectivity curve(s), and estimating parameters of the VBGF. Parameters were altered independently with a separate simulation run for each parameter being tested. In addition, we assessed the interaction of growth and selectivity parameters on estimation bias of VBGF parameters fit to length-at-age data sampled with a selectivity function by randomly assigning both growth and selectivity parameters and comparing bias to overlap in length distributions among age classes and the steepness of the selectivity function (asymptotic selectivity) or relative area under the curve compared to all available selection space (dome-shaped selectivity).

6.2 Methods

To test the impact of the growth model and selectivity on estimation of the VBGF parameters, we created simulations where the following procedure was performed at each iteration: create a basic population, assign paired length-at-age data using the VBGF, sample the population using both selectivity types as well as full samples (*i.e.* no selectivity), and estimate the VBGF parameters (*i.e.* L_∞, k, t_0) for each sampling type. The “No Selectivity" scenario simply used all length-at-age data produced by the population and growth simulation (*i.e.* a constant selectivity of 1). We performed this same basic procedure across a range of values for each parameter in the growth function as well as both selectivity functions. Populations in each simulation were created as age- and length-structured with number-at-age calculated as: $N_a = 1000e^{(-0.2a)}$. We then assigned length (*i.e.* sets of paired observations of length-at-age) using the VBGF for mean length-at-age and a normal distribution with variance as a hockey-stick function of age, which reasonably approximates the length distribution at age for many fishes (*pers. obs*, using data from Icelandic fishes, Hafrannsóknastofnun, 2018).

The basic calculations for a base case of paired length-at-age data was as follows:

$$
\begin{aligned}
L_{a_i} &\sim \mathrm{N}(L(L_\infty, k, t_0, a), \sigma_a^2) \\
L_\infty &= 100 \text{ (cm)} \\
k &= 0.25 \text{ (yr}^{-1}\text{)} \\
t_0 &= -0.5 \text{ (yr)} \\
\sigma_a &= \begin{cases} a \text{ (cm)} & a \leq 5 \\ 5 \text{ (cm)} & a > 5 \end{cases}
\end{aligned}
$$

where $L(\cdot)$ is the VBGF, and a is age from 1 to 10. However, for simulations that tested the effects of growth parameters these values were used as central values and a range above and below them were tested (Table 6.1).

To test how selectivity biased growth parameter estimates for both asymptotic and dome-shaped selectivity we sampled the population simulated as described above using both types of selectivity curves. For asymptotic selectivity we used the logistic function:

$$
S(l) = \frac{1}{1 + e^{-\alpha(l - l_{50})}}
$$

where S is the level of selectivity, ranging from 0 to 1, for a particular length l, α is the steepness parameter for the ascending limb, and l_{50} is the inflection point of the ascending limb.

To test dome-shaped selectivity the gamma selection function was used.

$$
S(l) = \left(\frac{l}{(\eta - 1)\beta\gamma} \right)^{\eta - 1} \exp\left(\eta - 1 - \frac{l}{\beta\gamma} \right)
$$

where S and l are as above, and η, β and γ are parameters controlling the steepness and doming of the selectivity curve.

We tested parameters both independently and interactively by performing two types of simulations: 1) altering a single parameter at a time while holding the others constant, and 2) randomly assigning growth and selectivity parameters over a number of iterations and then assessing bias. These simulations are described in greater detail below.

6.2.1 Independent Tests of Parameter Bias

For each of the parameters in the growth function and each of the selectivity functions a vector of evenly spaced values was established (Table 6.1). At each respective parameter value 100 replicate data sets of length-at-age were produced with N_a number of fish using the population structure and growth model with the growth parameters specified in Table 6.1 (except when a given growth parameter was being altered in which case a different value of that parameter was used at each iteration). For each parameter tested, 100 replicate data sets were produced for a particular value within the range according to the values found in Table 6.1 for that respective parameter while holding all other parameter values constant at the base case values. To alter variance in length distributions a scalar multiplier, m_σ, was used to scale values of σ_a, which either increased or decreased variance in the simulated population. Within each iteration of the simulations that tested for the effects of k, t_0, and m_σ sampling was performed using both selectivity functions with parameter values of $\alpha = 0.15$ and $l_{50} = 30$ for the logistic selectivity function and $\eta = 5$, $\beta = 0.7$, $\gamma = 20$ for the gamma selection function. The simulations that tested bias across values of L_∞ only sampled data using the asymptotic selectivity function. It would not make sense to increase maximum length by increasing L_∞ and not update the selectivity function appropriately (*i.e.* we wanted to test the effects of altering the L_∞ parameter relative to an appropriate selectivity – not against a single selectivity curve). This could be achieved easily with the asymptotic selectivity function but not by the dome-shaped selectivity function. Therefore, we only used asymptotic selectivity when altering values of L_∞ and used an l_{50} value of half the L_∞

value used to produce length-at-age data for a particular iteration.

Table 6.1. Parameters and range of values tested as well as the base case values that were held constant while testing other parameters. The first four are VBGF parameters where m_σ refers to the scalar multiplier that σ_a was multiplied by. Asymptotic and dome-shaped in the Function column refers to the type of selectivity whereas growth refers to the VBGF

Parameter	Function	Lower	Upper	Interval	Base Case Value
L_∞	Growth	50	150	1	100
k	Growth	0.01	0.5	0.01	0.25
t_0	Growth	-2	0	0.02	-0.5
m_σ	Growth	0.05	5	0.05	1
α_l	Asymptotic	0.01	1	0.01	0.15
l_{50}	Asymptotic	0	80	1	30
η	Dome-shaped	2	8	0.06	5
β	Dome-shaped	0.1	1.1	0.01	0.7
γ	Dome-shaped	10	30	0.2	20

At each parameter value ranging from the lower to the upper limit 100 datasets of paired length-at-age data were produced and VBGF parameters estimated by minimizing the negative log-likelihood (NLL) of the VBGF function. NLL was fit using paired length-at-age observations from both un-sampled data as well as data sampled with asymptotic and dome-shaped selectivity (except for the scenarios that ranged in L_∞, which selected data only with asymptotic selectivity). Only the parameters of L_∞, k, and t_0 were estimated in these simulations. σ_a was assumed known and was set to the observed variance of length values for each age in the likelihood function.

To assess bias in both L_∞ and k we calculated relative error RE of the estimated parameter:

$$RE = \frac{(\hat{\theta} - \theta)}{\theta}$$

where θ is the true parameter value and $\hat{\theta}$ is the estimated parameter value.

In order to avoid confusion with interpretation of the t_0 parameter (since it largely has negative values) we used absolute error to assess bias:

$$\text{Error} = (\hat{\theta} - \theta)$$

t_0 is generally a negative parameter and is almost always close to 0, so using absolute error should not cause issues with interpretation but will aid in the discussion of such.

6.2.2 Interactive Tests of Parameter Bias

To test how parameters interactively induced bias into length-at-age data and in turn affected growth parameter estimates we performed simulations using the same population, growth, and sampling structure as above, but randomizing the growth and selectivity parameters at each of 10,000 iterations and estimating VBGF parameters on the data sampled with selectivity. Randomized parameter values were sampled independently from the same range of values used in the independent tests (Table 6.1) with the exception of m_σ, which was reduced to an upper maximum of 3 to ensure that the variance in length-at-age was not unrealistically large at low values of L_∞. In addition, we fixed t_0 to 0 to ensure that lengths were lower at younger ages (*e.g.* a greatly negative t_0 along with a high k produces fish that are essentially at L_∞ at age 0). We again computed error values for the VBGF parameters, but instead of comparing directly to the parameter values themselves we used the following metrics to assess both the length-at-age distributions and the overall shape of selectivity curves. To get a general metric that indicated degree of overlap in length-at-age distributions within a simulated age-length structured population at each iteration we used the overlapping coefficient for probability distributions by comparing the probability distribution of lengths from one age to the next (*i.e.* 1 to 2, 2 to 3, etc.) and summed overlapping coefficients across ages as follows:

$$O_i = \sum_{a=1}^{A-1} \int_{-\infty}^{\infty} \min[f_L(l,a), f_L(l,a+1)]dl$$

$$f_L(l,a) = \frac{1}{\sqrt{2\pi}\sigma_a} \exp\left(-\frac{(l - L(L_\infty, k, t_0, a))^2}{2\sigma_a^2}\right)$$

where O_i is the total overlapping coefficient for iteration i, lengths l are normally distributed around a mean determined by the VBGF ($L(\cdot)$) and variance σ^2 for ages a in 1 through A. Therefore, all else being equal, a population with greater variance in the length-at-age distributions will have a higher total overlapping coefficient. The same would likely be true for slow growing fishes with a lower k in the VBGF.

To obtain a general metric for how the asymptotic selectivity function passed through the length-at-age distributions in each iteration we simply took the derivative of the logistic function at the mean length within each age (determined by the VBGF in that iteration) and recorded the steepest slope that occurred at the mean length within any age.

$$d_{max_i} = \max_a \frac{d}{dl}\left[\frac{1}{1 + \exp(-\alpha_i(l_a - l_{50_i}))}\right]$$

where d_{max} is the maximum derivative in each iteration i of the logistic selectivity function found among mean length l at age a for all ages A produced by the VBGF in iteration i.

To understand how the overall level of doming affected VBGF parameter estimation bias we created a metric that assessed how steeply domed the gamma function was for an iteration by dividing the area under the dome-shaped selectivity curve by the total possible area that the selectivity curve could occupy and taking the inverse of this number as follows:

$$\text{Doming}_i = \frac{200}{\int_0^{200} \left(\frac{l}{(\eta_i-1)\beta_i\gamma_i}\right)^{\eta_i-1} \exp\left(\eta_i - 1 - \frac{l}{\beta_i\gamma_i}\right) dl}$$

where l is length and η, β, and γ are the parameters for the gamma selectivity function at iteration i. We set the upper limit of the definite integral to 200 as the probability of a fish obtaining this length essentially becomes 0 even at the highest L_∞ and m_σ. While there are potential issues with this metric of doming such as curves with very different shapes having a similar area, all of our selectivity functions were required to have a maximum value of 1 at some point along the curve. Therefore, a smaller doming value will generally be indicative of a curve that is not very steeply domed, whereas greater doming values will generally be indicative of a steeply domed selectivity curve in this study.

We compared the VBGF parameter estimation bias by plotting the error values for each estimated parameter against both O and either d_{max} (for asymptotic selectivty) or Doming (for dome-shaped selectivity). All simulations and analyses were performed in R3.4.3 (R Core Team, 2017).

6.3 Results

6.3.1 Effects of Parameters Used to Simulate Growth

In general, growth parameters used to produce length-at-age data did not greatly alter VBGF parameter estimates (Figs. 6.9–6.11). Some exceptions to this were along the range of L_∞ used to produce length-at-age, which underestimated L_∞ and overestimated k. Estimates of t_0 were positively biased at lower values of L_∞, but this decreased with increasing L_∞ and became negatively biased at high values of L_∞.

In addition, L_∞ and k showed consistent bias across values of t_0 when the VBGF was fit to data sampled with dome-shaped selectivity (Figs. 6.9–6.10, panel for t_0). L_∞ was consistently underestimated across values of t_0 in the scenarios sampled with dome-shaped selectivity, and k was overestimated in these scenarios and across values of t_0 but becoming slightly less so as t_0 approached 0.

Higher values of m_σ resulted in increased bias of all parameters (m_σ panel in Figs. 6.9–6.11). This bias was much greater under dome-shaped selectivity, but was also apparent for asymptotic selectivity, especially in estimates of t_0. Additionally, the error in parameter estimates increased multiplicatively with m_σ when there was no selectivity, but not in a single direction *per se*. The effect of m_σ on estimation bias of all parameters was most apparent with dome-shaped selectivity, where increasing variance (mediated through increased m_σ) seemed to have an exponential effect on parameter estimation bias. At high values of m_σ the dome-shaped selectivity curve consistently underestimated L_∞, and overestimated k. t_0 also tended to be more positively biased with greater m_σ. The asymptotic curve resulted in relatively less bias at high values of m_σ compared to dome-shaped selectivity especially for L_∞ and k, but did cause some bias in the estimation of t_0 especially at the mid-ranges of m_σ.

6.3.2 Effects of Asymptotic Selectivity

Data sampled under asymptotic selectivity did not produce much bias in the VBGF parameter estimates relative to dome-shaped selectivity. The most noticeable exceptions to this were in the k and t_0 parameters where increasing α in the logistic function biased estimation of both of these parameters (Figs. 6.10–6.11). Parameters estimated with data sampled from asymptotic selectivity under different values of l_{50} remained relatively unbiased until values of l_{50} became greater than about 40, which resulted in an exponential increase in bias for all parameters (Fig. 6.9-6.11). This particular value of 40 for l_{50} is of relatively little importance as the bias induced by this parameter would vary with the life-history and growth of the organism being sampled.

6.3.3 Effects of Dome-shaped Selectivity

The effect of dome-shaped selectivity across parameter values in the gamma selectivity function biased some VBGF parameter estimates, but not all (Figs. 6.9–6.11; bottom tier of panels). Lower values of η in the gamma selectivity function resulted in underestimated L_∞ (Fig. 6.9), whereas the bias for estimated t_0 was fairly close to 0 at lower values of η, but increased with increasing η (Fig. 6.11). Certain values of β resulted in high bias of both L_∞ and k estimates with the greatest amount of bias being between β values of roughly 0.15-0.5. These values of β resulted in a hump- or U-shaped trend depending on the VBGF parameter assessed. Estimates of t_0 also had some bias when fit to data sampled at these same levels of β, but bias was less consistent. γ produced bias in VBGF parameter estimates similarly to the upper values for β in the gamma selectivity function, which makes sense as these parameters are simply factors of one another.

6.3.4 Interactive Effects of Parameters

In the simulations that randomly assigned parameter values for both growth and selectivity processes there was a general trend in bias for all estimated parameters where error increased steeply at the highest levels of overlap in the length-at-age distributions (Fig. 6.12). Under asymptotic selectivity the estimates with the greatest error also tended to occur at lower levels of d_{max}, which would be indicative of a more gradually sloping selectivity curve. For dome-shaped selectivity there did not seem to be a single, consistent trend among doming and bias for any parameter.

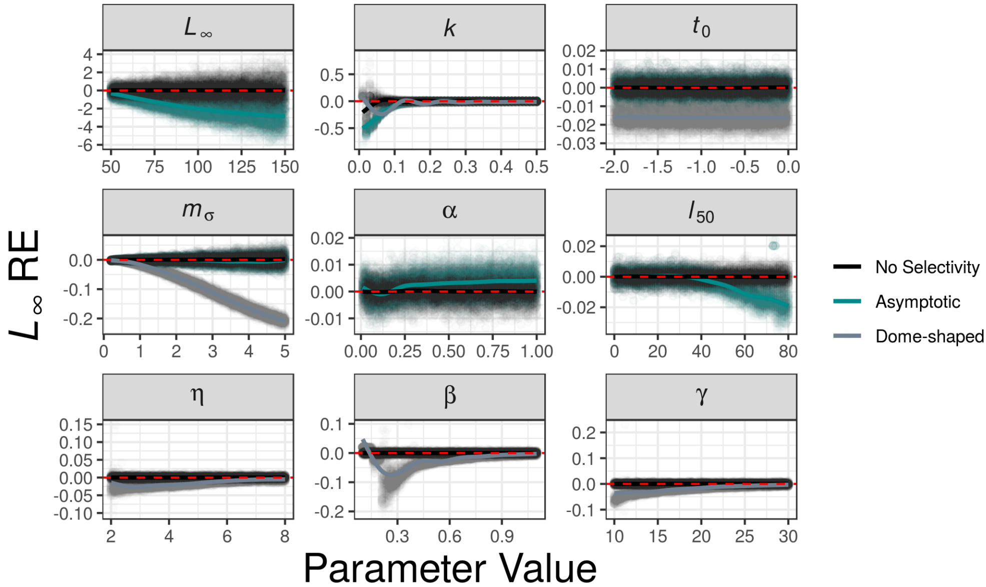


*Figure 6.9. Relative error of L_∞ for different values of growth and selectivity parameters and using data under either asymptotic or dome-shaped selectivity or using all data (*i.e. *no selectivity). L_∞, k, and t_0 are parameters used in the VBGF, m_σ is the scalar multiplier applied to variance in length-at-age, α and l_{50} are parameters in the logistic selectivity function, and η, β, and γ are parameters used in the gamma selectivity function. The dashed horizontal line at 0 indicates no bias. The solid lines showing trend in each sampling type is a GAM-smoothed line produced by* `geom_smooth()` *in the ggplot2 package in R. m_σ refers to the scalar value used to multiply the vector of variance values.*

6.4 Discussion

This study illustrates that size-based selectivity indeed biases parameter estimates of growth models (Ricker, 1969; Sampson, 2014), but the factors leading to that bias are more intricate than perhaps previously assumed (Ricker, 1969; Thorson and Simpfendorfer, 2009; Sampson, 2014). These findings show that biases in growth parameter estimates due to size-based selectivity depends not only on the shape of the selectivity function (Thorson and Simpfendorfer, 2009; Schueller et al., 2014), but also partially on growth processes as well. For instance, greater values of m_σ resulted in greater bias

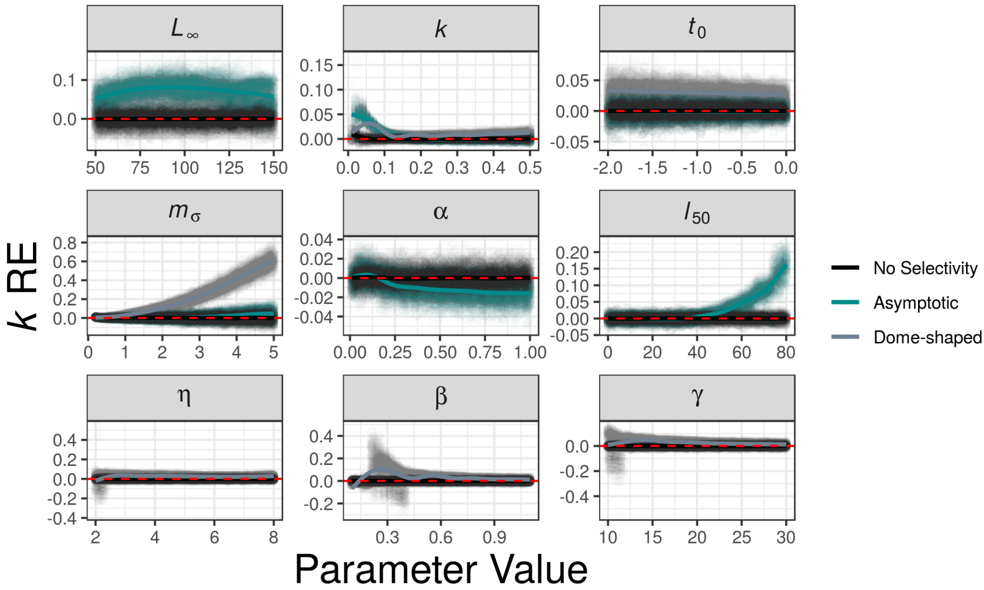


Figure 6.10. Relative error of k for different values of growth and selectivity parameters. Panels, x-axes, and scenarios are the same as Fig. 6.9.

in all VBGF parameters that were estimated using data sampled under dome-shaped selectivity (*i.e.* a greater variance in length-at-age led to greater bias for each parameter – either positive or negative depending on the parameter). This could be an important factor when estimating growth in fish that have high variance in size-at-age such as longer-lived fishes that tend to grow slower and have greater overlap in size among ages (Bowker, 1995). Similarly, Hulson and Hanselman (2014) found that uncertainty and bias in the parameter estimates of stock assessments differed among life history types even when the same selectivity functions were used. Differences in the variation of size-at-age distributions among life-histories may perhaps be a reason for that.

Different values of L_∞ also biased estimates of VBGF parameters when using data sampled with an asymptotic selectivity curve. This could be an issue related to the cumulative effect of mortality across age resulting in fewer numbers of sampled fish at older ages, which could cause difficulties in estimation of L_∞. However, bias existed

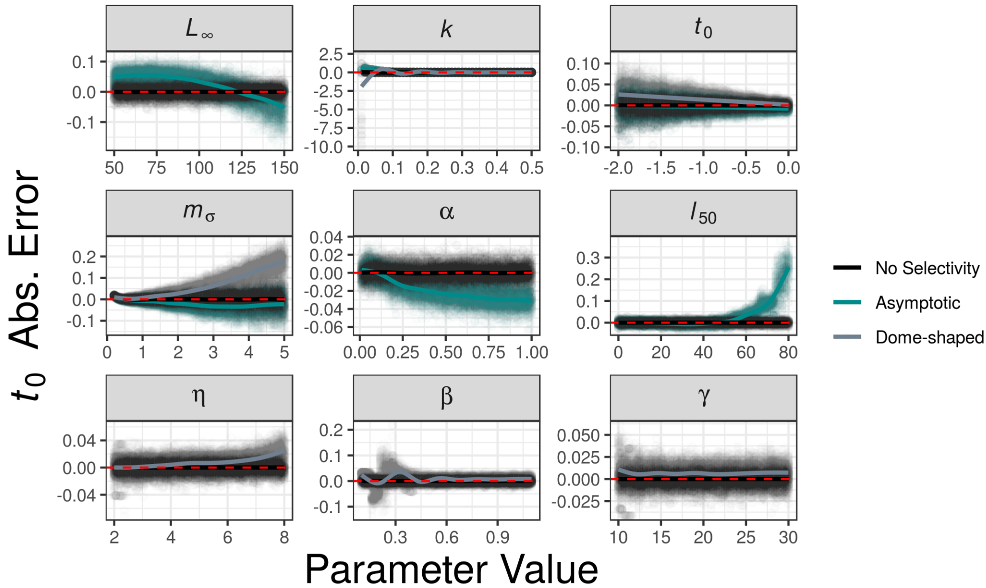


Figure 6.11. Absolute error of t_0 for different values of growth and selectivity parameters. Panels, x-axes, and scenarios are the same as Fig. 6.9

for both k and t_0 as well across different values of L_∞ used to simulate length-at-age. Parameters of the VBGF are notorious for being highly correlated, and the functions used in this study are no exception to that. It is possible that higher values of L_∞ caused fewer young fish to be sampled as we also increased α in the asymptotic selectivity function in proportion to L_∞. This could make t_0 ill-defined and in turn may cause problems for the estimation of L_∞ and k. The l_{50} parameter in the logistic selectivity function was altered on a sliding scale proportional to L_∞ in these simulations, so the phenomenon of an ill-defined t_0 is certainly possible.

The only other growth parameter that seemed to bias parameter estimates of the VBGF was t_0. This, however, is more likely an artifact of dome-shaped selectivity itself rather than the impact of different values of t_0. The underestimation of L_∞ and overestimation of k across different values of t_0 both make sense when thinking about the statistical properties of the dome-shaped selectivity function (Hulson and Hanselman, 2014). On

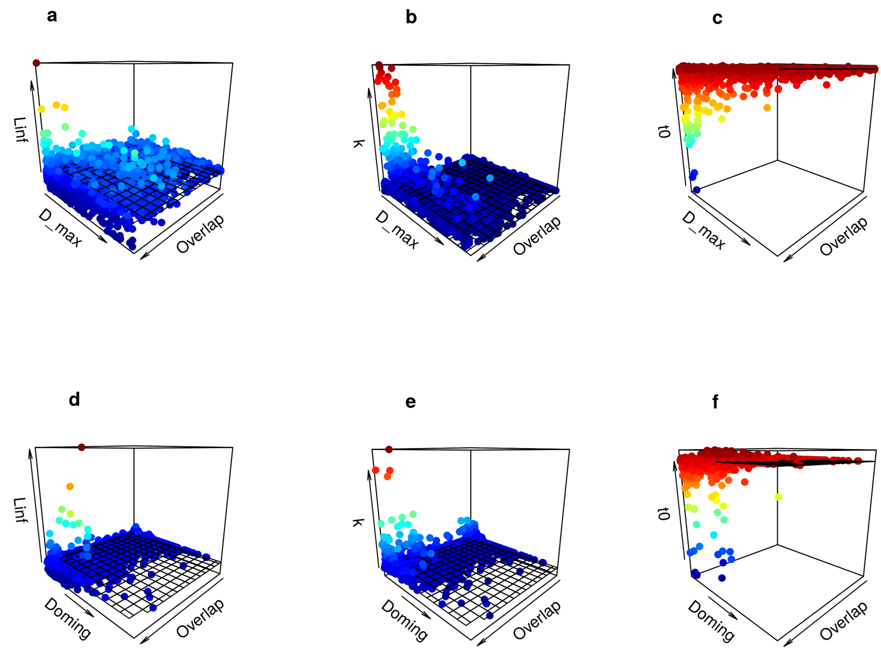


Figure 6.12. Error in growth parameters L_∞ (a,d), k (b,e), and t_0 (c,f) from simulations with randomly chosen growth and selectivity variables for asymptotic selectivity function (a-c) and dome-shaped selectivity function (d-f). "Doming" refers to the metric used to calculate the level of doming in the gamma selectivity function. "Overlap" refers to the combined amount of overlap among length distributions from one age to the next. Data in the figures are subset to exclude extreme outliers to better visualize general trends.

the ascending portion of the selectivity curve a relatively greater number of fish at higher than mean length-at-age will be selected (Sampson, 2014), which will increase the expected mean length for that age in the sampled data and bias estimates of k upward. This translates into greater estimates of the growth coefficient k. On the descending limb a relatively greater number of fish lower than mean length-at-age will be selected, which translates into a lower estimated L_∞ parameter. These respective influences will become greater as the ascending and descending limbs become steeper if they pass through the bulk of the distribution of size-at-age. While some may attribute this bias to correlation in the growth function parameters it should be noted that this consistent bias of L_∞ and k only occurred under dome-shaped selectivity, while the optimization using an asymptotic sampling scheme maintained relatively unbiased across values of t_0.

Among the bias introduced by different values of parameters for the selectivity curves, asymptotic selectivity seemed to produce the least amount of bias overall. The biased estimates of both k and t_0 with increasing values for α makes sense as a steeper selectivity curve will result in a lower probability of selection at the smallest ages and lengths and fewer of those fish being sampled. This in turn will cause t_0 to be estimated at lower (*i.e.* more negative) values which will increase the height of the curve at age 0. A lower value of t_0 translates to a greater value for size at age 0 and forces the k parameter to correspondingly decrease as the curve no longer needs to increase as rapidly to obtain the appropriate L_∞ value. α alone caused bias in VBGF parameter estimates, but if we altered values for both α and l_{50} we might see strong bias for a particular size-at-age distribution where steep values for α pass through the middle of that distribution. The l_{50} parameter, on the other hand, introduced very little bias below values of about 40. At values of $l_{50} \gtrapprox 40$ all parameters began showing an exponential trend in bias. This is not entirely surprising as the L_∞ value used to produce size-at-age data was 100. As l_{50} approaches this value the number of smaller fish sampled decreases, which correspondingly introduces increasingly greater bias.

Dome-shaped selectivity seemed to, by and large, introduce the greatest amount of bias into the estimation of VBGF parameters in the independent tests. This was true for all estimated parameters of the VBGF (*i.e.* L_∞, k, and t_0) and across parts of the range for each parameter in the gamma selectivity function (*i.e.* η, β, and γ). This is consistent with other findings that bias induced by dome-shaped selectivity is often more complicated than that of asymptotic selectivity (Hulson and Hanselman, 2014; Francis, 2016) and introduces greater bias into parameter estimates of not just growth but stock assessment parameter estimates in general (Catalano and Allen, 2010). Dome-shaped selectivity did not seem to produce much bias in the interactive tests. This is surprising given the results of the independent tests where dome-shaped selectivity produced the greatest amount of bias on growth curve estimation. This is perhaps a result of our doming metric, which we discuss below.

Parameters of the dome-shaped selectivity function can be generalized into two classes: horizontal shifting (η) and dome-control (β and γ). The η parameter of the gamma selectivity function tends to control how far left or right the whole shape of the curve moves. At first glance it may seem that the β and γ parameter produced drastically different estimation biases, but when viewing the abscissa and ordinates of Figs. 6.9–6.11 notice that these two parameters actually produce quite similar results. η and γ are factors of each other in the gamma selectivity function and with lower values of γ we would likely see the curvilinear response of greater bias and then error values back towards the true VBGF parameters (but with greater variance in estimates) as seen in β. Therefore, the effects of η, β, and γ all biased VBGF parameter estimates quite similarly. Lower values of the parameters for the gamma selectivity function can be thought of as causing more of a dome-shape in the selectivity function (*i.e.* steeper ascending and descending limbs). Therefore, it is no surprise that estimation of VBGF parameters using data sampled at these levels induced the greatest estimation bias. A highly domed curve will result in little data sampled at both the lower and upper sizes with the chance for greatly altered size distributions in the sampled data due to the steepness in the ascending and descending portions of the curve (Sinclair et al., 2002a; Sampson, 2014).

While many of the parameters of both growth and selectivity processes seemed to cause at least some bias when tested independently, the collective metric of overlap seemed to produce the greatest amount of bias when testing estimation bias interactively. Error in all parameters increased consistently with greater overlap among length-at-age distributions. A greater overlap in the distributions of length-at-age could be caused by a number of parameter combinations that produce the age-length structured populations used in these simulations, but would most likely be affected greatest by variance in length-at-age and the k parameter of the VBGF. An increased variance in the length-at-age distributions would directly result in a higher overlap value with all other parameters held constant, but a lower k could also have a similar effect as it would only increase the mean length-at-age by a relatively small amount each year. This impact of k on overlap would depend on variance in the length-at-age distributions as well.

Interestingly, some of the most biased parameter estimates were obtained with parameter combinations that had only gradually increasing slopes as indicated by a lower d_{max}. We would have predicted the opposite; that steeper slopes passing through a length-at-age distribution result in more biased data. The result of greater bias at lower d_{max} is likely due to the fact that an asymptotic selectivity curve that is shallower will pass through a greater number of length distributions representing different ages and will bias the data from multiple ages as opposed to a sharply sloping asymptotic curve which would most likely only pass through a length distribution for a single age and only biasing length data for that particular age.

Also somewhat surprising was that the level of doming did not cause much of an estimation bias in the interactive estimates of VBGF parameters. Based on the independent tests it might have been assumed that doming to a greater degree would bias parameter estimates to a greater degree. There was not an obvious trend of this in the interactive tests sampled with the dome-shaped selectivity, though. It is possible that our metric to detect level of doming is not accurate, but it seems intuitive that a more steeply domed curve will produce a smaller area relative to the possible overall space of selectivity – all else being held equal, which it was. It would not be pragmatic to only look at the derivative of the ascending limb of a dome-shaped selectivity curve as the steepness of the descending limb also produces bias. Perhaps a better metric for doming would reveal a greater biasing effect, but we did not endeavor to undertake such a venture in this particular study.

Our study does not include the effects of measurement error on the bias of length-at-age data. Aging error in particular has been shown to have a large impact on both growth curves and stock assessment models (Reeves, 2003; Cope and Punt, 2007). An interesting avenue to pursue this research further would be to study the effects of aging error in combination with various growth and selectivity parameters to see if the bias from one particular aspect outweighs the others. It is possible that the bias from one aspect of either growth, selectivity, or measurement error completely outweighs the bias from all other sources in which case improvements to that particular component should be made first. Alternatively, it is possible that all processes may equally or interactively bias length-at-age data in which case the path forward to a corrective solution becomes quite a bit more difficult.

Conclusions

We found a biasing effect on the parameter estimation of the VBGF both for some aspects used to simulate growth as well as some parameters used to compute selectivity. Parameters of asymptotic selectivity contributed to growth parameter estimation bias less so than dome-shaped selectivity, but this is dependent on how the asymptotic selectivity curve passes through size-at-age distributions. All parameters of dome-shaped selectivity consistently introduced bias into growth parameter estimates and seemed to do so to a greater extent for independent parameter tests when the selectivity curve was highly domed with steep ascending and descending limbs, but this trend was not detected when parameters were tested interactively. The overlap in length distributions among ages had the most consistent biasing effect when parameters were tested interactively. These results could potentially be used to structure species-specific surveys if growth and size distributions are known well. Our recommendations for reducing the bias in growth parameter estimation due to size-based selectivity would be to: 1) alleviate doming as much as possible, 2) attempt to maintain a constant (*i.e.* flat) selectivity curve (which is unlikely to happen), and 3) structure survey gear in a species-specific manner such that the selectivity gear passes steeply in between size distributions of two distinct ages thereby minimizing the weighting impact that selectivity will have on sampled size-at-age. This will not be feasible for all species as the variance in size-at-age for some species will cause these distributions to overlap to a large extent. While we would like to offer solutions to improve growth estimation either for stock assessment or management purposes we conclude by stating that once size data are biased by some combination of the growth and selectivity processes it is rather difficult to correct those data. The scope of this paper was namely to illustrate that selectivity-based bias in size-at-age data is not a problem of selectivity alone, but rather a collective process of the shape of a selectivity curve as well as how size-at-age is distributed across ages.

Acknowledgements

Funding for the lead author to undertake his Ph.D. studies was received from the European Union's Seventh Framework Programme for research, technological development and demonstration for the project MareFrame [Grant Agreement No. 613571] and from the European Commission's Horizon 2020 Research and Innovation Programme for the project Science, Technology, and Society Initiative to minimize Unwanted Catches in European Fisheries (Minouw) [Grant Agreement No. 634495] as well as from the Icelandic Research Fund (Rannís, [No. 152039051]). The lead author is grateful to the second author for acquiring said funding.

7 Paper II - Comparison and evaluation of approaches aimed at correcting or reducing selectivity bias in growth parameter estimates for fishes

Paul N. Frater [1] and Gunnar Stefansson [1]

[1]University of Iceland

Dunhaga 5

107 Reykjavik

Iceland

Abstract

Several techniques have been proposed and tested to alleviate the common problem of selectivity bias in fish size-at-age data used for fitting growth models. Many of these techniques have not been directly compared to each other, though. Here, we test six methods for correcting selectivity bias in fish growth models using two selectivity shapes and across three life history types. We also include an example of fitted growth curve parameters from each method using data on Atlantic cod (*Gadus morhua* L.) from standardized surveys in the waters surrounding Iceland. Lastly, we present simulations that incorporate three types of misspecification on each method, selectivity type, and life-history strategy to discern the impacts of incorrect assumptions stemming from each method.

Age structured populations are simulated, and the basic von Bertalanffy growth curve is used as the mean to assign normally-distributed length-at-age data which are sampled under two selectivity shapes. Growth parameters are then estimated using these data for each of six bias-correction methods. In general, likelihood-based bias-correction methods resulted in growth parameter estimates that are closer to true values used as well as more similar to parameter estimates derived from data sampled with no selectivity. The method introduced by Troynikov (1999) performed best for both simulation and misspecification scenarios. This result is true of both dome-shaped and asymptotic selectivity as well as across life-history types.

Keywords: selectivity bias; growth; bias-correction; likelihood; von Bertalanffy; age-length data

7.1 Introduction

Understanding the growth of fishes is important for fisheries biology, stock assessment (Maunder et al., 2016), and management of species (Lorenzen and Enberg, 2002; Lorenzen, 2016). Applications of growth curves range from as basic as having a simple biological understanding of the life-history of fish species to as sophisticated as determining appropriate size-at-age within integrated assessment models (Francis, 2016). Growth of a fish species relates to everything from natural mortality (Pauly, 1980; Gislason et al., 2010) to the shape of maturity ogives (Roff, 1984; Jensen, 1996) to reference points for the management of a species (Beverton and Holt, 1957; Deriso, 1987). Despite the importance of having a solid understanding of growth and how it relates to fisheries research and management it can nonetheless be assumed that our knowledge of growth for any given fish species is imperfect even in data-rich scenarios. There are several reasons for having an incomplete understanding of growth such as: time-varying growth (Lorenzen, 2016; Thorson and Minte-Vera, 2016), growth variability among individuals (de Zárate and Babcock, 2016; Webber and Thorson, 2016), and gear selectivity from sampling or commercial harvest of fish (Sampson, 2014; Schueller et al., 2014).

Of these impediments to gaining a complete understanding of growth in fishes, length-based selectivity is the process that directly places a sampling bias on subjects (Fig. 7.13, Sampson, 2014; Francis, 2016). Selectivity is the probability of sampling fish of a certain size or age (either by a fishery or a scientific survey), and is a product of three processes: contact selection, availability, and population selection, which is the net effect of the former two (Sampson, 2014). Contact selection is relatively straightforward – fish of a certain size are either caught or not in fishing or sampling gear due to either size limitations (e.g., smaller fish swim through nets or trawls of a certain mesh size, Somerton et al., 2011) or behavioral traits (e.g., trawl avoidance, Handegard and Tjøstheim, 2005; Ryer, 2008). For example, flatfish have a life-history predation strategy of burying themselves in sandy substrate to ambush prey, which might make them less available to entry into a trawl net (Ryer, 2008). Availability is simply how available fish of a certain size are to fishing gear. As an example, certain fish occupy different spatial areas throughout their life rendering them unavailable to certain gear types at durations of their existence. Population selection is a slightly more complex concept that can combine contact selection and availability and often emerges at larger spatial scales as the result of fish of different sizes or ages using different spatial areas or habitat types (Sampson and Scott, 2011). Sampson (2014) showed that even in a fishery where all fleets imposed asymptotic selectivity on fish the overall population selection was dome-shaped under certain fishing scenarios. This phenomenon was due to different size or age classes utilizing different spatial areas in the model.

The above processes may interact together to form an overall selection (for either

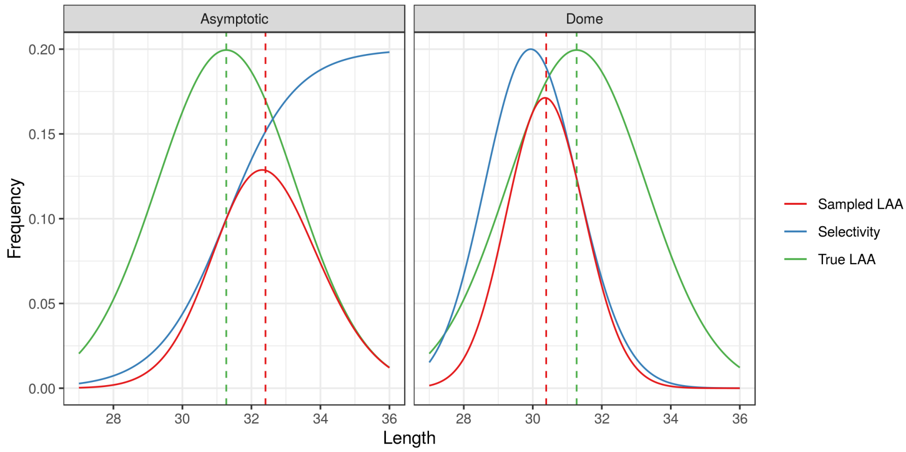


Figure 7.13. Diagram illustrating the biasing effect of selectivity on length-at-age distributions for asymptotic selectivity and dome-shaped selectivity. The true length-at-age distribution is shown in green. The sampled length-at-age distribution appears in red and is biased by the selectivity curve shown here in blue. Dashed vertical lines indicate the mean of the respective distributions.

commercial fishing or surveys) for a particular fishery, and this selection imposes a bias on growth data (Sampson, 2014). In size-based asymptotic selectivity curves, the smallest fish (which tend to be the youngest age-classes) are typically underrepresented, and those age classes with a size distribution corresponding to the steepest part of the selection curve will have a sampled mean length greater than the population mean length as proportionally more fish of a larger size will be sampled in those respective age-classes (Ricker, 1969; Gwinn et al., 2010; Wilson et al., 2015).

A similar but opposite phenomenon occurs with the largest (and typically oldest) fish with dome-shaped, length-based selectivity. Longer fish sampled under the descending limb of a dome-shaped size-based selectivity curve tend to have a mean length lower than the population mean (for that age class) as proportionally more fish of a smaller size in these length distributions will be sampled (Taylor et al., 2005). The very nature of size-based selectivity curves creates a statistical sampling bias which, as is shown below, can be corrected for in the likelihood function used for estimation of the growth parameters for a particular set of data if the size-based selectivity function is known.

Several methods exist to attempt to correct for selectivity-biased growth data (Troynikov, 1999; Troynikov and Koopman, 2009; Gwinn et al., 2010; Schueller et al., 2014; Wilson et al., 2015), but a direct comparison of these methods is lacking (see Gwinn et al., 2010, for the most comprehensive comparison to date). Correction or reduction of selectivity bias generally falls into one (or more) of three categories: 1) fixing parameters of the growth function, 2) removing poorly sampled data, or 3) altering the likelihood function of the growth model to account for selectivity or poorly sampled data.

Methods from these three categories of selectivity bias-correction have not been tested simultaneously. Gwinn et al. (2010) presented an evaluation of four methods for reducing selectivity bias on fitted growth curve parameters. All of the methods tested by Gwinn et al. (2010) were in the first and/or second type described above (i.e., removing poorly sampled data or fixing growth curve parameters, respectively). These four methods were: 1) fix the t_0 parameter of the VBGF to 0, 2) remove poorly sampled data, 3) combine options 1 and 2, and 4) fix the L_∞ parameter of the VBGF to the maximum observed length. Gwinn et al. (2010) found that none of these methods completely eliminated the effect of selectivity bias. Schueller et al. (2014) introduced a method based on McGarvey and Fowler (2002) that adjusts the probability of length-at-age by how well a length is thought to be sampled (i.e., the normal probability is divided by the definite integral of the normal where the upper and lower limits are chosen cutoff lengths). Troynikov (1999) used Bayes' theorem to derive a likelihood function that incorporates selectivity into growth curve parameter estimates by multiplying the probability of length-at-age by the selectivity and dividing this by its integral across all lengths. The method of Schueller et al. (2014) is actually a special case of the Troynikov (1999) method when selectivity is uniform over a specified range. This paper directly compares the above methods using three life-history types and across three sample size scenarios. We provide a short background on the statistical aspects of selectivity bias below, but do not delve into the details of all methods described above as they may be found in the respective references. Basic likelihood descriptions and references are found in Table 7.2.

Our objectives for this study were to: 1) compare the performance of published selectivity bias-correction methods across life-history types, 2) assess the performance of these methods under misspecification scenarios, and 3) provide an example of these bias-correction likelihood methods using data from an actual stock.

Table 7.2. Methods evaluated in this study. Type refers to the general technique: type 1 are methods that fix parameters in the likelihood function, type 2 are methods that remove poorly sampled data, and type 3 are methods that somehow alter the likelihood function based on either selectivity (L6 method) or a user-determined length truncation criteria (L5). Normal in the NLL column indicates the normal distribution as the likelihood (i.e. $\prod_a \prod_i \mathrm{N}(\mu_{ai}, \sigma_{ai})$), and VBGF indicates the von Bertalanffy growth function as described by eqn. 1. $f_a(l)$ and $\tilde{f}_a(l)$ in the NLL column refer to the true and biased density functions. Est. Params. displays which parameters were estimated for each likelihood type.

Method	Lik. ID	Type	Reference	NLL	Est. Params.
Normal (i.e. no correction)	L0	–	–	Normal	$L_\infty,k,t_0,\sigma_\infty,b,a'$
Fix t_0	L1	1	Gwinn et al. (2010)	Normal	$L_\infty,k,\sigma_\infty,b,a'$
Rm Small Samples	L2	2	Gwinn et al. (2010)	Normal	$L_\infty,k,t_0,\sigma_\infty,b,a'$
Rm Small Samples/Fix t_0	L3	1,2	Gwinn et al. (2010)	Normal	$L_\infty,k,\sigma_\infty,b,a'$
Fix L_∞	L4	1	Gwinn et al. (2010)	Normal	k,t_0,σ_∞,b,a'
Lik$_{Trunc}$	L5	3	Schueller et al. (2014)	$\prod_a \prod_{i=1}^{N} \frac{\mathrm{N}(\mu_{ai}, \sigma_{ai})}{\int_{min}^{max} \mathrm{N}(\mu_{ai}, \sigma_{ai})}$	$L_\infty,k,t_0,\sigma_\infty,b,a'$
Troynikov	L6	3	Troynikov (1999)	$\prod_a \prod_{i=1}^{N} \frac{\tilde{f}_a(l_{ai})S(l_{ai})}{\int_{l'} f_a(l')S(l')}$	$L_\infty,k,t_0,\sigma_\infty,b,a'$

7.2 Methods

7.2.1 Description of established selectivity bias-correction likelihood methods

The methods that fall into both type 1 and type 2 (as described above and in Table 7.2) utilize the regular likelihood that would be used if no selectivity bias-correction was being sought. In the present study this is the normal p.d.f. Both L1 and L4 are not statistically consistent with the other methods as a parameter is fixed in each of these methods.

The likelihood types which remove poorly sampled data (i.e., type 2 – methods L2 and L3) simply eliminate length-at-age data thought to be poorly sampled by the selectivity curve (e.g., shorter lengths for asymptotic selectivity and both the shorter and longest lengths for dome-shaped selectivity).

The last type of methods that aim to correct selectivity bias during the estimation of growth curve parameters are those that somehow alter the likelihood (i.e., type 3 – methods L5–L6). The likelihood in these methods is actually defined to account for the sampling process while fitting growth parameters, and this generally involves some sort of adjustment that incorporates selectivity into the uncorrected likelihood function. For example, the L5 method divides the normal likelihood by the definite integral of the normal likelihood with truncation points for the minimum and maximum that are assumed to be poorly sampled. The L6 method divides the product of the normal likelihood and selectivity at length by the integral of such.

7.2.2 Statistical description and correction of the effect of selectivity on size at age data

Mean size-at-age (or, synonymously, growth) can be described using a growth function. Arguably the most common is the von Bertalanffy growth function (VBGF) for length:

$$\begin{aligned} L_a &\sim \mathrm{N}(\bar{L}_a, \sigma_a) \\ \bar{L}_a &= L_\infty(1 - \exp(-k(a - t_0))) \end{aligned} \tag{1}$$

where L_a is length at age a, $\bar{L}_a$ is mean length-at-age determined by the following parameters: L_∞, which is the theoretical maximum length as age a approaches infinity; a steepness parameter, k; and t_0, which is a horizontal scaling parameter that determines the theoretical age when length equals 0. σ_a is the age specific variance associated with lengths L at each age. Lengths-at-age are often assumed normally distributed, which produces true frequency distributions that are normally distributed around the mean:

$$f_a(l_{ai}) = \frac{1}{\sqrt{2\pi}\sigma_a} \exp\left(-\frac{(l_{ai} - \mu_a)^2}{2\sigma_a^2}\right) \tag{2}$$

where l_{ai} is the length of fish i of age a, μ_a is the mean length-at-age produced by eqn. 1, and σ_a the variance in length at age a.

Eqn. 2 can be used with length-at-age data as a likelihood to produce growth parameter estimates. The problem with this is that length-at-age data are often biased by selectivity (Fig. 7.13), which will in turn produce parameter estimates for the VBGF that are also biased. Selectivity functions are often either asymptotic or dome-shaped. A common asymptotic selectivity function is:

$$S(l) = \frac{1}{1+\exp(-\alpha(l-l_{50}))} \tag{3}$$

where l is length, and α and l_{50} are the steepness and half-saturation parameters, respectively. There are several dome-shaped selectivity functions, but a relatively simple one with few parameters can be represented with the following, which is often referred to as the gamma selectivity function (Begley, 2012):

$$S(l) = \left(\frac{l}{(\eta-1)\beta\gamma}\right)^{(\eta-1)} \exp\left(\eta - 1 - \frac{l}{\beta\gamma}\right) \tag{4}$$

where l is length and η, β and γ are parameters controlling the shape and steepness of the dome-shape.

A selectivity function can be thought of as a statistical weighting on the true size-at-age distribution. Uncorrected growth parameters could be estimated by using the above p.d.f. as a likelihood function to estimate the parameters used to compute μ_a, but length data sampled under a selectivity curve are biased by the selectivity curve whereby the true distribution of length-at-age is skewed by the selectivity function. That is:

$$\tilde{f}_a(l_{ai}) = \frac{f_a(l_{ai})S(l_{ai})}{\int_{-\infty}^{\infty} f_a(l')S(l')dl'} \tag{5}$$

where $\tilde{f}_a(l_{ai})$ models the density of the length of fish i of age a caught under selectivity curve $S(l) \in (0,1)$, which is a length-based selectivity function. The integral in the denominator ensures that the area under the curve is equal to one since the true density in the numerator is scaled by the selectivity curve. The likelihood presented in eqn. 5 is actually the L6 method (Troynikov, 1999, though in that paper it was derived using

Bayes' theorem).

7.2.3 Life History Types

As a direct comparison of the effectiveness of each of the presented bias reduction methods we simulated age-length structured fish populations based on three life-history types: cod (*Gadus morhua* L.), capelin (*Mallotus villosus*), and golden redfish (*Sebastes norvegicus*). Capelin is a small, pelagic fast growing fish with generally low variance in length distributions and high mortality; cod is a moderately fast growing benthopelagic fish that exhibits intermediate variance in length distributions and moderate mortality; and golden redfish is a long-lived, slow-growing fish with high variance in length-at-age and low mortality. Parameters for growth and mortality in these three life-history types were taken both from assessments performed for these species by the Icelandic Marine and Freshwater Research Institute (Hafrannsóknastofnun, 2019b) and from FishBase (Froese and Pauly, 2017b) and can be found in Table 7.3.

Table 7.3. Various parameters for creating age-length structured populations and sampling under the different selectivity curves. Parameters belong to the respecitve functions listed, which are described in the Introduction and Methods sections.

Parameter	Function	Unit	Cod	Capelin	Redfish
L_∞	Growth	cm	125.00	25.00	60.00
k	Growth	year^{-1}	0.15	0.40	0.07
t_0	Growth	year	-0.50	-0.10	-1.00
σ_∞	LAA Variance	cm	1.00	0.50	0.50
b	LAA Variance	cm	4.00	4.00	9.00
a'	LAA Variance	year	5.00	2.50	4.00
M	Mortality	year^{-1}	0.20	0.40	0.10
Min. Age	Age	year	0.00	0.00	0.00
Max. Age	Age	year	20.00	5.00	45.00
α	Asymptotic Selectivity	–	0.10	0.40	0.10
L_{50}	Asymptotic Selectivity	cm	40.00	10.00	20.00
η	Dome-Shaped Selectivity	–	7.00	5.00	6.00
β	Dome-Shaped Selectivity	–	0.50	0.30	0.30
γ	Dome-Shaped Selectivity	–	20.00	10.00	20.00

7.2.4 Simulation Scenarios

To test the effectiveness of the bias-correction methods we carried out simulation studies for each method using simulated length-at-age data for the three life-history strategies and sampled this data under the two selectivity curves shown in eqns. 3 and 4. We then used that data to estimate growth curves (using eqn. 1). For each life-history type, we simulated length-at-age data, performed sampling under various selectivity scenarios, and fit growth curves to this sampled data using the regular likelihood (i.e., L0–no bias-correction) and each of the bias-corrected likelihood approaches presented above. For comparison we also estimated growth curve parameters using data that were sampled with a constant selectivity of 1 across length. The VBGF (eqn. 1) was used to produce mean length-at-age, and individual lengths were distributed as normal for each age (see below). Populations were simulated as $N_{as} = N_0 \exp(-M_s a)$ where N_{as} is the number N at age a for life-history type s, M_{as} is the natural mortality for life-history type s at age a (as specified in Table 7.3), and N_0 is the initial number at age 0, which in this study was equal to 10,000 for all life-history types. These individuals were then assigned length according to a normal distribution – $\mathrm{N}(\mu_a, \sigma_a^2)$ where μ_a is the mean length at age a determined by the VBGF for each species and its respective growth parameters and σ_a was assigned as a hockey stick function of age a respective to each species as follows:

$$\sigma_a = \begin{cases} a\left[\dfrac{(\sigma_\infty - b)}{a'}\right] & a < a' \\ \sigma_\infty & a \geq a' \end{cases} \tag{6}$$

where σ_a is the age-specific variance in length at age a, σ_∞ is a maximum upper limit to σ_a, a' is the age at which the function reaches σ_∞, and b is the y-intercept for the left hand (i.e., ascending) side of the function. This function was used in simulations, and the parameters b, a', and σ_∞ were estimated in the likelihood optimizations.

The data produced by these simulations were then sampled with both asymptotic and dome-shaped selectivity curves (eqns. 3 and 4, respectively), and a vector of length-at-age data from the above simulations was multiplied by the appropriate selectivity for a respective length (see Supp. Materials 1 (S1) for an example data set and S2 for growth curves). We then employed three "sample size" scenarios by multiplying the sampled vector by a scalar, either 1, 0.1, or 0.01. Paired length-at-age data sampled from these simulations were used to estimate VBGF parameters under all bias-correction likelihood methods listed in Table 7.2, including the regular normal likelihood (i.e., L0–no correction), using data sampled with asymptotic, dome-shaped, and constant selectivity (with constant of 1) for comparison. Selectivity and growth parameters were carefully chosen for each life history type so as to reasonably induce some selectivity bias into the length-at-age data as both selectivity and growth can have an impact on the amount of selectivity bias (Frater and Stefansson, 2019). For methods L2 and L3 we used the correctly specified selectivity curve and parameters to determine the data that should be removed. For L5 we based truncation points for the definite integral off of lengths that were poorly sampled by the selectivity curve. For the L2 and L3 methods, poorly sampled data were assumed as those length values that were sampled at a selectivity of $\leq 50\%$ (consistent with Gwinn et al., 2010), while truncation points for the definite integral in the L5 method were defined as the L_{50} parameter times 0.6 for asymptotic selectivity or lengths that were sampled at a selectivity of greater than 0.3 for dome-shaped selectivity. The truncation points for L5 were determined somewhat arbitrarily, but represent a selectivity of roughly 0.3 for both selectivity types. We decided on this level as a selectivity of 0.5 (as was used in L2 and L3) often did not result in a sufficient number of data points for dome-shaped selectivity. The choice of cutoff for the L5 method relies on scientific expertise and available data (Schueller et al., 2014), so we utilized these to our best ability to create a realistic simulation for this particular method.

For each life-history type, sample size, likelihood type and selectivity curve shape we replicated the above procedure 1000 times and estimated growth curve parameters using the `optim()` function in R (R Core Team, 2017). Parameters for the classic VBGF from eqn. 1 were estimated as were parameters for eqn. 6 used to produce σ_a. The default Nelder-Mead optimization algorithm was used in R's `optim()` function. If any method did not converge according to the convergence criteria returned by `optim`, then all estimates for that iteration were dropped. Likelihood scenarios were coded explicitly as described in Table 7.2.

7.2.5 Misspecification Scenarios

Since many of these methods use selectivity, either directly or indirectly, in the bias-correction technique it was of interest to determine how each of these methods performed when selectivity was incorrectly specified. To test performance under misspecification we performed additional simulations using bias-correction methods where misspecification would be applicable and comparable (methods L2, L3, L5, L6). In order to test the performance of each of the bias-correction methods under selectivity misspecification we also performed simulations where selectivity was incorrectly assumed. For each likelihood method that either included selectivity or truncated poorly sampled data, we performed three different misspecification types: 1) incorrect L_{50} parameter for asymptotic selectivity, 2) incorrect doming for dome-shaped selectivity, and 3) misspecified selectivity curve shape. Age-length structured populations were simulated and sampled in the same manner as above. For asymptotic selectivity, L_{50} was assumed as either 1) 0.5x the correct L_{50}, or 2) twice the correct L_{50} parameter. For dome-shaped selectivity, domes were assumed as either shallower or steeper. Dome-shape was assumed steeper by multiplying the η parameter by 2 and dividing the β parameter by 2, while they were assumed shallower by dividing the η parameter by 2 while multiplying the β parameter by 2. The true selectivity curves along with misspecified ones can be seen S2. For scenarios where shape of the selectivity curve was misspecified we simply assumed the wrong shape (i.e., assumed asymptotic selectivity for dome-shaped samples and assumed dome-shaped selectivity for data sampled under asymptotic selectivity). In these instances the values used for incorrectly specified selectivity functions were those shown in Table 7.3.

Each of the above misspecification scenarios was replicated 1000 times for each life history type using the same procedure as for the simulation scenarios. Iterations where any optimizations did not appropriately converge were again discarded.

7.2.6 Performance Metrics for Simulations

To determine the performance of each bias-correction likelihood method and comparatively evaluate them we looked at the error in estimated parameters compared to true values. For both L_∞ and k we computed relative error (RE) as:

$$RE = \frac{(\hat{\theta} - \theta)}{\theta}$$

where $\hat{\theta}$ is the estimated parameter and θ is the true parameter. To assess the bias and bias-correction of t_0 we used true error:

$$\text{Error} = \hat{\theta} - \theta$$

to avoid confusion and aid in interpretation as t_0 is generally a negative parameter. All simulations, analyses, and figures were performed and produced in R-3.6.1 (R Core Team, 2017).

7.2.7 Selectivity bias-correction in cod

In addition to the above simulations we provide an example using real data for the Atlantic cod stock in Iceland. Growth curves were fit to data from Atlantic cod captured in Icelandic waters using each method from Table 7.2. These data were collected during annual spring groundfish surveys performed by the Icelandic Marine and Freshwater Research Institute (Hafrannsóknastofnun, 2018). These surveys are benthic trawls that have been performed annually in March since 1985, and are specifically designed to sample cod. To determine the appropriate selectivity curve for use in the bias-correction methods we fit a stock assessment model to cod using data from these spring trawl surveys as well as autumn trawl surveys and commercial port sampling. The stock assessment model was fit using the modeling software Gadget (Begley, 2012) and is similar to the model results presented in Taylor et al. (2007). The best fitting model utilized asymptotic selectivity for the spring trawl surveys (same equation as eqn. 3) with an α of 0.055 and an L_{50} of 50 cm. Using this selectivity curve and the length-at-age data from spring trawl surveys we estimated growth parameters for cod using each method as well as the regular normal likelihood (i.e., no assumed selectivity) for comparison. L2 and L3 were both greatly affected by the presence of an outlier with length 72 cm at age 1, which is an unrealistic length-at-age for Atlantic cod. This data point was therefore removed.

7.3 Results

7.3.1 Simulation Scenario Results

The majority of optimizations for the various likelihood types converged appropriately in the correctly assumed simulation scenarios. Almost all of the methods compared had convergence rates well above 99%. The L1 method in redfish had the lowest at 96.4% convergence with asymptotic selectivity and sample size scenario using a scalar of 1, and 98.6% convergence with asymptotic selectivity and a 0.1 proportional sample size. All other scenarios had convergence above 99% (see S3).

Selectivity bias in the simulation scenarios induced some error in growth curve parameter estimates, especially for the k and t_0 parameters and more so under dome-shaped rather than asymptotic selectivity (Fig. 7.14). When error due to selectivity bias was present the type 3 methods (L5–L6, Table 7.2) tended to correct the selectivity bias-induced error better than the type 1 or 2 methods (L1-L4). The L6 method especially showed low bias and variance in parameter estimates with the bias decreasing towards 0 as sample sizes increased. For most life-history types, the type 1 and especially type 2 methods tended to induce far greater error into growth parameter estimates even compared with the non-corrected likelihood (L0). In general, increased sample sizes led to decreases in both bias and variance for most methods.

The L6 method seemed to correct estimates of σ_a and reduced the variance in these estimates as well (Fig. 7.15). This method along with the L5 method tended to outperform the others at estimating length-at-age, although the L6 method seemed to have greater variance in estimates of older aged-fish than the L5 method for both selectivity types and across life-history types (Fig. 7.16).

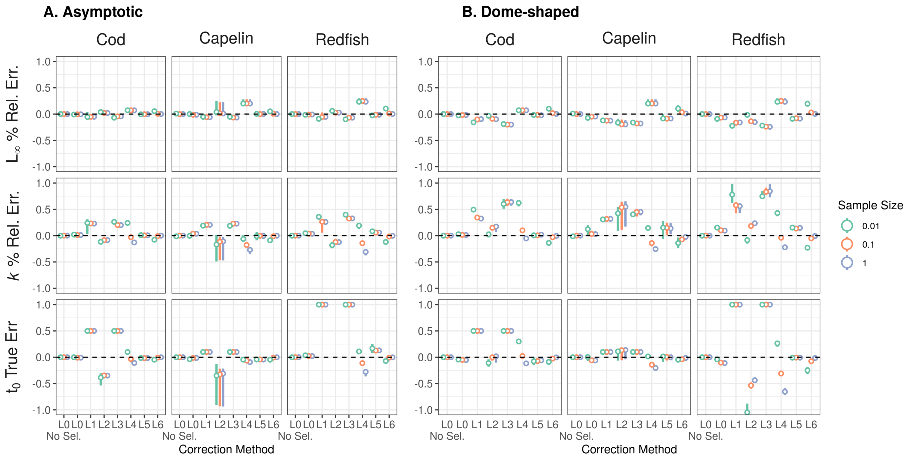


Figure 7.14. Median and 95% error in growth curve parameter estimates using all methods presented for both A) asymptotic selectivity and B) dome-shaped selectivity. Each column of figures represents one of the life-history types and each row represents a growth curve parameter.

7.3.2 Misspecification Scenarios

The type 2 methods (L2–L3) seemed to have the greatest error and variance in growth curve parameter estimates under misspecification scenarios (Figs. 7.17–7.19). The L6 and L5 methods performed similarly with both having parameter estimate bias near 0 for most parameters and life histories as well as low variance in error. The L5 method seemed to outperform the L6 method at lower sample sizes. In general, the L6 method tended to show much more of an effect of decreasing bias with increased sample sizes, and the performance of the L6 method seemed to improve with increasing sample size more so than the other methods. Interestingly, parameter estimates seemed to maintain a certain overall direction of bias during the misspecification scenarios rather than trending positive bias for one type of scenario and negative bias for another, or vice versa. Parameter estimates tended to be biased in a single direction across misspecification scenarios rather than showing opposite bias for selectivity curves that were either to one side of the true inflection point or another (asymptotic) or

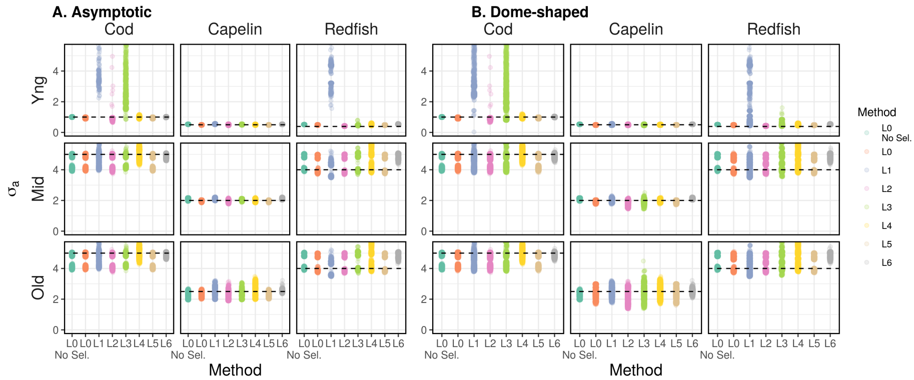


Figure 7.15. Estimated σ_a for three representative ages across each method and selectivity type. Ages are represented as young (yng), middle age (mid), and maximum age (old) and are respectively as follows for each life history type: cod – 0, 10, 20; capelin – 0, 3, 5; redfish – 0, 25, 45.

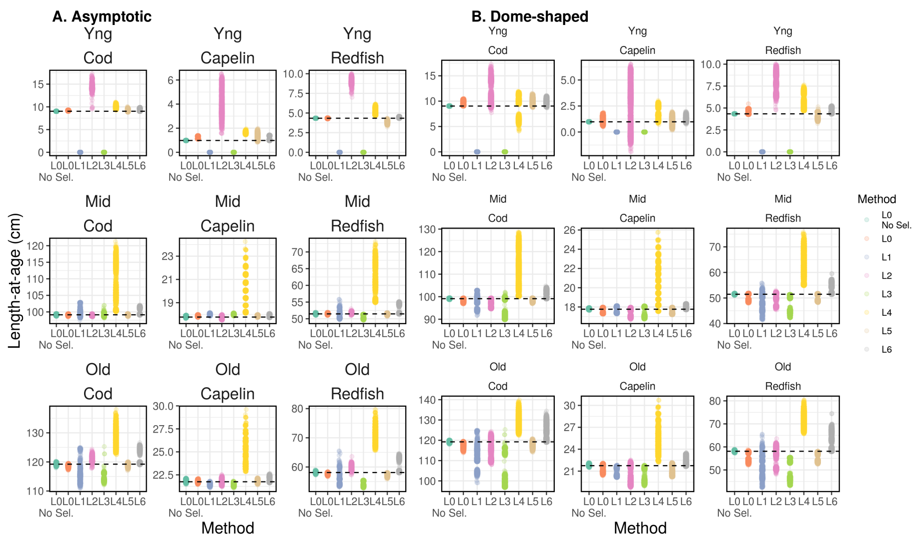


Figure 7.16. Estimates length-at-age for three representative ages for each method and selectivity type. Ages correspond with the same age classes in 7.15

Table 7.4. Absolute value of growth parameter estimates for Atlantic cod sampled in Icelandic waters on annual spring bottom trawl surveys.

Method	Lik. ID	L_∞ (cm)	k (yr^{-1})	t_0 (yr)	b	a'	σ_∞
Normal (i.e. no correction)	L0	127.44	0.15	0.02	3.35	3.00	10.26
Fix t_0	L1	132.95	0.14	0.00	3.10	3.12	10.31
Rm Small Samples	L2	125.76	0.15	0.04	0.00	3.00	9.99
Rm Small Fix t_0	L3	120.02	0.17	0.00	9.10	1.52	10.12
Fix L_∞	L4	129.00	0.15	0.13	4.68	1.81	10.12
Lik$_{Trunc}$	L5	124.06	0.17	0.44	0.00	3.90	10.36
Troynikov	L6	139.19	0.13	0.14	1.02	4.54	10.89

steeper/shallower (dome-shaped). Overall, bias was not as severe as we expected it to be under misspecification.

7.3.3 Cod Example

All methods estimated growth parameter values for cod that were fairly similar to each other (Table 7.4). The L3 method estimated a much higher b parameter for the hockey stick function used to compute σ_a, which translates into a much higher variance in lower ages than the other methods, and the L5 method estimated a t_0 that was much more positive than the other methods. The L6 method estimated a L_∞ that was greater than, and a k that was less than, the other methods. Absolute values are listed in Table 7.4. Growth curves fit to the data used to optimize parameter estimates for each method can be seen in Fig. 7.20

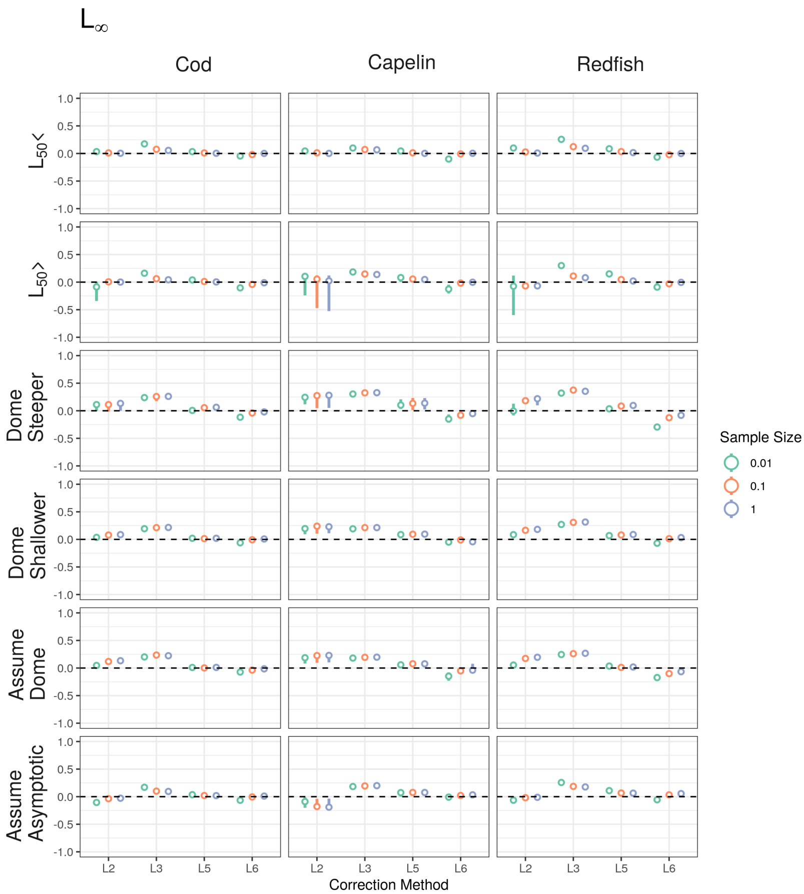
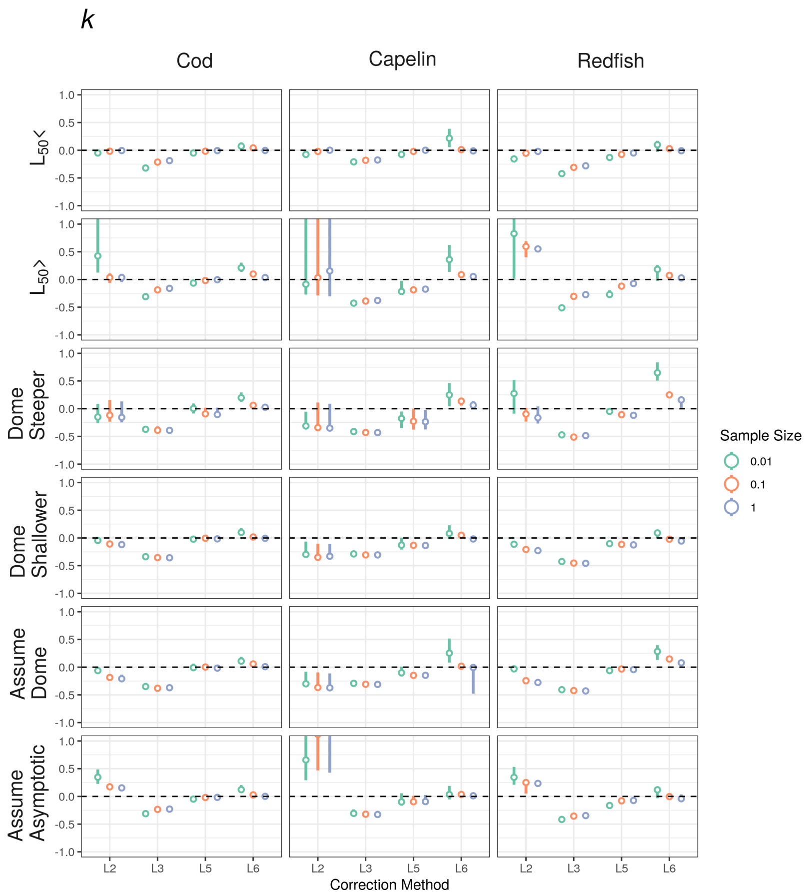


Figure 7.17. Median and 95% CI error in L_∞ estimates when selectivity was misspecified in the likelihood function. Each column represents a species and each row is a misspecification scenario. $L_{50} <$ and $L_{50} >$ are scenarios in which the L_{50} parameter was either decreased or increased, respectively, relative to the true value used to sample data with asymptotic selectivity. "Dome steeper" and "dome-shallower" represent scenarios in which the dome-shaped selectivity was assumed either steeper or shallower than correct, and "assume dome" and "assume asymptotic" are scenarios in which the incorrect selectivity curve shape was assumed ("assume dome" means that samples from an asymptotic selectivity curve were used for data, but dome-shaped selectivity was assumed in the likelihood).



s

Figure 7.18. Median and 95% CI error in k estimates when selectivity was misspecified in the likelihood function. Rows, columns, and methods are the same as in Fig. 7.17.

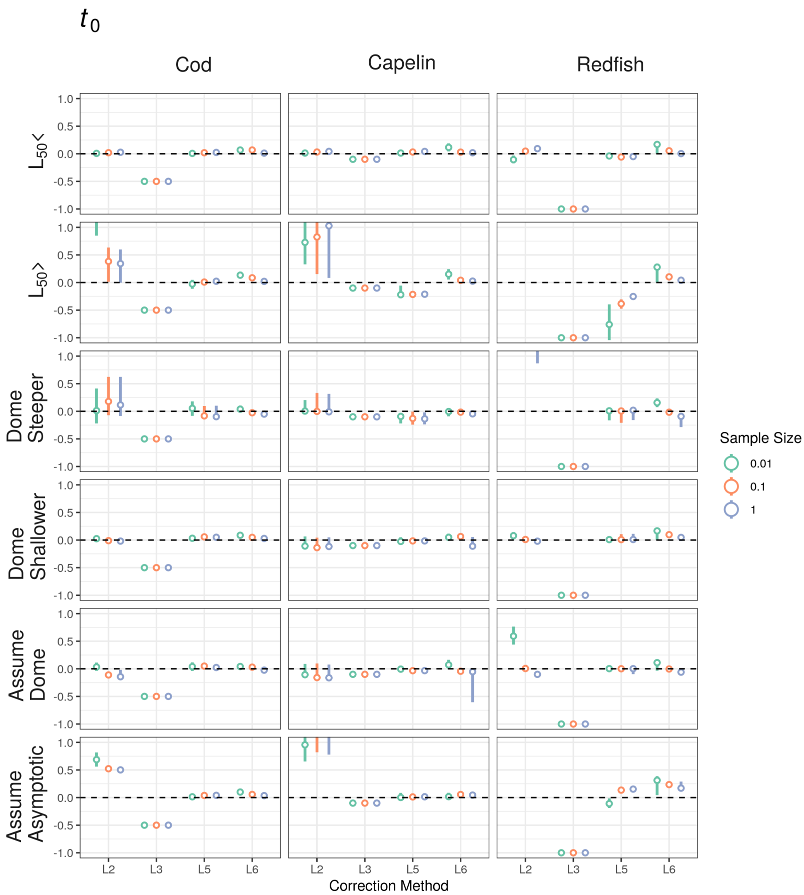


Figure 7.19. Median and 95% CI error in t_0 estimates when selectivity was misspecified in the likelihood function. Rows, columns, and methods are the same as in Figs. 7.17 and 7.18.

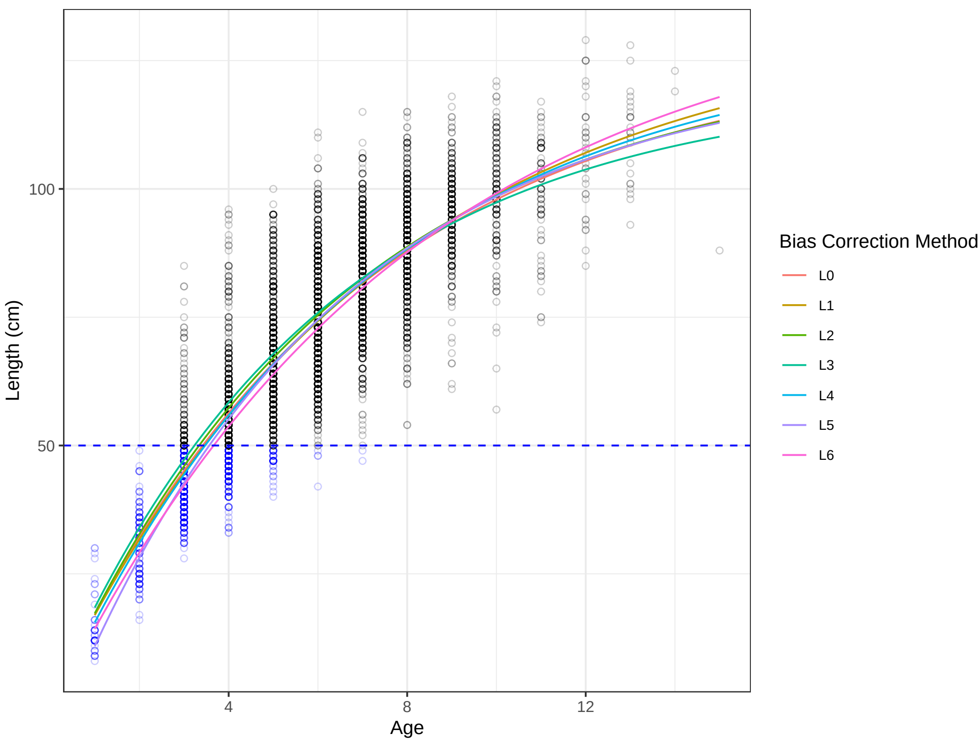


Figure 7.20. Fit of the VBGF to standardized spring trawl survey length-at-age data. Each colored line represents one of the methods from Table 7.2. The blue points are data that were removed for the L2 and L3 methods.

7.4 Discussion

This paper represents the most comprehensive comparison of selectivity bias-correction methods for growth parameter estimates to date. Overall in the simulation scenarios, the L6 method, the likelihood method proposed by Troynikov (1999), generally had the lowest bias in parameter estimates. The L5 and L6 methods all had similar error bias with some (usually) minor differences in variance in certain parameter/life-history combinations although the L6 method exhibited consistently low variation in parameter estimate bias. In general, the type 1 and 2 methods, which fix parameters or remove small samples sizes, respectively, performed less well than the type 3 likelihood methods. This was true across life history types and selectivity curve shapes. In Gwinn et al. (2010), the L4 method performed best out of the four methods tested in that paper, at least for dome-shaped selectivity. In our study the L4 method estimated parameters rather poorly across life-history and selectivity types. This is perhaps due to the fact that estimated parameters of the VBGF are notoriously correlated with each other (Ratkowsky, 1986), such that fixing one parameter might lead to overcompensation by one or several of the others. In general this issue could be avoided by using a re-parameterization of the VBGF as in Schnute (1981), but since some of the methods in this study fixed parameters specific to the classic VBGF (such as L4) we maintained use of the original equation.

The cod and redfish life-history types had the greatest bias and variance in results especially for the L_∞ and k parameters. This is likely due to the combination of both selectivity and length-at-age distributions of these species as aspects of both growth and selectivity have been shown to influence selectivity bias in length-at-age data (Frater and Stefansson, 2019).

In general, the type 3 methods outperformed the other method types at reducing error in the estimates of growth parameters, and, of these, the L6 method performed probably a bit better than the L5 method especially with greater sample sizes. However, there were several scenarios, especially under misspecification, where the L5 method outperformed the L6 method at lower sample sizes. The type 3 methods generally performed fairly well in misspecification scenarios overall as both methods had consistently low bias and variance in errors across misspecification scenarios.

The superior performance of the Type 3 methods is likely due to the fact that they capture the statistical aspects of selectivity bias on growth within the likelihood to reverse its effect. Size-based selectivity induces bias into the distribution of size-at-age by acting as a weighting factor on the frequency distribution of the true size-at-age. The L6 method attempts to dampen this effect by including selectivity directly in the likelihood equation. This method can be used with any probability density function for size-at-age and any selectivity curve (as long as selectivity is known). The L5 method is similar to the L6 method except that it divides the normal p.d.f. likelihood by a truncated integral of the normal p.d.f. (McGarvey and Fowler, 2002; Schueller et al., 2014). The truncation is specified by the user and is dependent on assumed shape of the selectivity curve as well as the lengths at which the observed size-at-age data are suspected to be poorly sampled due to selectivity. For example, if dome-shaped selectivity is assumed for the sampling procedure then one would truncate lengths at both the lower and upper ends of the observed values.

All six methods except the L2 and L3 methods estimated fairly similar growth curves to the length-at-age data for Atlantic cod from Iceland. The L2 method estimated a curve that was much too shallow and seemed too high at both the lower and upper ends, and the L3 method estimated a growth curve that appears as though L_∞ might be too low. The other methods were all fairly close to each other, though.

This study focused on likelihood-based methods for correcting selectivity bias in fish, but alternative methods exist. One way to correct for selectivity bias is by circumventing it altogether as is done in some integrated stock assessments. Certain integrated assessment models such as Gadget (Begley, 2012) incorporate both growth and selectivity into the models, and the output should be an unbiased representation of the stock (assuming the model is correct). These parameter estimates, along with recruitment and estimates of F, could be suitable then to use for something like calculating yield-per-recruit and MSY (Elvarsson et al., 2018). This is an effective method as any effect of selectivity on growth should be accounted for in the model.

The main issue with the L6 method, and to some degree, the L5 method as well is the necessity of a knowledge of selectivity. This is indeed an issue and has been criticized in the past (Gwinn et al., 2010). For stocks that do have a knowledge of selectivity these could be useful methods for correcting growth parameter estimates. Alternatively, the L6 method could prove useful as a diagnostic test for examining the accuracy of selectivity parameters estimated by a stock assessment model. If one were to compare growth parameters estimated from this method (using the selectivity estimated in an assessment model) to the growth parameters estimated or used by that same assessment model, then they should be comparable. If they are not, then either growth or selectivity (or both) in the assessment model can be assumed to be misspecified.

We have shown here that the bias in growth parameter estimates caused by selectivity can be reduced and/or reversed by factoring the effect of selectivity out of the likelihood used to estimate growth parameters. The methods presented here are potentially useful for stocks in which selectivity is known and where growth curve parameter estimates are desired outside of an integrated assessment model. The ability of the type 3 methods to accurately estimate true growth curve parameters is partially dependent on the type and shape of the selectivity curve and the life history of the species. The L6 method also warrants further investigation as a potential diagnostic tool for understanding confoundedness of growth and selectivity in integrated assessment.

Acknowledgments

Funding for the lead author to undertake his Ph.D. studies was received from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 613571 for the project MareFrame and from the European Commission's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 634495 for the project Science, Technology, and Society Initiative to minimize Unwanted Catches in European Fisheries (Minouw) as well as from the Icelandic Research Fund (Rannís, No. 152039051). The lead author is grateful to the second author for acquiring said funding. We are also grateful to Bjarki Elvarsson and Pam Woods at the Icelandic Marine and Freshwater Research Institute for providing survey data on Atlantic cod used in this paper.

7.5 Supplementary Materials

S1. Example of simulated length-at-age datas

Table 7.5 shows example simulated length-at-age data under selectivity used for sampling. Number shows the true number of a given length at a given age. The Asymptotic Sel. and Dome-shaped sel. columns provide the respective selectivity at that length, and the corresponding sample columns provide the number that are sampled under each respective selectivity type.

Table 7.5. Example of length-at-age data that were used as well as selectivity and number of fish sampled under selectivity.

Age	Length	Number	Asymptotic Sel.	Asymptotic Samples	Dome-Shaped Sel.	Dome-Shaped Samples
1	23	91	0.15	14	0.13	12
1	26	153	0.20	30	0.20	30
2	40	84	0.50	42	0.65	54
2	44	28	0.60	17	0.77	22
3	50	47	0.73	34	0.91	43
3	53	52	0.79	41	0.96	50
4	62	50	0.90	45	1.00	50
4	75	1	0.97	1	0.85	1
5	62	3	0.90	3	1.00	3
5	75	25	0.97	24	0.85	21
6	82	12	0.99	12	0.72	9
6	90	1	0.99	1	0.57	1
7	82	18	0.99	18	0.72	13
7	92	8	0.99	8	0.53	4
8	84	7	0.99	7	0.68	5
8	94	12	1.00	12	0.49	6
9	88	7	0.99	7	0.61	4
9	95	9	1.00	9	0.48	4
10	96	11	1.00	11	0.46	5
10	98	13	1.00	13	0.42	6

S2. Growth and Selectivity

The true growth used to simulate length-at-age data are shown below.

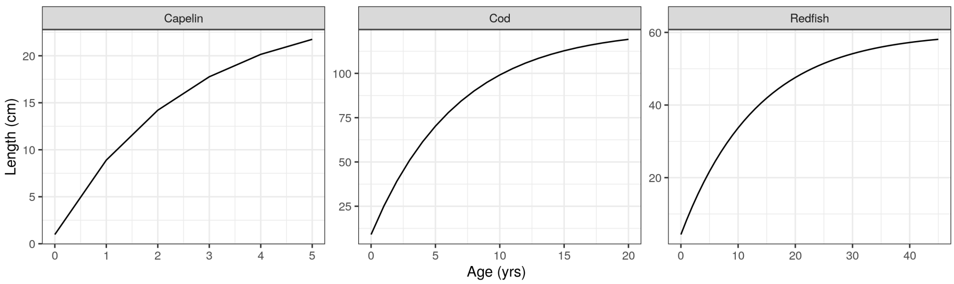


Figure 7.21. Growth curves used to simulate populations for each life history type. Note that difference in both x- and y-axes across facets.

The selectivity curves used to sample the simulated populations created by growth curves in Figure 7.21 are shown in blue below. The incorrectly assumed selectivity curves used during the misspecification scenarios are indicated by the dashed red lines.

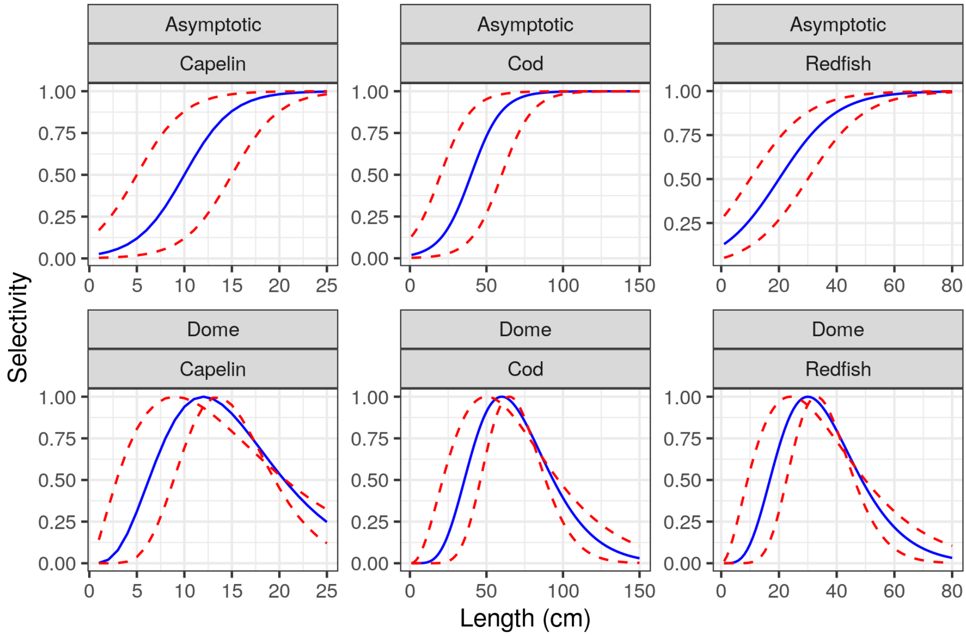


Figure 7.22. Selectivity used to sample simulated populations. The blue line represents the true selectivity used in sampling and the dashed red line represent the the incorrect selectivity shapes assumed during the misspecification scenarios.

S3. Convergence Rates

Table 7.6 shows the convergence rates for all likelihood methods included in this study under each life history type, sample size, and selectivity scenario. Convergence rates for all methods under all scenarios were quite high, with the lowest being 96.4%. Most methods and scenarios were $\geq 99.9\%$

Table 7.6. Convergence rates for each likelihood method under each life history type, sample size, and selectivity scenario.

Species	Sample Size	Selectivity	L0	L1	L2	L3	L4	L5	L6
Cod	0.01	Dome-shaped	1	0.99	1	0.996	1	1	1
Cod	0.01	Asymptotic	1	1	0.998	0.998	1	1	1
Cod	0.1	Dome-shaped	1	0.999	0.999	1	1	0.996	1
Cod	0.1	Asymptotic	1	1	1	0.999	1	1	0.999
Cod	1	Dome-shaped	0.999	1	1	0.999	1	1	1
Cod	1	Asymptotic	1	1	0.996	0.999	1	0.999	0.999
Capelin	0.01	Dome-shaped	0.998	1	0.999	1	1	0.999	1
Capelin	0.01	Asymptotic	1	1	0.998	0.999	1	1	1
Capelin	0.1	Dome-shaped	0.998	1	0.998	1	1	0.999	1
Capelin	0.1	Asymptotic	1	1	1	1	1	1	1
Capelin	1	Dome-shaped	0.999	1	0.999	0.998	1	1	1
Capelin	1	Asymptotic	1	1	1	1	1	1	1
Redfish	0.01	Dome-shaped	1	0.997	0.998	1	1	1	1
Redfish	0.01	Asymptotic	0.997	0.996	0.995	1	0.999	1	0.998
Redfish	0.1	Dome-shaped	0.999	1	0.999	0.999	0.999	1	0.999
Redfish	0.1	Asymptotic	1	0.986	1	0.999	1	0.999	1
Redfish	1	Dome-shaped	0.998	0.999	1	1	0.999	1	1
Redfish	1	Asymptotic	1	0.964	1	1	0.999	1	1

8 Paper III - Drivers of growth for Atlantic cod (*Gadus morhua* L.) in Icelandic waters – a Bayesian approach to determine spatiotemporal variation and its causes

Frater, P.N.[1] Hrafnkelsson, B.[2] Elvarsson, B.Th.[3] Stefansson, G.[2]

[1]University of Iceland, Dunhagi 5, 107 Reykjavik, Iceland

[2]Faculty of Physical Sciences, University of Iceland, Dunhagi 5, 107 Reykjavik, Iceland

[3]Icelandic Marine and Freshwater Research Institute, Skúlagata 4, 101 Reykjavik, Iceland

Abstract

Growth in Atlantic cod (*Gadus morhua* L.) has been well studied, and the presence of time-varying growth has been noted. However, drivers for variation in growth across time has been hotly debated, and, while there are a number of hypotheses about what causes time-varying growth in *G. morhua*, no consensus seems to have been reached. Recent interest in growth studies, especially time-varying growth, has led to a resurgence in the study of growth in fishes across time. Here we use a Bayesian approach to gain an understanding of differences in growth across time, among areas, and between genders in *G. morhua* found in Icelandic waters. We compare twelve models that fit the von Bertalanffy growth function (VBGF) to length-at-age data using the deviance information criteria (DIC). We found that the "best" models (those with the lowest DIC scores) were those that accounted for time-varying growth (both by year and cohort), differences among areas, and differences between gender. Standardized mean fitted parameter values were compared to potential explanatory variables using a redundancy analysis (RDA). The RDA indicated that capelin recruitment and spawning stock biomass (SSB) are highly correlated with patterns in the k parameter of the VBGF, which controls the steepness of the growth curve (i.e. how fast fish grow at younger ages). The L_∞ parameter of the VBGF was correlated with cod landings in the model that incorporated for time-varying growth by year, but L_∞ was correlated with estimated cod recruitment in the model that includes cohort for time-varying growth. Additionally, the t_0 parameter was negatively correlated with sea surface temperature, which indicates that at warmer temperatures predicted length at age 0 is greater. These results indicate that growth of *G. morhua* is driven by different mechanistic factors throughout the stages of their lives.

Keywords: Atlantic cod, *Gadus morhua*, growth, von Bertalanffy, Bayesian, Iceland

8.1 Introduction

The study of growth in organisms dates back decades to the seminal work of von Bertalanffy (1938) and even long before this. Since then numerous studies of somatic growth described by the von Bertalanffy growth function (VBGF) have been employed in a wide variety of organisms including, but not limited to: fish (Lockwood, 1974; Kimura, 1980; Sainsbury, 1980; Chen et al., 1992), trees (Pienaar and Turnbull, 1973; Borah and Mahanta, 2018), birds (Ricklefs, 1968), reptiles (Shine and Charnov, 1992), corals (Chadwick-Furman et al., 2000), and a number of other organisms (Stokes, 1996). Within fisheries the VBGF is one of the most fundamental and widely used models

(Francis, 2016). Growth of a fish species has been shown to affect or be correlated with a wide range of other life-history aspects such as natural mortality (Pauly, 1980; Gislason et al., 2010), the shape of maturity ogives (Roff, 1984; Jensen, 1996), reference points for the management of a species (Beverton and Holt, 1957; Deriso, 1987), reproductive output (Barneche et al., 2018), and even food web trophic structure (Barneche and Allen, 2018).

The concept of growth is central to fisheries management and stock assessment procedures (Sparre and Venema, 1998). Historically, time-varying growth in stock assessment models has been included as either empirical weight- or length-at-age (Kuriyama et al., 2016; Lorenzen, 2016), but the recent push towards using integrated assessments recommends internally modeling as many processes as possible (Maunder and Punt, 2013). Within integrated assessments where growth is modeled internally the temporal variation in growth is often not taken into account (Stawitz et al., 2019). That paradigm is increasingly being challenged (Lorenzen, 2016) as more studies are reporting issues of bias or imprecision with misspecifiation in growth (Stawitz et al., 2019); however, having an understanding of if and how growth varies across time, and identifying potential mechanisms, is crucial for further development.

A number of variables can induce growth plasticity and variability across both time and space such as: fisheries-induced evolution (Heino et al., 2015), environmental forcing (Brander, 1995; Clark and Hare, 2002), and population density (Mehl and Sunnanå, 1991). In turn, the assumption of using either time-varying or constant growth can have a substantial impact on the outcome of assessment or management advice (Whitten et al., 2013). It is important to have an understanding of how growth changes across time in managed stocks. In addition, spatial variation in growth has been noted in *G. morhua* and other species (Brander, 2000, S. Campana *pers. comm*). There could be a number of reasons why growth in *G. morhua* might change across time. Increases in temperature have been shown to increase growth in *G. morhua* up to a certain extent (Brander, 1995;

Björnsson et al., 2007), and changes across time in the ocean temperature associated with climate change (Wijffels et al., 2016) could lead to increased growth and/or growth rate in *G. morhua*. Alternatively, increased fishing pressure over the past half-century could lead to a decline in the number of fish at older ages (Hsieh et al., 2010) or fisheries-induced evolution of life-history traits such as growth (Heino et al., 2015). In this study we fit Bayesian models of the VBGF to length-at-age data from Atlantic cod (*Gadus morhua* L.) caught in Icelandic waters and compare models to assess whether growth varies across both time and space as well as between gender and compare the results of these Bayesian models to potential explanatory variables to gain insight into what drives time-varying growth in *G. morhua* in Iceland.

Within the past decade or so a number of studies have taken a Bayesian approach to fitting the VBGF to describe growth within fisheries (including bivalves, molluscs, and other non-fishes often associated with fisheries; Zhang et al., 2009; Alós et al., 2009; Jiao et al., 2010; Linde et al., 2011; Tang et al., 2014; Lopez Quintero et al., 2017). This approach not only allows for an effective way to fit VBGF parameters to data, but also provides a credible interval (CI) on parameter values (Gelman et al., 2014). While there have been a number of studies that have fit growth data using Bayesian techniques, most have described the growth of individual fishes using one of two methods. Fish are either: a) measured for length throughout time using mark-recapture studies (Bal et al., 2011; Tang et al., 2014) along with a variant of the VBGF known as the Fabens version (Fabens et al., 1965), or b) harvested destructively, measured for length, and the otoliths (or other aging parts) used to back calculate age and length from measurements of annulii in the growth rings (Alós et al., 2009; Linde et al., 2011; de Zárate and Babcock, 2016; Contreras-Reyes et al., 2018; Harris et al., 2018). Surprisingly, not that many studies have fit a basic growth model to direct conditional length-at-age data using a Bayesian framework (i.e. individuals measured directly for both length and age at a single time point, but see Thorson and Minte-Vera, 2016; Lopez Quintero et al., 2017);

however, in many cases a wealth of this data exists from standardized surveys and there is useful information to be garnered from this.

Here, we present point estimates and CIs of parameters for the VBGF associated with *G. morhua* found in Icelandic waters. In addition we analyze point estimates of parameters against possible explanatory variables to determine correlations that might influence the growth of *G. morhua*. These results are important for understanding the growth of cod in Icelandic waters, which is necessary for management of the stock.

Parameter estimates and CIs from these models could be especially useful as the VBGF is commonly used to predict growth within stock assessment models where growth parameters are often estimated as part of an optimization procedure (Begley, 2012) or by using Bayesian Markov chain Monte Carlo (MCMC, Methot and Wetzel, 2013). Therefore, when fitting a stock assessment on a species it would be useful to have CIs of growth parameters to guide the setting of upper and lower parameter bounds (if optimizing an objective function) or to have a distribution for growth parameters to use as a prior if performing MCMC estimation of the parameters.

8.2 Materials and Methods

The Icelandic Marine and Freshwater Research Institute (Hafrannsóknastofnun) has performed standardized bottom trawl surveys for groundfish every spring since 1985 and autumn since 1995 (Hafrannsóknastofnun, 2018). Given the temporal extent of this data, twelve variations of growth models for *G. morhua* were explored and compared. These multi-level models included aspects accounting for time-varying growth (by either year or cohort), spatial differences, and gender differences. The most basic is a simple VBGF fit to the mean length-at-age for all data (M0). VBGF models were also fit to mean length-at-age for each year, cohort, spatial division, gender, and all combinations thereof except for year by cohort (Table 8.7).

8.2.1 Data Collection

The data used for this study was sampled from both standardized trawl surveys dating to 1985 and commercial catch samples dating to 1960. Trawl surveys were originally designed to assess the abundance of *G. morhua*, and a random subsample of fish caught in each trawl are measured for age and length. Age of fish (at time of sampling) from these surveys was determined by counting annuli growth rings in the otolith of sampled fish. Fish caught in the standardized trawl surveys are typically 1-12 year old fish, so to supplement the data to include older ages we also included length-at-age data sampled from commercial catches of *G. morhua* caught in Icelandic waters, which often includes larger fish of older ages.

Data on length-at-age from fish caught on vessels other than standardized trawl surveys (i.e. commerical vessels) for *G. morhua* have been sampled dating back to 1899 (Hafrannsóknastofnun, 2018). Samples are taken throughout each year from commercially landed fish as port samples and are measured for age and length. Commercial fishermen are often interested in catching larger fish to maximize their fishing efforts and offset costs. This additional data supplements the standardized trawl survey by providing length-at-age at older ages. However, this data must be used with caution as it is neither standardized nor randomly sampled due to commercial interest. The use of commercial catch sample data could introduce bias into the growth models in this study because of the targeted nature of commercial catches (i.e. if schools of fish tend to aggregate by length rather than age and are spatially disjunct, then vessels would likely target schools of greater length, assumed to correlate with biomass, regardless of whether they are older or not). In addition, commercial catch samples would be more likely to not contain younger fish due to selection on larger fish, and younger fish that are caught by commercial vessels are also likely to be longer than the mean for that age class as selectivity curves tend to sample fish larger than the mean of a respective yearclass for those sizes where the selection curve is steepest (Sampson, 2014).

Despite these biases in the data there is useful knowledge to be gained from them. For instance, all of the length-at-age data for older fish (i.e. 15-20 years old), which greatly aids in the estimation of the parameter for maximum length, come from commercial catch samples. In addition, including these data extend the range of sampling back to the earlier part of the 20th century which allows for a comparison of growth in *G. morhua* across time.

As will be shown below we fit models of VBGF to length-at-age data for both sexes; therefore, we subset data to those entries where gender was known. In addition, we calculated the year class of each measured fish by subtracting the age from the year sampled. For example, if a fish was aged to be 16 years in 1986, then that fish is from the 1970 cohort. This is not a perfect solution for calculating the year a fish was born given that error in age measurements exists, but we were interested in looking at trends over longer time periods so the proposed method should be sufficient for the objectives at hand. The number of samples prior to the 1960 cohort were variable and generally low, so only data from the 1960 cohort and afterward were used. In addition, the number of older fish in more recent cohorts is few as fish have not yet aged older than the number of years before the current sampling date (i.e. fish born in 2010 can be no older than 5 years if sampled in 2015). Given these constraints only data from cohorts between 1960 and 2007 were used. Within this timeframe some years or cohorts consistently resulted in VBGF parameter estimates that were unrealistically high (e.g. $L_\infty > 300$), so these years or cohorts were removed. In addition, a few years or cohorts in the late 1970s/early 1980s were drastically skewed by initial sampling design created by the

Icelandic Marine and Freshwater Research Institute. The years and cohorts that were removed from the models for these various reasons were: year = (1968, 1981); cohort = (1966-1967, 1976-1977, 1984-1985)

Data subset to known gender and the above specified years and cohorts were used for all models. This resulted in 363,904 individual samples measured for age, length, and gender across 46 years representing 40 cohorts and among 14 ecologically and hydrologically unique areas (Fig. 8.23).

8.2.2 Models

We set up a series of Bayesian models to assess growth in *G. morhua* across both time and space and between genders. The growth curve used is the basic deterministic VBGF

$$L_a = L_\infty(1 - \exp(-k(a - t_0)))$$

where L_a is mean length at age a, L_∞ is the length as age approaches infinity (i.e. the maximum theoretical length attainable by fish), k is a steepness parameter controlling how quickly the curve reaches L_∞, and t_0 is a scaling parameter that controls the length at age 0.

Given that some age classes had skewed or heavy-tailed distributions (especially both old and young see Supporting Information (SI) Fig. 8.26) a log-normal distribution was used to model the data. Therefore, the data distribution for a single length L_a was:

$$p(L_a|L_\infty, k, t_0, a, \sigma_a^2) = \text{log-N}(L_a|L(L_\infty, k, t_0, a), \sigma_a^2)$$

where σ_a was treated as a nuisance variable, but was modeled internally. Fish often exhibit an asymptotic relationship between variance in length and age, and our data were no exception (see dispersion plots in SI; Figs. 8.27–8.31).) However, when variance of log length is plotted against age the relationship tends to be more hump-shaped (see example R code in SI). To appropriately model this we used a modified Ricker curve which consisted of the basic Ricker curve plus an intercept: $\sigma_a = (\alpha a e^{-\beta a}) + c$, where a is age, α and β the parameters of the Ricker model, which were modeled internally, and c was an intercept (also modeled internally). The intercept was added as not all hypothetical variances at age 0 are equal to 0 – similar to the t_0 parameter of the VBGF. Parameters for modeling σ_a as a Ricker function are not reported here.

Since *G. morhua* is such a heavily studied species we were able to find enough data on growth to use informative priors for the parameters of the VBGF. We obtained data on growth parameters from Fishbase (see *Gadus morhua* growth on Fishbase, Froese and Pauly, 2017b) to use as prior distributions for the VBGF parameters . Parameter values for L_∞, k, and t_0 were pulled from FishBase and subset by country to obtain recorded values of these parameters for *G. morhua* in the North Atlantic Ocean (i.e. we subset to include only values from Iceland, Norway, Faroe Islands, and Greenland). From these parameter estimates we chose distributions that appeared to well match the densities for each parameter to use as prior distributions. For L_∞ and k we decided to use gamma distributions (Figs. 8.32-8.33), but the histogram for t_0 did not fit a standard distribution (i.e. data density was multimodal and skewed, but generally centered around 0; Fig. 8.34). t_0 should generally be close to 0, so we decided to use an uninformative prior represented as the uniform distribution from -3 to -0.01 (setting t_0 directly to 0 caused issues with the log-normal data distribution as $\log(0) = -\infty$). For L_∞ and k we obtained prior density parameter estimates by minimizing the negative log-likelihood for the respective distributions using the `optim` function in R. This resulted in the following prior distributions

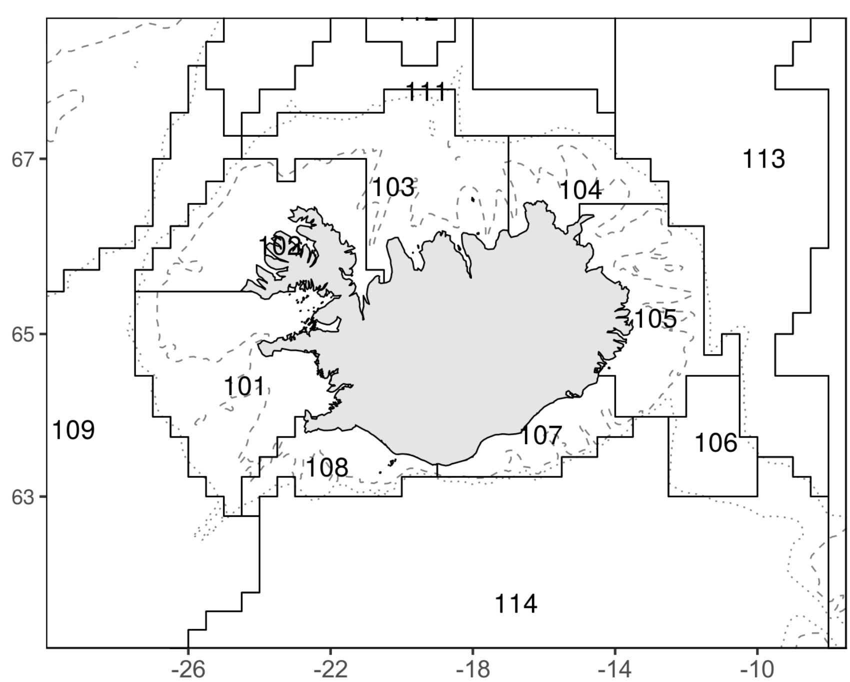


Figure 8.23. Map of Iceland and surrounding waters with areas used in the models delineated. These areas are formally known as divisions and are those originally described in (Stefánsson et al., 1997)

.

$$L_\infty \sim \text{Gamma}(22.61, 0.187)$$

$$k \sim \text{Gamma}(4.499, 25.761)$$

$$t_0 \sim \text{Unif}(-3, -0.01)$$

Models were parameterized using the above prior and data distributions. The base case model (M0) was fit using all data combined. Models were then fit to data resolved to time (cohort or year), area (division), or gender (M/F) as well as combinations of these factors (e.g. M1 - Year; M5 - Year x Area; M10 - Gender x Year x Area;) and separate parameters were fit for each particular data subset respective to the model. The exception to this was that cohort and year were not combined as these both represent aspects of time-varying growth. These combinations of factors resulted in twelve distinct models (Table 8.7).

Posterior distributions for the models were obtained using Markov Chain Monte Carlo (MCMC) to estimate parameters. To perform MCMC estimation we utilized JAGS within R (v3.4.3, R Core Team, 2017) with package R2jags (Su and Yajima, 2015). For each model, four MCMC chains were run for 50,000 iterations with a burn-in of 5,000 and a thinning rate of 10. Gelman-Rubin statistics ($\hat{R}$) were obtained from `jags` output to ensure values were close to 1, which is a metric for convergence in MCMC chains (Gelman et al., 1992, 2014). Effective sample sizes were also obtained to assess convergence (Gelman et al., 2014). Models were compared to assess the goodness of fit using DIC (Lunn et al., 2013), which was also obtained from the output of the `jags` function. The model with the lowest DIC is considered to have the best fit to the data.

8.2.3 Analyzing Model Parameters

To assess drivers of deviations in model parameters across time and space we performed a redundancy analysis (RDA) of model parameters against potential explanatory variables. RDA is essentially a constrained principal components analysis (i.e. where dependent variables are constrained on the explanatory variables in multivariate space) and is useful for determining correlations among multiple dependent and independent variables. Since parameters of the VBGF are notorious for being correlated with each other this method is well-suited for describing correlations of VBGF parameters with explanatory variables while taking into account the correlation among parameter values themselves. The variables that we included in the RDA were: sea surface temperature (SST), measured zooplankton biomass (Zoo), deviations in temperature (T_{dev}) and salinity (S_{dev}), mean fishing mortality F_{ave}, cod landings in Iceland, estimated cod recruitment (Rec_{cod}) and spawning stock biomass (SSB_{cod}), and estimated capelin recruitment (Rec_{cap}) and spawning stock biomass (SSB_{cap}). Values for SST were monthly mean estimates (°C) pulled from NOAA ESRL (2019) and were spatially resolved to 1° latitude x 1° longitude. This data was used to compute mean annual SST for each year and area, which was used in RDA models. Zoo, T_{dev}, and S_{dev} were all values taken on an annual basis from five stations at the Siglunes section north of Iceland (Hafrannsóknastofnun, 2019a). Zoo was measured as the average zooplankton biomass (g dry weight m^{-2}), and both T_{dev} and S_{dev} are the deviations in temperature (°C) and salinity (PSU) from the 1961-1980 average. Unfortunately, data on Zoo, T_{dev}, and S_{dev} were not as spatially resolved to the extent that cod length-at-age data were collected, so they each represent a single reading per year (i.e. no area mean). Cod landings in thousands of tonnes was collected each year by the Icelandic Marine Research Institute. Rec_{cod}, SSB_{cod}, Rec_{cap}, and SSB_{cap} were estimated by the Icelandic Marine Research Institute each year as part of the country's assessment procedure. Data on cod landings and estimates of recruitment and SSB were also single values per year. In the RDA models parameter values were scaled and centered (using the `scale()` function in R). This was done because VBGF parameter values tend to be quite disparate and scaling made visualization and interpretation of the RDA results easier.

8.3 Results

Most MCMC chains for the models converged appropriately based on the Gelman-Rubin statistic ($\hat{R}$, Table 8.7). Models with higher average $\hat{R}$ tended to also have a greater standard error for $\hat{R}$ as well as a greater model DIC. In general, the two factor models (i.e. M5–M9), but also M1, the model that only incorporated year, seemed to have a greater average $\hat{R}$ indicating that these models likely did not converge as well as some of the others. The three factor models, M10 and M11, had some of the lowest values for average $\hat{R}$, and these models also had the lowest DIC. M10 and M11 are the models that are considered best-fitting and will be the only models further analyzed.

Table 8.7. Model variables, $\hat{R}$ statistics, N_{eff} statistics, DIC, and p_D for all Bayesian models fitted with MCMC and compared using DIC. The two models with the lowest DIC were those that used all variables: M10 and M11.

Model	Model Variables	$\hat{R}_{ave}$	$\hat{R}_{SE}$	$\bar{N}_{eff}$	$N_{eff_{SE}}$	DIC	p_D
M0	Base	1.004	1.31e-03	2188.3	876.0	2.647e+06	6.04
M1	Year	1.038	1.44e-02	3696.8	286.0	2.613e+06	329.91
M2	Cohort	1.019	5.45e-03	4444.7	387.0	2.618e+06	317.64
M3	Area	1.025	8.05e-03	6473.3	751.0	2.589e+06	96.64
M4	Gender	1.003	9.82e-04	6054.2	1860.0	2.645e+06	12.04
M5	Year x Area	1.025	3.95e-03	14321.0	82.6	2.932e+06	51960.41
M6	Cohort x Area	1.031	5.38e-03	13830.0	95.1	2.921e+06	27656.89
M7	Gender x Area	1.001	2.63e-05	14903.0	342.0	2.936e+06	123.13
M8	Gender x Year	1.053	2.58e-02	14526.0	206.0	3.007e+06	1440.39
M9	Gender x Cohort	1.051	1.43e-02	14237.0	236.0	3.015e+06	3905.90
M10	Gender x Year x Area	1.009	5.85e-04	9129.2	84.1	2.535e+06	5855.32
M11	Gender x Cohort x Area	1.005	1.92e-04	8573.7	87.6	2.548e+06	5668.39

Both M10 and M11 incorporated time, area, and gender into fitting the VBGF. M10 used year as the time-varying component in the model, and M11 used cohort. The median $\hat{R}$ value for both of these models was 1.001, and the standard error of $\hat{R}$ values for both of these models was lower than any other model. When plotting parameter values from these models across time, area, and gender there appear to be difference across all of these components (Fig. 8.24). Note that for both M10 and M11 some of the divisions had very little data in certain years (e.g. division 113 prior to 1980). In these particular time-area combinations parameter values tended to take on the values of the prior distribution rather than establishing a posterior distribution based more on the data.

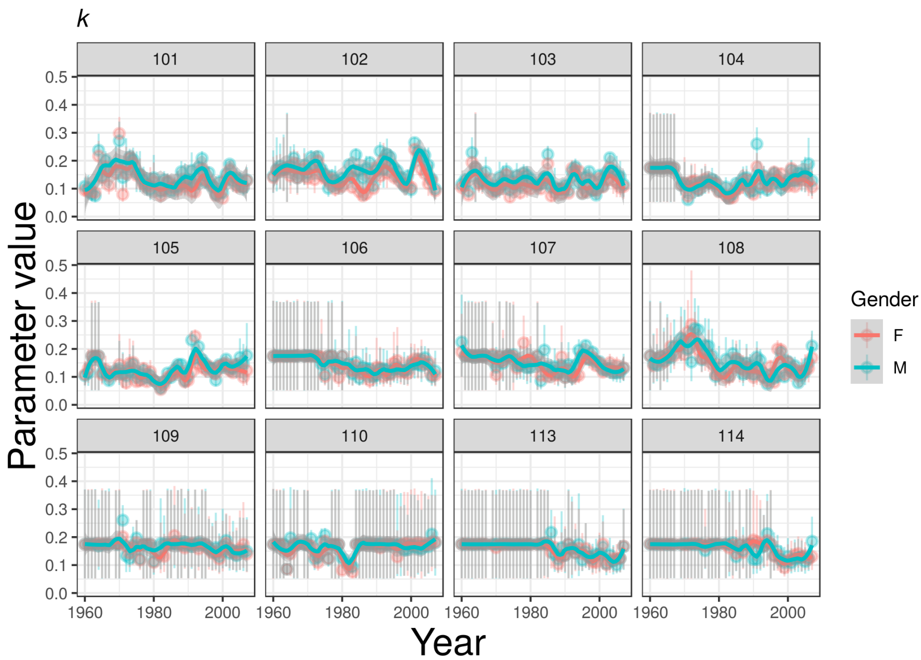


Figure 8.24. Posterior distributions for the values of the k parameter in M10 - gender x year x area where parameter values vary across time. Each panel represents a different region around Iceland corresponding to the divisions shown in Fig. 8.23. This figure does not explain what is causing spatiotemporal variation and differences in growth between gender; merely reflecting that it exists.

Results of the RDA for both M10 and M11 indicated that parameter values L_∞, k, and

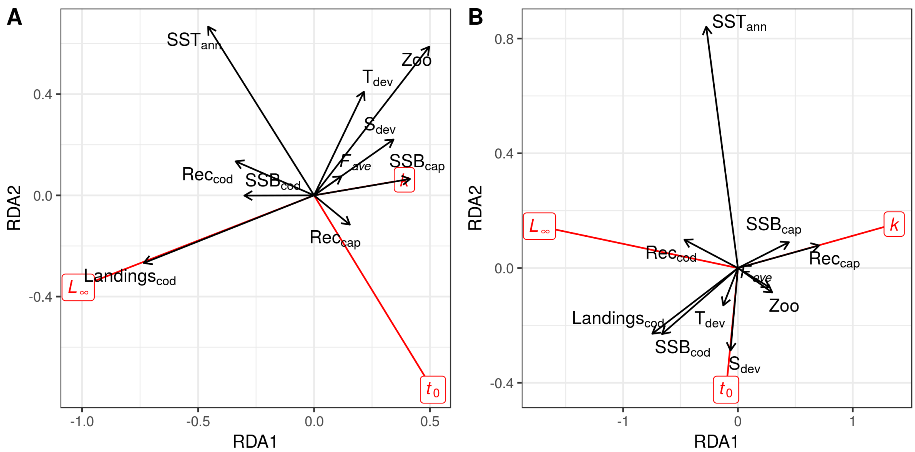


Figure 8.25. Biplots of first and second ordinate components of a redundancy analysis (RDA) for scaled and centered parameters of the VBGF fitted for models: (A) Gender x Year x Area (M10), and (B) Gender x Cohort x Area (M11). The k parameter of the VBGF is highly correlated with estimated capelin SSB in both M10 and M11 and with estimated capelin recruitment in M11. L_∞ is highly correlated with cod landings in M10 and is correlated with estimated recruitment of cod in M11.

t_0 were correlated with each other, but after accounting for this there were certain explanatory variables that were highly correlated with certain VBGF parameters. In M10, SSB_{cap} was highly correlated with k, and cod landings was highly correlated with L_∞ (Fig. 8.25A). SST also seemed to be negatively correlated with t_0 in M10. In M11, the k parameter for the VBGF was correlated with SSB_{cap} and was highly correlated with Rec_{cap} (Fig. 8.25B). L_∞ was also correlated with Rec_{cod}, and t_0 was highly correlated with S_{dev} and negatively correlated with SST in M11. The proportion of variance explained by the RDAs was low for M10 (4%), but was higher for M11 (12%). The results from both RDAs indicate that prey abundance has a strong influence on the k parameter of the VBGF, density dependence is a potential driver for L_∞. In addition, SST seems to be negatively correlated with t_0 indicating that *G. morhua* at age 0 will be longer when sea surface temperatures are greater.

8.4 Discussion

This study assesses fit of the VBGF to length-at-age data from *G. morhua* in Icelandic waters using Bayesian models and DIC. Models that incorporated some aspect of time-varying growth (i.e. year or cohort), spatially-varying growth, and differences between gender (models M10 and M11) had the best fitting models both in terms of DIC and $\hat{R}$ values. Models that used one or none of these specifications tended not to fit the data too poorly, but models that only used two out of these three factors seemed to have the poorest fit to the data both in terms of DIC and average $\hat{R}$. Here, we show that growth not only varies across time in *G. morhua*, but also across space and between genders. In addition, we show, using constrained ordination, that estimated capelin abundance, density dependence, and SST are potential drivers of growth in cod.

Growth plasticity across time is increasingly recognized in fishes (Thorson and Minte-Vera, 2016) and can be attributed to both ecological and evolutionary forces (Lorenzen, 2016). Temperature has been shown to have a direct and substantial impact on the growth rate of *G. morhua* (Brander, 1995; Björnsson et al., 2007). The impact of this environmental forcing combined with increasing ocean temperatures (Seidov et al., 2017) could reasonably lead to the hypothesis that growth rate steepness (k), maximum length-at-age (L_∞), or both would increase across time. However, we did not see this trend in our data. The k parameter in the best fitting models, however, was highly correlated with estimated recruitment and SSB of capelin. Capelin is a key prey item and major food source for *G. morhua* in the North Atlantic (Bogstad and Gjøsæter, 2001; Rose and O'driscoll, 2002; Vilhjálmsson, 2002). The liver condition of northeast Arctic cod has been linked to capelin biomass (Yaragina and Marshall, 2000), and other studies have indicated that capelin biomass could be important for growth of cod (Dalpadado and Bogstad, 2004). Additionally, density dependence has been shown to impact the growth of *G. morhua* with higher abundance leading to depressed growth (Sinclair and

Swain, 1996; Anderson et al., 2002). We found a correlation among L_∞ and either cod landings (in M10) or Rec_{cod} (in M11). Both of these correlations are coherent with their respective models. In the model that uses cohort to address time-varying growth (M11) Rec_{cod} is correlated with L_∞, but in the model that uses year (M10) landings is highly correlated. The strong recruit classes that are driving L_∞ in the cohort-based model are likely ending up as greater landings in the year based model. In either case the abundance of cod is positively correlated with L_∞. This is counterintuitive to the hypothesis that increased abundance depresses growth in *G. morhua*. This is likely an artifact of either sampling or model fit. In years with greater landings there are likely to be more fish caught that are larger and of older age simply due to probability. These older, larger fish would have a strong effect on driving the estimate for the L_∞ parameter as there are not always large samples for these older, larger fish. Hence, when abundances are high, more fish are caught in fisheries and in surveys, and the L_∞ parameter is estimated larger. An alternative to this explanation is that abundance of *G. morhua* is driving growth more so in younger fish, which influences estimates of the k parameter. That is, greater abundance of *G. morhua* leads to slower growth at young ages, which corresponds to a lower value for k in that year (or cohort). Since k is correlated with L_∞ in the VBGF, then a lower k will likely be associated with a higher L_∞ making it appear as if abundance of *G. morhua* is positively correlated with L_∞. This negative correlation between abundance of *G. morhua* and estimated k parameter values seems more parsimonious as density dependence would likely have the greatest effect on growth of *G. morhua* at younger ages, which would be captured by the k parameter.

Lastly, the t_0 parameter was negatively correlated with SST. There has been much research on the effects of temperature on growth of *G. morhua* (Brander, 1995) with optimum growth generally occurring at some intermediate temperature (Björnsson and Steinarsson, 2002; Björnsson et al., 2007) for larger sized fish. However, smaller fish tend to have an increased growth rate with increasing temperature (Brander, 2000). The negative correlation between t_0 and SST falls in line with these previous findings as estimation of the t_0 parameter is going to be largely driven by fish in the youngest age classes and smallest fish. Björnsson and Steinarsson (2002) and Björnsson et al. (2007) found a negative relationship between optimum temperature for growth and size of *G. morhua* whereby smallest fish tended to have higher optimum temperatures for growth. Our results seem to fall in line with this. A negative relationship between t_0 and SST would translate to a greater length at age 0 at warmer SST (i.e. length at age 0 increases as t_0 decreases).

In addition to the effects of environmental forcing, fishing has also been shown to play a role in driving the evolution of life history traits in fish (Andersen and Brander, 2009) including growth (Sinclair et al., 2002b; Enberg et al., 2012; Hidalgo et al., 2014). *G. morhua* is a species with a long history of heavy exploitation (Cook et al., 1997), so an effect of fishing on heritable traits controlling growth seems plausible. We did not find evidence of an effect of fishing on growth of *G. morhua* as F_{ave} was not correlated with any of the VBGF parameters in our RDA models.

Based on our Bayesian models and the constrained ordination of estimated VBGF parameters on explanatory variables it seems that the growth of *G. morhua* is limited or affected by different drivers throughout the different stages in life. At the youngest ages temperature seems to have the greatest effect, whereas at young-middle ages prey abundance and possibly density dependence seems to matter most. It is not clear what is driving the differences in L_∞ as this parameter was correlated with abundance of *G. morhua*, but it is possibly an artifact of either the model (i.e. correlation with k) or sampling.

The paradigm shift in thinking about growth in fishes is important. Time-varying growth is speculated to have a large impact on the population dynamics and overall biomass of a stock (Shin and Rochet, 1998; Rice, 2011), and it is increasingly suggested that plasticity of growth be recognized and accounted for in stock assessment and fisheries management (Kuriyama et al., 2016; Lorenzen, 2016; Stawitz et al., 2019) as it can have potentially large and far-reaching impacts on management decisions.

The results of our Bayesian models are similar to those found by Thorson and Minte-Vera (2016) who performed a similar analysis on 25 species of marine fishes and found that a majority of them exhibited variation in year effects for growth. This model was parameterized slightly differently than that of Thorson and Minte-Vera (2016) as they looked at weight-at-age data and used a re-parameterization of the VBGF. However, in the discussion section of that paper they state that results might change if the model was parameterized differently. We look at length-at-age data in the current study and both cohort and year effects and the model parameterization is different from that of Thorson and Minte-Vera (2016). We did still find that the best-fitting models included time-varying growth. This lends credence to the shifting paradigm that time-varying growth is present (and possibly ubiquitous) in fishes (Lorenzen, 2016). Growth plasticity has also been previously shown for other groundfish stocks as well (Mehl and Sunnanå, 1991; Clark and Hare, 2002), and growth and condition in *G. morhua* has even been shown to be correlated with stock collapse (Brander, 2007; Morgan et al., 2018). A general theory of the causes of time-varying growth in fishes has not been established, but has been attributed to temperature (Brander, 1995), food availability (Weatherley, 1990; Gjøsæter et al., 2002) and density dependence (Swain et al., 2003) among other things (Lorenzen, 2016). The present study found that growth in *G. morhua* is possibly driven by different causes at different points in a fishes life with temperature having the greatest impact at the youngest ages and density dependence and prey abundance driving growth at young-middle ages.

There are two sources of concern with the study that we have presented here. The first is that data from commercial catch samples are used to fit VBGF models. The second is that the estimated parameters for VBGF parameters are correlated with each other. We address each of these in turn.

This study makes use of data from commercial catch samples. These data are therefore not random and are likely to be inherently biased as fisheries generally desire to maximize profit in any given outing. This will most likely lead to the data being comprised of not only older fish, but also data skewed towards larger sizes at younger ages (Sampson, 2014). This happens because the fishery tends to be size selective rather than age selective (e.g. size selectivity due to mesh or hook size). This is not likely to be much of an issue for older fish, but the younger fish that are harvested in commercial catch will likely be larger than the population mean at a given age. Despite this apparent bias the use of commercial catch data allows the time span of this study to greatly exceed what it would have been otherwise. Whereas standardized surveys in Iceland date to 1985 the data from commercial catches allowed us to analyze cohorts back to 1955 and years to 1960. We saw the use of this data as complementing and extending the standardized survey data available not only temporally but also across ages as standardized trawls only tended to catch fish up to about 12 years of age (which makes them somewhat biased in their own right). The commercial catch samples provided age data up to 20 years in some cases, which can greatly aid in fitting the L_∞ parameter. Combining gear types such as scientific surveys and fishing catches has also been shown to reduce bias and increase precision of growth parameter estimates (Wilson et al., 2015).

The other source of concern in this study, the correlation of VBGF parameters, is somewhat troubling as it could lead to the appearance of differences in growth across cohorts when perhaps the parameters for each cohort simply shifted in the MCMCs. It is visually notice-able that the L_∞ parameter for the 1976 and 1977 cohorts are greatly inflated compared to all other cohorts. Additionally, the k value for these cohorts are the two lowest estimated values for k among all cohorts. The results for these two cohorts likely stems from a relative paucity of data within those cohort years, which is reaffirmed by their relatively large credible intervals on the L_∞ parameters. However, the results are not sensitive to these two particular cohorts years. The DIC in the cohort model is substantially reduced when they are excluded from analyses, but M2 is still the best-fitting model even when they are included. Correlated parameters in the VBGF are nothing new, and alternative parameterizations have been suggested (Ratkowsky, 1986). We decided to use the traditional form of the VBGF as this is used in many stock assessment models (Begley, 2012) and one of the objectives for implementing this exercise was to produce a guide for bounds or prior distributions for potential use in stock assessment models. It would be interesting to test alternative parameterizations of the VBGF to see if all lead to time-varying growth. The alternative parameterization used by Thorson and Minte-Vera (2016) also led to time-varying growth (as a year effect) having the lowest DIC in the majority of stocks, but alternative parameterizations were not tested in that study. However, they also mention that their results could possibly change with different model formulation. We recommend implementing such a test with different model parameterizations to see if alternative formulations of the VBGF suggest or deter evidence towards time-varying growth.

Conclusions

This paper reports results of fitting the VBGF to length-at-age data for *G. morhua* caught in Icelandic waters. Using a Bayesian approach and DIC model comparison we found that a model fit to individual cohort length-at-age data was the most parsimonious approach to describing the growth of *G. morhua* while maximizing fit to the data. This study reveals the presence of time-varying growth in this stock and lends evidence to the importance of recognizing time-varying growth as present and possibly prevalent in fishes. We hope that these results will be of use to future studies on *G. morhua* growth as well as assessment and management of this stock.

Acknowledgements

The primary author would like to sincerely thank Giri Gopalan for technical assistance while assembling Bayesian models. Funding for the lead author to undertake his Ph.D. studies was received from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 613571 for the project MareFrame and from the European Commission's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 634495 for the project Science, Technology, and Society Initiative to minimize Unwanted Catches in European Fisheries (Minouw) as well as from the Icelandic Research Fund (Rannís, No. 152039051).

8.5 Supporting Information

8.5.1 Data Distributions

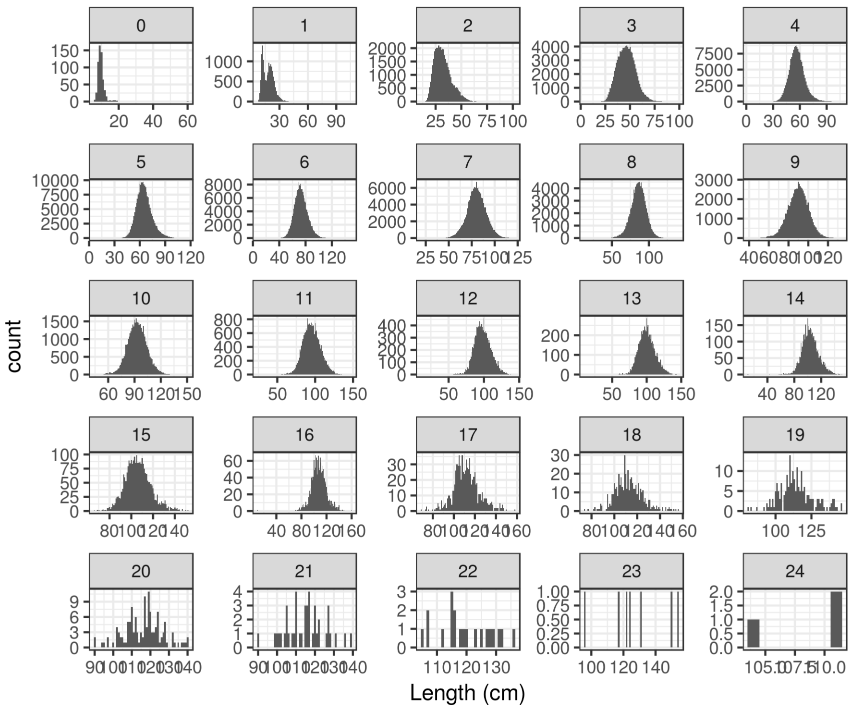


Figure 8.26. Histograms of length for each age in Atlantic cod from Icelandic waters. These were the data used to fit the VBGF in each model.

8.5.2 Dispersion Plots

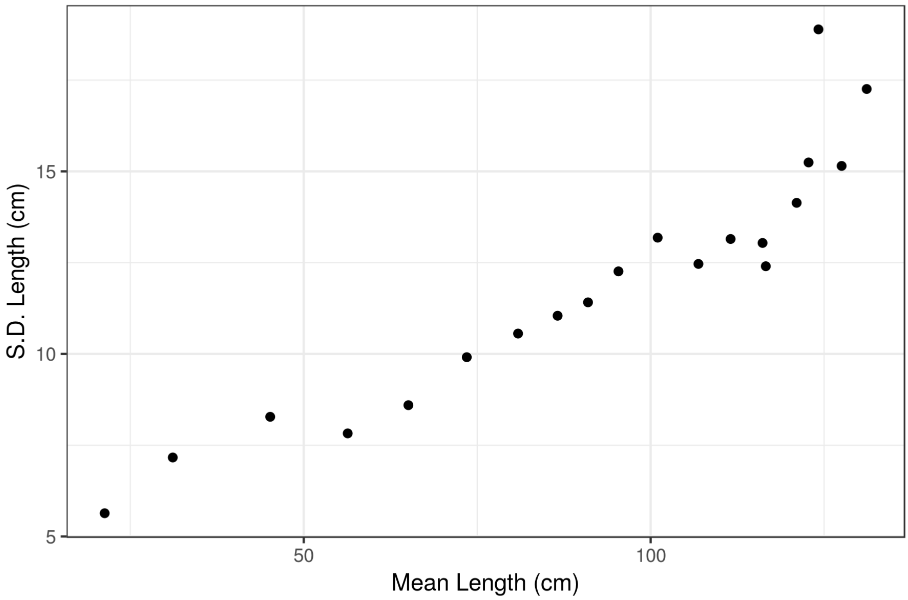


Figure 8.27. Dispersion plot of overall data. Each point represents an age.

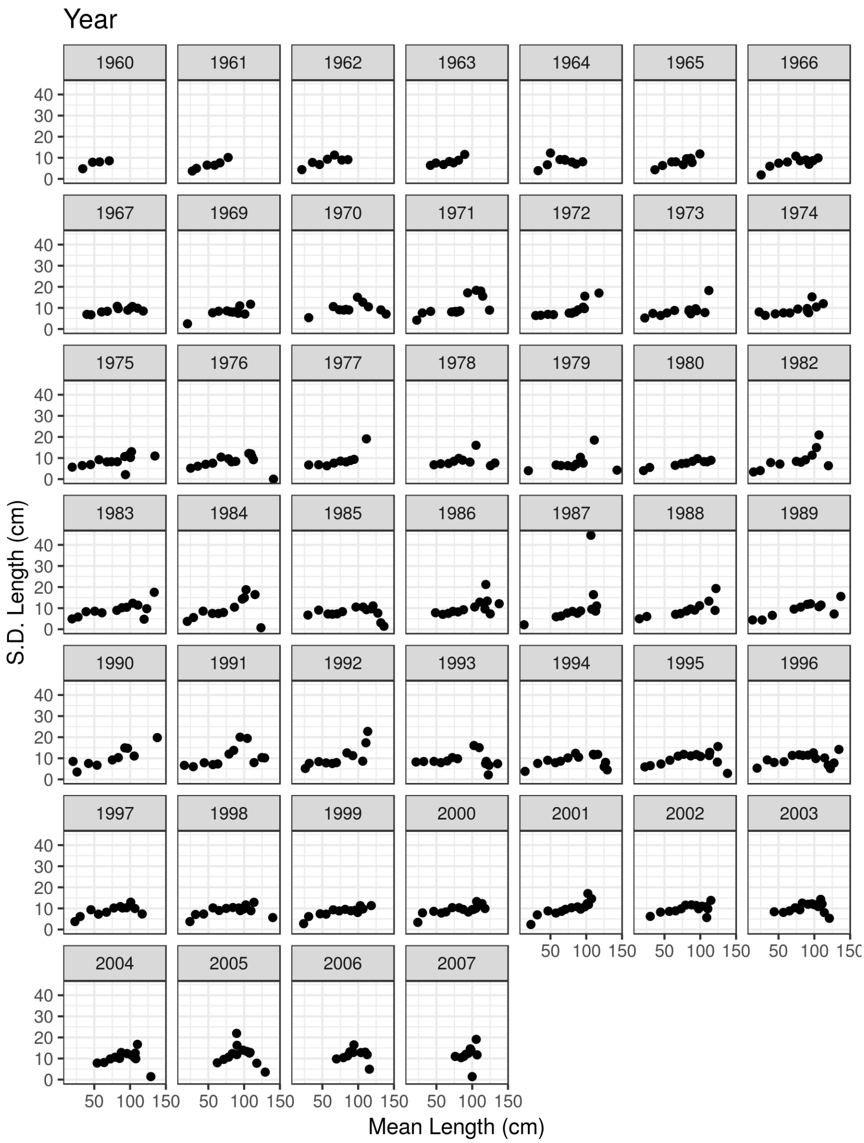


Figure 8.28. Dispersion plot of data for each year included in the models. Each point represents an age and each panel a year.

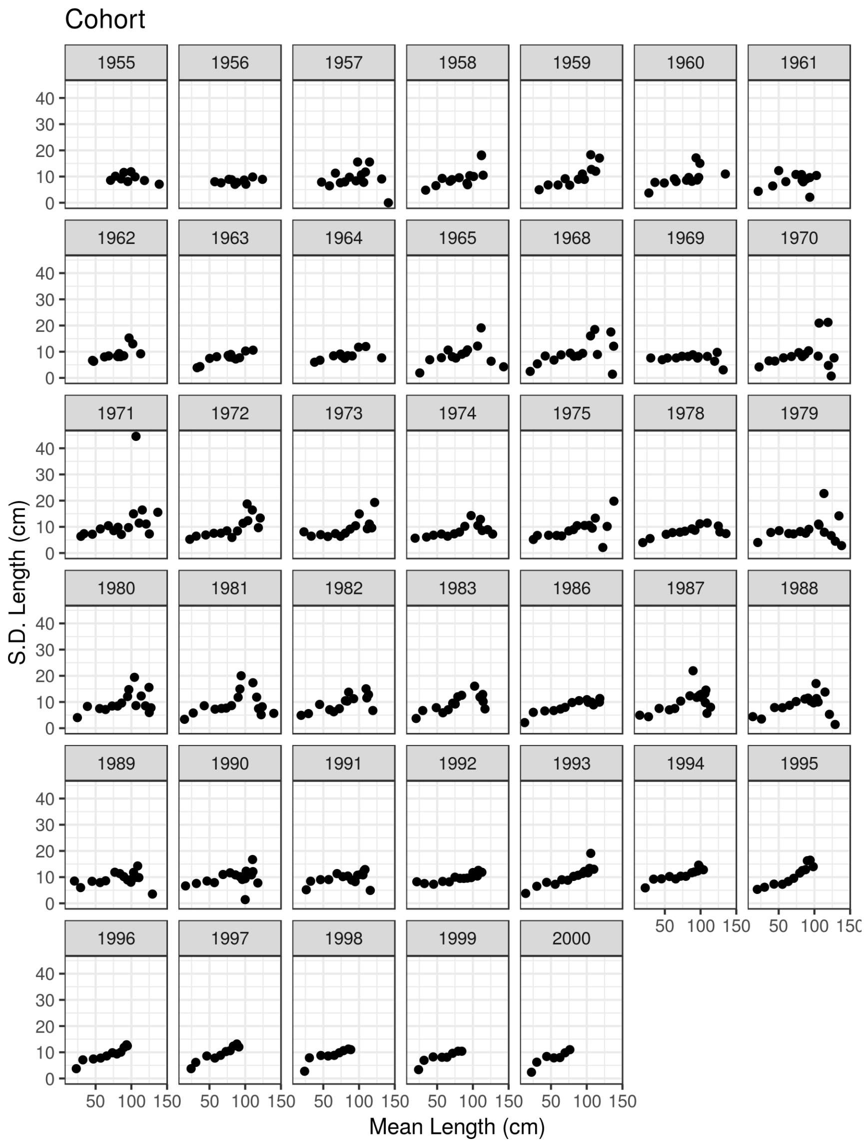


Figure 8.29. Dispersion plot of data for each cohort included in the models. Each point represents an age and each panel a cohort.

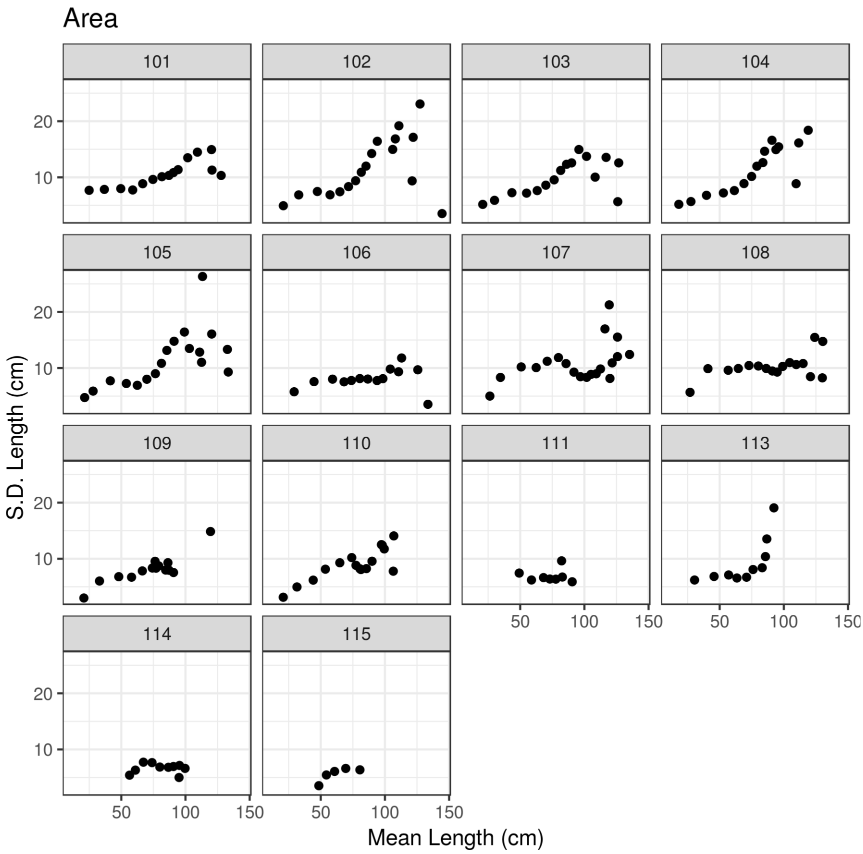


Figure 8.30. Dispersion plot of data for each year included in the models. Each point represents an age and each panel a division.

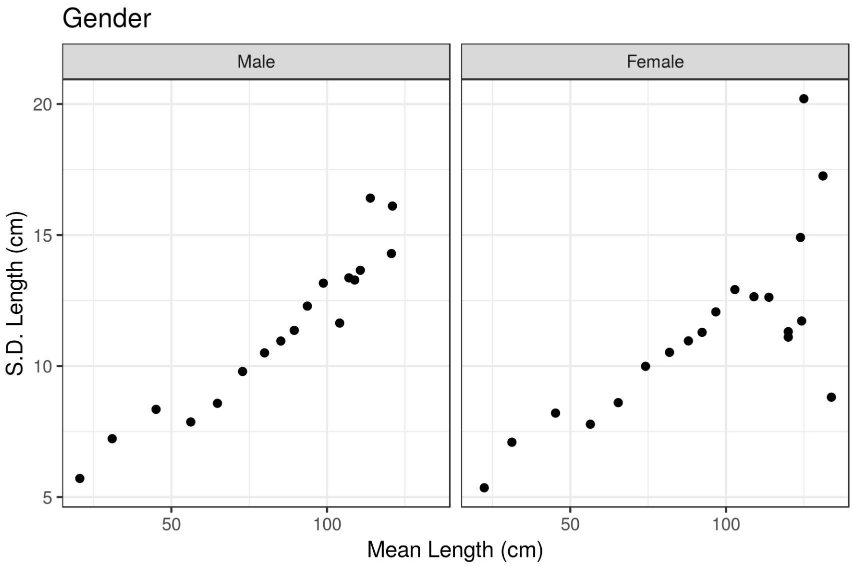


Figure 8.31. Dispersion plot of data for each year included in the models. Each point represents an age and data are facetted to gender.

8.5.3 Prior Distributions

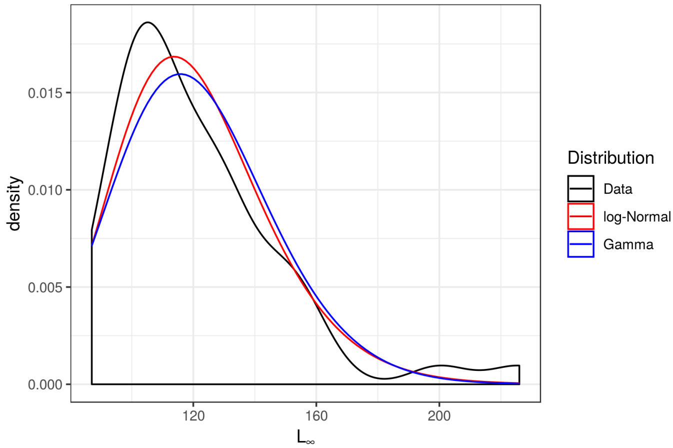


Figure 8.32. Prior distribution used for L_∞ based on data for the parameter collected from FishBase. Black line is the density of the L_∞ parameters collected from FishBase; the red line is the NLL minimized values of the log-N distribution fit to those data ($\mu = 4.775; \sigma = 0.204$); the blue line is the NLL minimized values of the gamma distribution fit to those data ($\alpha = 22.61; \beta = 0.187$).

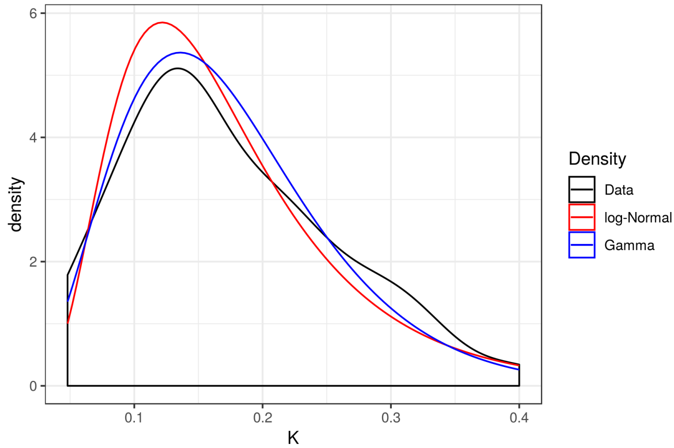


Figure 8.33. Prior distribution used for k based on data for the parameter collected from FishBase. Black line is the density of the k parameters collected from FishBase; red line is the NLL minimized values of the log-N distribution fit to those data ($\mu = -1.86; \sigma = 0.495$*); blue line is the NLL minimized values of the gamma distribution fit to those data (*$\alpha = 4.499; \beta = 25.761$*).*

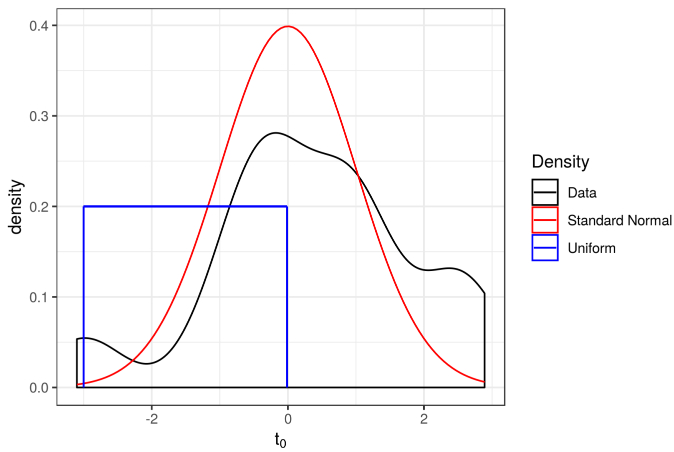


Figure 8.34. Prior distribution used for t_0 based on data for the parameter collected from FishBase. Black line is the density of the t_0 parameters collected from FishBase; red line is the the standard normal distribution ($\mu = 0; \sigma = 1$*); blue line represents an uninformative prior of a uniform distribution from (-3,-0.01).*

8.5.4 Example Showing Variance in log Length at Age

```
library(tidyverse)

# define ages and von Bertalanffy growth function
age <- 1:10
vb <- function(linf, k, t0, age) {return(linf * (1 - exp(-k *
    (age-t0))))}

set.seed(123)

# simulation length-at-age data to simple age-structured
    population
len_dat <-
  lapply(age, function(x) {
     # creates N_a and assigns length
    len <- round(rnorm(1000 * exp(-0.2*x), vb(125, 0.25, -0.5,
        x),
                     sd = ifelse(x <= 5, x, 5)))
    return(data.frame(age = x, length = len))
  }) %>% do.call("rbind", .) %>%
  mutate(log_length = log(length)) %>%
  gather(key = type, value = length, -age) %>%
  group_by(age, type) %>%
  summarize(mnlen = mean(length),
            varlen = sd(length))

# plot variance in length at age and variance in log lenth at
    age
lenvar_plot <-
  ggplot(data = len_dat, aes(x=age, y=varlen)) +
  geom_line() + facet_wrap(~type, scales = "free")
```

9 Paper IV - Systematically testing the relative importance of age and length composition data in an integrated assessment model

Paul N. Frater[1] Gunnar Stefansson[1]

[1]University of Iceland, Dunhaga 5, 107 Reykjavik, Iceland

Abstract

The effects of composition data sample size on performance and bias of integrated assessment models has been well studied. However, past research projects in this realm focus on the performance of models using simulations of sampling scenarios from different fisheries around the world. While this approach is useful for determining composition data collection scenarios that have the best performance or minimize bias, a systematic and comprehensive approach is lacking. We test the relative importance of composition data by systematically altering sample sizes of both age and length data across three life-history types and for three fishery harvest scenarios. We found that both age and length data are important, but that adding even a small amount of age data reduced bias in both cod and capelin life-history types even when length data was relatively lacking. Model results for flatfish tended to be relatively unbiased, so increasing sample size of age and/or

length data in this life-history type changed model performance relatively little. These results are important in showing that more data is better at fitting fisheries integrated assessment models, but perhaps finding an appropriate balance between age and length data sample sizes is more important.

9.1 Introduction

Composition data (i.e. age-, length-, and conditional age-at-length data) have been shown to be important in stock assessment modeling and integrated analysis (Ono et al., 2015; He et al., 2016; Monnahan et al., 2016). Particularly, composition data are crucial for internal estimation of parameters for biological model components such as growth (Lorenzen, 2016), recruitment (Maunder and Punt, 2013; Maunder and Piner, 2015), and natural mortality (Lee et al., 2011; Sippel et al., 2017). In addition, composition data can be important for estimating selectivity of both fisheries and standardized surveys (Maunder and Punt, 2013; Crone and Valero, 2014; Maunder and Piner, 2015). Recently, a flurry of simulation studies have emerged that test the importance of composition data on various aspects of stock assessment performance or component estimation (Ono et al., 2015; He et al., 2016; Minte-Vera et al., 2017; Stewart and Monnahan, 2017, among others). Many of these studies are performed via simulation and are tested under various fishery and fishery research development scenarios (i.e. quasi-realistic scenarios of fishery-independent data collection). While this type of setup is beneficial for understanding the importance of age and length data under various research regimes. it does not necessarily afford a systematic comparison. In addition, many of these types of studies have been carried out using the integrated analysis platform Stock Synthesis 3 (SS3; Methot and Wetzel, 2013). While it is incredibly beneficial and informative to have such a broad working knowledge of integrated analysis based on a single, common platform it would also be beneficial to understand similar effects across modeling platforms as well. Stock assessment models differ in terms of performance and estimation (Ralston et al., 2011; Deroba et al., 2015), and it seems reasonable to expect that various outcomes relating to the effects of composition data might differ across model types. In the present study we carry out a systematic test of the relative importance of age and length composition data in the integrated assessment environment Gadget. The impact of age and length data sample size is tested via simulation under a range of data collection scenarios whereby sample sizes are decreased from a supposed data-rich scenario and the effect on parameter and biomass estimation bias in the model is measured.

Gadget (the Globally applicable Area Disaggregated General Ecosystem Toolbox) is an integrated analysis framework for fitting single species stock assessment or multi-species community models (Begley, 2012). It is widely used throughout Europe and is gaining in popularity elsewhere as well (Taylor and Stefansson, 2004; Plagányi, 2007; Taylor et al., 2007). However, it has been recommended that if Gadget is to be used more widely it should undergo rigorous testing similar to the suite of simulation studies undertaken with SS3 (Elvarsson et al., 2018). This study represents the first of what will hopefully be a long line of simulation studies testing various aspects of Gadget and its component data on model performance.

Gadget is an age- and size-structured statistical model that can incorporate a wide variety of data types such as landings, relative abundance indices, age, length, and conditional age-at-length data as well as tagging data (Taylor et al., 2007). While it is very similar to other stock assessment models certain differences exist (Elvarsson et al., 2018) and further testing should be done within Gadget to see if differences exist in model performance under different scenarios of data inclusion or model mis-specification. At its core Gadget is an age structured model that implements size-based processes on the population(s). Growth and mortality are both directly related to age, but processes such as selectivity, maturity, and predation are length-based. In this study we use Gadget as both operating model (OM) and estimation model (EM) to test the effects various sample sizes of both age and length data on model performance under time-varying recruitment.

Numerous aspects of stock assessment or integrated analysis have been tested with regards to the effects of composition data. Tests involving some aspect of composition data can be generally placed into one or more of four categories: 1) importance of data quantity and/or quality, 2) data conflict/identifying appropriate weighting schemes, 3) model mis-specification, and 4) model structure. Studies under the first category generally focus on model performance under different data scenarios and assess the importance of composition data by looking at how model performance improves (or degrades) with more/better (or less/poorer) data (e.g. data-rich vs. data-poor or some other sampling scenario Wetzel and Punt, 2011; Ono et al., 2015; He et al., 2016). Those that assess data conflict or attempt to identify appropriate weighting schemes usually focus on how composition data conflicts with abundance data and how to properly weight compositional data to avoid such conflict (Francis, 2011; Crone and Valero, 2014; Minte-Vera et al., 2017; Maunder et al., 2017; Punt, 2017; Stewart and Monnahan, 2017; Wang and Maunder, 2017). Designs for studies aimed at understanding the effects of model mis-specification on model performance are seemingly almost limitless. The mis-specification itself is usually related to either structural differences in models (e.g. assessing an age-size structured model with only an age-structured model Punt et al., 2017) or by misspecifying certain processes within a model (e.g. replacing asymptotic with dome-shaped selectivity Crone and Valero, 2014; Ichinokawa et al., 2014; Wang and Maunder, 2017). Composition data are often incorporated into these studies under different categorical scenarios of data quantity or quality. Lastly, studies undertaking the assessment of model structure look at model performance under different structural setups and also often include scenarios of composition data quantity or quality (Monnahan et al., 2016; Kuriyama et al., 2016; Szuwalski, 2016).

This study can be placed under the first category – the importance of composition data. Here we test the relative importance of age and length data by altering length data sample sizes and the proportion of those data that are also sampled for age data.

Previous tests studying the effects of different scenarios of age data have tended to focus on small sample sizes along scenarios of common sampling schemes (Wetzel and Punt, 2011; Ono et al., 2015). In this study we were interested in testing how the sample sizes of age and length data contribute to bias in the parameter and biomass estimation of models. This was achieved by sampling operating models under the same sampling scheme but altering the sample sizes of composition data that were sampled. In many ways this could be thought of as strictly a data-rich scenario as we look at consistent sampling over the final 30 years of the OM; however, we believe this sort of study still offers important insight into the importance of composition data. The OMs that we use to systematically test the importance of age and length data are based on three different life-history types (cod-like, flatfish-like, and the small, pelagic baitfish capelin) as well as three different fishery development scenarios. Life history types were based on life-history parameters from: Atlantic cod (*Gadus morhua*), European plaice (*Pleuronectes platessa*), and capelin (*Mallotus villosus*). We will hereafter refer to these different life history types as: cod, flatfish, and capelin.

While many other simulation studies present different scenarios of sampling history or frequency (for example, Ono et al., 2015) we opted not to do this as it can make interpretation of results seem less systematic. The objectives for this study were rather to study the effect of age and length data sample size in a systematic way by maintaining a consistent sampling scheme for all scenarios and only adjusting age and length data sample size. This type of setup allows for a direct comparison of different sample sizes of age and length data to one another.

9.2 Methods

To test the effects of composition data sample size we used a simulation-estimation approach using Gadget as both an OM and EM. Using different models for OM and EM makes model fitting notoriously difficult so we used Gadget as both OM and EM to remove any potential bias induced by structural differences in models, which makes interpretation of the effects being tested more straightforward. For both OM and EM we used Gadget 2.2.00 (Begley, 2012). Additionally, we developed an R package, *gadgetSim* (Frater, 2018a), that easily connects R and Gadget for OM and EM setup as well as reading results using the R statistical software environment (R Core Team, 2017).

Operating Models

Operating models were based on the three life-history types described above. Parameters for these models were loosely based on data and assessments for the respective species from the Icelandic Marine and Freshwater Research Institute (Hafrannsóknastofnun, 2018). Parameter values needed that were not available were supplemented either from the literature or using FishBase (Froese and Pauly, 2017b) in cases where the literature was not informative (Table 9.8).

Gadget is an age- and size-structured model where fish in each cohort are reduced according to fishing and natural mortality.

$$N_{a,y} = \begin{cases} N_{a-1,y-1}e^{-F_{y-1}-M_a}, & a < a_{max} \\ N_{a-1,y-1}e^{-F_{y-1}-M_a} + N_{a,y-1}e^{-F_{y-1}-M_a}, & a = a_{max} \end{cases}$$

where N_{ay} are numbers at age a in year y, M_a is the age-specific natural mortality, and F is asymptotic length selective fishing mortality.

Spawning in the OMs for each year was based on the Beverton-Holt spawner-recruit function (Beverton and Holt, 1957) but included multiplicative error as follows to create time-varying recruitment:

$$R_y = \frac{\mu S_y}{\lambda + S_y} \cdot e^{\mathrm{N}(0,\sigma_R)}$$

where R_y is the mean recruitment in year y, S_y is SSB in year y, μ and λ are species specific parameters controlling the steepness and maximum level of R, and σ_R is a species specific error term used to introduce time-varying recruitment. The value of σ_R for each species was chosen to induce enough process error to be clearly evident, but not so much that it masked the trajectory of each respective fishery development scenario ($\sigma_R = 0.1$ for cod $\sigma_R = 0.15$ for flatfish; and $\sigma_R = 0.2$ for capelin). The resulting values were used to fix a recruitment value for each year within the OMs.

Growth was implemented as a variant of the von Bertalanffy growth curve:

$$\Delta L = (L_\infty - L_i)(1 - e^{-k\Delta t})$$

where L_∞ is the mean length L as time t approaches infinity, and k is the usual Brody growth coefficient. Variation in mean size-at-age was implemented using a beta-binomial distribution, and weight was based on the length-weight relationship: $W = \alpha L^\beta$. Initial numbers of fish for each age class at year 1 was given as $N_{a,0} = 1\mathrm{e}08 e^{-M\times a}$. For further details on how populations are structured in Gadget, see Begley (2012).

For each life history type and fishery development scenario we simulated a model that ran for 120 years with the first 20 years serving as a burn-in period to allow OMs and EMs to stabilize. For ease of interpretation the burn-in period was discarded and hereafter each year is referred to as the respective year minus 20 (i.e. the final year is 100 and the beginning year at the end of the burn-in period is 0.

The fishery development scenarios were the commonly used fishing trajectories two-way-trip, fish down, and flat MSY. These trajectories were created specific to each life-history type by calculating the fishing mortality at maximum sustainable yield (F_{MSY}) and setting up the development scenarios in relation to that. We calculated F_{MSY} by computing yield curves across values of F and defined F_{MSY} as the value for F that resulted in the maximum yield. Other reference points used to determine fishery development scenarios will be referred to as F_{high} and F_{low} which were defined as the values for F that resulted in yield being 0.8 times the yield at F_{MSY}, where $F_{low} < F_{MSY} < F_{high}$ (Table 9.9). Fishing for all fishery development scenarios began in year 20. For the fish down scenario, the fishery began with an F of 0 and increased linearly to a high of F_{high} in year 100. The two-way-trip scenario began similarly, but reached an F of F_{high} in year 68 and then decreased to F_{MSY} in year 100. The flat MSY scenario began fishing in year 20 at F_{MSY} and remained there for the duration of the model.

Sampling the OM

We used output from the OMs as “data" on which to minimize the objective function in the EMs. OMs were sampled with asymptotic selectivity similar to the fishery selectivity, but with an inflection point scaled lower in order to sample a greater number of smaller-sized fish. Starting in year 70 we sampled OM output biannually by applying asymptotic selectivity on absolute numbers at year, timestep, age, and length combinations up to a maximum of 0.001% and with multiplicative error as follows:

$$I_{ytal} = S_l N_{ytal} e^{\mathrm{N}(\mu,\sigma)}, \qquad \mu = 0, \sigma = 0.2$$

where I_{ytal} is the index of abundance at each year (y), timestep (t), age (a), and length

Table 9.8. Parameters used in the operating model for each life-history type. Dots for the Length$_{SD_{init}}$ parameter indicates a repitition of the final value until maximum age of the stock. Parameters for the growth, maturity, spawning, and fleet selectivity parameter types refer to the Gadget functions "lengthvbsimple", "exponential" length selection in the spawning file, Beverton-Holt spawner-recruit process, and "exponentiall50" suitability curve, respectively.

Parameter	Parameter Type	Cod	Capelin	Flatfish
Min. Age	Stock	1	1	1
Max. Age	Stock	20	5	30
Min. Length	Stock	1	1	1
Max. Length	Stock	150	25	100
Length$_{SD_{init}}$	Stock	2, 4, 6, 8, 10,...	1, 1.5, 2,...	2, 4, 6,...
M	Mortality	0.2	1	0.1
L_∞	Growth	125	20	55
k	Growth	0.15	0.6	0.14
t_0	Growth	-0.35	0	-1.3
α_{LW}	L-W Rel.	6.7e-06	7.1e-06	9.6e-06
β_{LW}	L-W Rel.	3.1	3.15	3
α_{mat}	Maturity	-0.2	-2	-0.25
$L50_{mat}$	Maturity	64	12.5	30
λ_{BH}	Spawning	1.07e+08	5e+07	1.07e+08
μ_{BH}	Spawning	3e+08	8e+10	2e+08
α_{fleet}	Fleet selectivity	0.1	1	0.2
$L50_{fleet}$	Fleet selectivity	45	7	30

(l), S_l is the selectivity at length up to a maximum of 1% of total numbers N.

We used these values as abundance data in our EMs. Composition data was sampled as a subset of the abundance index data in sample sizes of 1000, 500, 200, 100, 50, 25, and 10. Both age and length data were sampled at these levels, but age data sample size were never greater than length data sample size for a given scenario. This was to reflect common sampling scenarios where age data is usually sampled from fish that have been measured for length. Length samples were randomly sampled from the samples selected for abundance indices and age samples were randomly selected from the length samples. Additionally, a scenario was included with no age data samples (Length Only). This resulted in a total of 35 unique age and length data scenarios.

Catches from fleets were recorded without error, and output from OM commercial fleets

Table 9.9. Fishing mortality (F) values used to determine fishery development scenarios. Both F_{low} and F_{high} are the F values that produce 80% yield compared to F_{MSY}.

Species	F_{MSY}	F_{low}	F_{high}
Cod	0.22	0.11	0.45
Flatfish	0.12	0.06	0.22
Capelin	1.08	0.53	1.82

were sampled similarly to survey sampling for composition data, but a size-selective curve was not applied as the fishery already implements length-based selectivity. Rather 0.001% of fishery output were sampled and used as length data regardless of size. Sampling these values to use as age data in the EMs was based on the appropriate scenario in accordance with the fishery-independent sampling described above. Survey and fishery samples were performed 100 times with random multiplicative error for both survey and fishery samples (therefore the length and age value inherit this sampling error – separate aging error was not placed on age values). This process produced 100 unique data sets for each life-history, fishery development, and sampling scenario, which resulted in a total of 31,500 estimation models.

Estimation Models

EMs were set up for each independent data set under each scenario and fit using the iterative re-weighting procedure described by Stefansson (2003) and implemented by Taylor et al. (2007) and Elvarsson (2015). This procedure consists of performing a separate optimization run for each likelihood component (e.g. indices of abundance, each composition data set, etc.) using a heavy weight on that component. Variance for that component optimization is estimated as the overall likelihood score (SS_c) divided by degrees of freedom (df_c) $\hat{\sigma}_c^2 = \frac{SS_c}{df_c^*}$ where the number of non-zero data points (df_c^*) serves as a proxy for df_c (Elvarsson, 2015). After all component-weighted optimization runs have finished a final optimization run is performed using the inverse variance from above as the weight on that respective component ($w_c = 1/\hat{\sigma}_c^2$). We used the R package

Rgadget (Elvarsson and Lentin, 2018) to automate this procedure. The likelihood that was used for composition data was the sum-of-squares function in Gadget. Abundance indices were fitted using a log-linear regression of estimated to observed values with a slope fixed at 1 and intercept estimated internally.

All model parameters except for M_a were estimated. These include parameters for growth (simplified von Bertalanffy function), growth implementation (beta-binomial), initial starting values, Beverton-Holt spawner-recruit parameters, spawning maturity, and selectivity parameters for surveys and fleets. It should be noted that spawning maturity was implemented as the proportion function in the Gadget spawning file and not as a separate stock that matures into a spawning stock.

For each data scenario, life-history type, and fishery development scenario we optimized 100 separate models each consisting of a different randomized data set. All models were assessed for convergence by ensuring that the convergence criteria in the optimization procedure was met.

Performance Assessment

We assessed the performance of EMs under each scenario using relative error (*RE*) and median absolute relative error (*MARE*) of estimated parameters, spawning stock biomass in the terminal year (SSB_{100}), and depletion (DEP = SSB_0/ SSB_{100}).

$$RE = (\hat{\theta} - \theta)/\theta$$

$$MARE = \text{median}(|\hat{\theta} - \theta)|/\theta)$$

RE values for each model are used to assess bias whereas *MARE* accounts for bias and precision simultaneously.

9.3 Results

Estimation of Depletion and SSB_{100}

Greater sample sizes of composition data showed a clear trend towards reduced bias as sample sizes of both age and length data increased. This was particularly evident for SSB_{100} in both capelin and cod (Fig. 9.35). Terminal biomass exhibited the greatest median bias and overall variance in bias when age data was absent and length data had the lowest sample sizes. However, when even slight increases were made to age data sample sizes this drastically improved the fit of the EM to the OM. Scenarios with age and length data of 25 or greater seemed to have a median bias fairly similar to even the most data rich scenarios. This same pattern was also apparent for depletion, but continued increases in composition data sample size seemed to continuously decrease bias in depletion at least for cod and capelin (Fig. 9.36).

The importance of composition data also seemed to rely somewhat on fishery development scenario as well. Median bias was highest in the fish down scenario for both cod and capelin, and having greater sample sizes of both age and length data decreased bias more noticeably in these scenarios. Median bias and variance in bias tended to be less, and increase to a lesser degree with low composition data sample sizes with the flat MSY and two-way-trip fishery development scenarios.

None of these results seemed to hold true for flatfish, which seemed to perform fairly similarly across composition data scenarios for both depletion and SSB_{100} . Bias in both of these metrics seemed to be about the same across all scenarios of age and length data sample size and no detectable trends appeared across fishery development scenarios.

Estimation of Parameters

Most estimated parameters did not show much variation across the different scenarios.

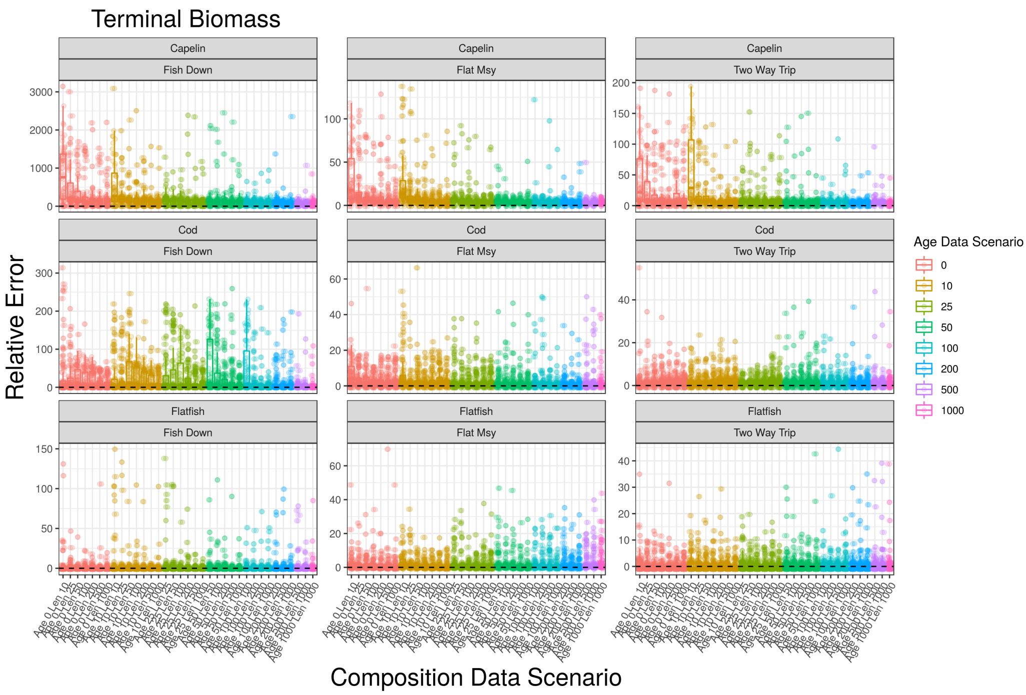


Figure 9.35. RE values for SSB_{100} across composition data sample size scenarios and for various life-history and fishery development scenarios. Composition data scenarios are listed along the x-axis, and each level of age sample size data are plotted as different colors.

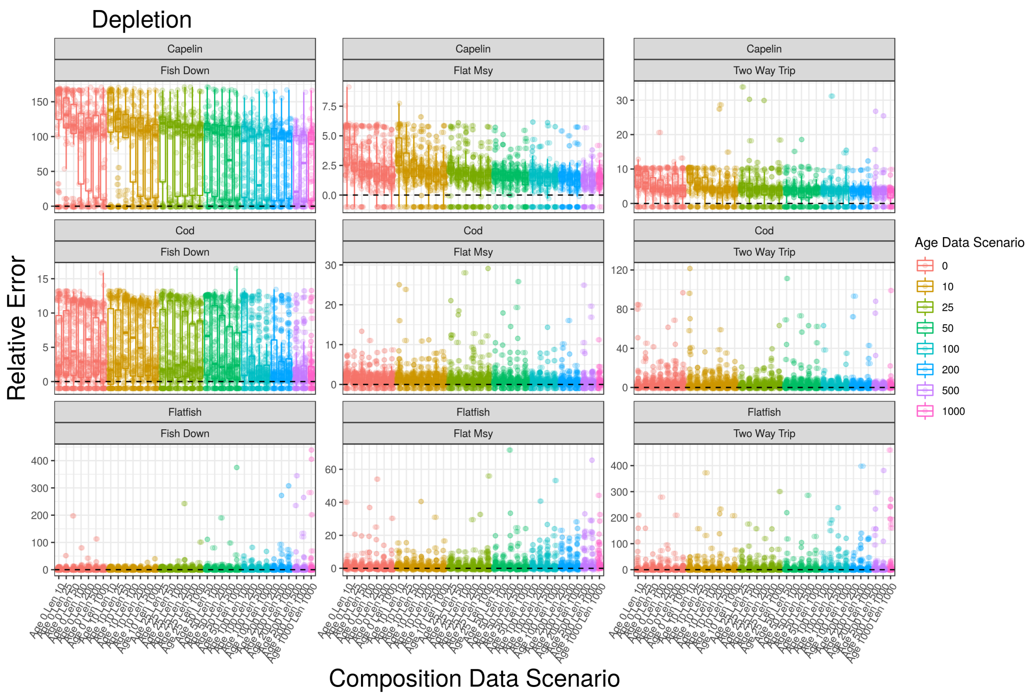


Figure 9.36. RE values for depletion among different composition data scenarios. Axes, colors and legend are all the same as Fig. 9.35

.

This did not hold true for estimated recruitment parameters, though. μ_{BH} showed considerable variation as age and length data sample sizes decreased (Fig. 9.37. This is likely also the cause for increased depletion and SSB_{100} for those scenarios as well. This pattern was most evident in all fishery development scenarios for capelin and in the flat MSY and, to some extent, the fish down scenarios for cod. The two-way-trip scenarios for cod, and all fishery development scenarios for flatfish actually showed increased bias in μ_{BH} estimates with increasing composition data sample size. Estimates of λ_{BH} did not exhibit increased bias at lower composition data sample sizes (not shown), but relative error in the ratio of $\frac{\mu_{BH}}{\lambda_{BH}}$ did decrease as composition data sample sizes increased (Fig. 9.38.)

Bias in estimates of growth parameters remained relatively constant across different sample sizes of age and length data. In addition, maturity parameters did not seem to drastically change among the various scenarios. Neither of these parameter types are shown here.

9.4 Discussion

We found that increasing composition data sample sizes decreases bias in biomass and parameter estimation in an integrated assessment model. This was as expected and is consistent with other studies that have tested the effects of composition data (Chen et al., 2003; Wetzel and Punt, 2011; Ono et al., 2015). Unlike many other studies that have tested the effects of composition data, though, we did not have fishery specific type sampling schemes, but rather opted for a consistent sampling strategy (i.e. biannual surveys for the last 30 years of model) while only altering the sample size of age and length data. This revealed some interesting and unexpected results. Firstly, age and

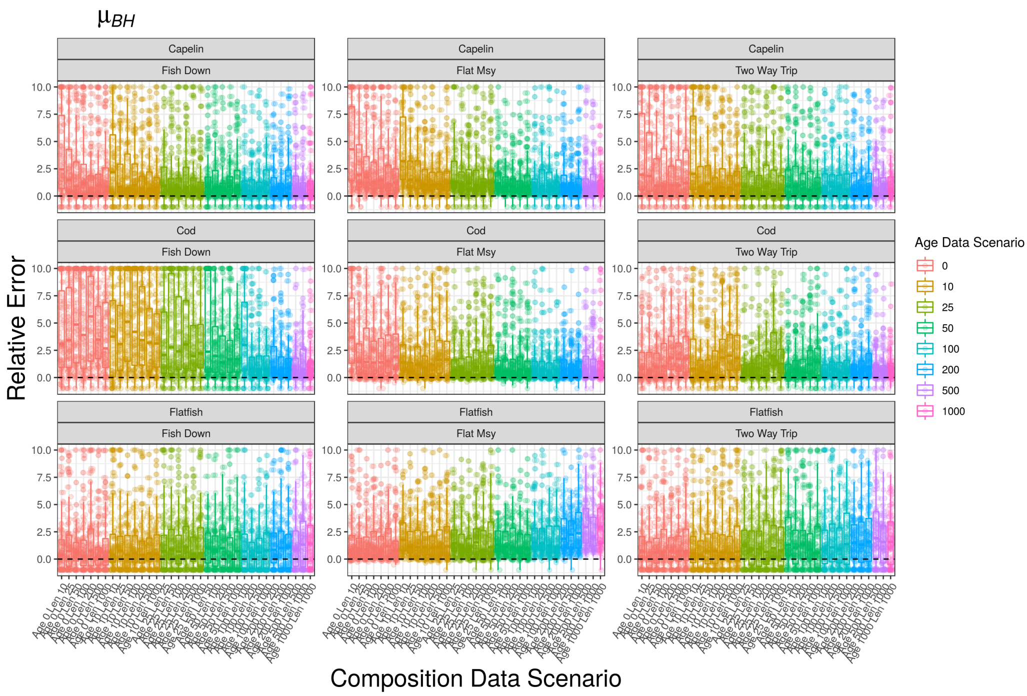


Figure 9.37. RE values for μ_{BH} among different composition data scenarios. Axes, colors and legend are the same as Fig. 9.35

.

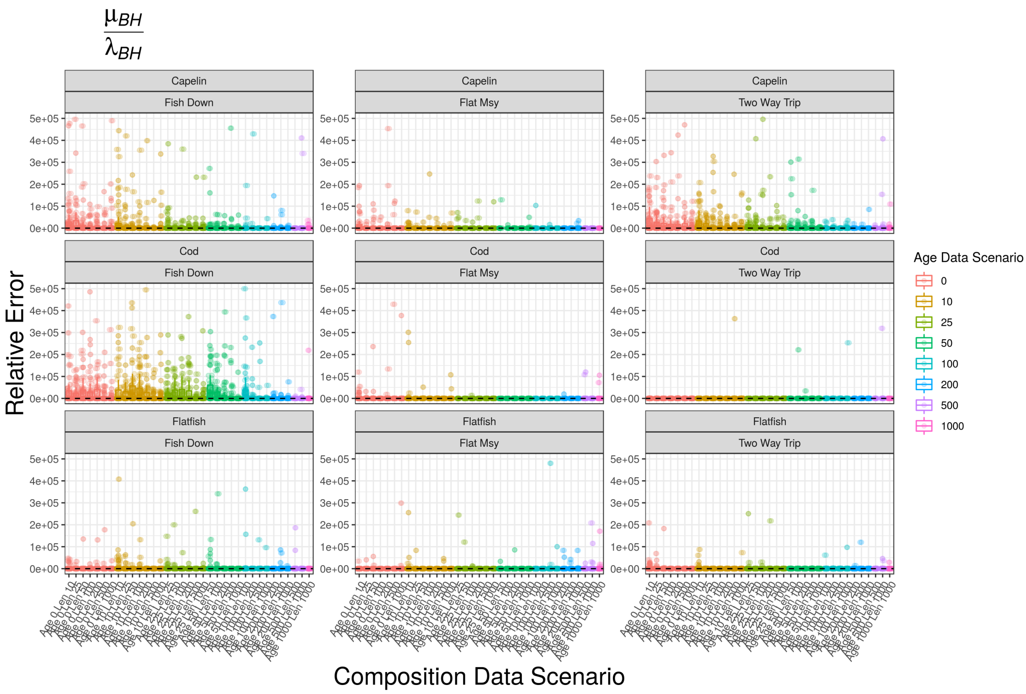


Figure 9.38. RE values for $\frac{\mu_{BH}}{\lambda_{BH}}$ among different composition data scenarios. Axes, colors and legend are the same as Fig. 9.35

length data seemed to have a greater impact on reducing bias in biomass estimation (i.e. depletion and SSB_{100}) for capelin the most, an intermediate impact on cod, and ostensibly no impact on flatfish whatsoever. These results are contrary to other studies that have assessed the importance of age and length data. Ono et al. (2015) found that age data improved stock assessment accuracy for all life-history types (which were similar to those used in this study) and did so particularly for a flatfish life-history type, but they noted that increasing the sampling frequency and sample size did not greatly affect the performance of estimation models. Similar results have also been produced by Wetzel and Punt (2011) and Chen et al. (2003) which show that inclusion of at least some fishery-independent age data is enough to greatly improve stock assessment performance. He et al. (2016) also performed a study simulation that is similar to the present one, but which did not include variable recruitment, and found that greatly reducing the amount of age data had minimal influence on accuracy of estimation models for some life-history types. In the above cases the authors noted that reducing age data tends to not affect estimation model results for life-history types that exhibit strong year-class signals in the length-frequency data (i.e. fast growing life-history types with high recruitment variability; Ralston and Ianelli, 1998). For long-lived, slow growing species age data tends to be important for estimation accuracy. Our results are similar to the above studies in showing that the inclusion of fishery-independent age data decreases bias in estimates of both parameters and biomass, but our results differ from many of the previously concurred results with regard to the differences in the importance of composition data among life history types. Other studies that have tested the importance of composition data argue that it is more important for life history types that are slow growing with much overlap among lengths between age classes (i.e. ages are easier to distinguish using length data alone for fast-growing species with low variance in length-at-age Ono et al., 2015). In this study we actually found the opposite. Composition data seemed to be most important for capelin, of intermediate importance for cod, and of almost no importance whatsoever for flatfish. These results are contrary to previous findings that explain the importance of length and, especially, age data via its ability to distinguish age classes when length-at-age is highly overlapping. There could be several reasons for the differences in trends among life histories with regards to composition data. Firstly, prior studies have concluded that the importance of age data were attributable to the ability of such to distinguish age classes and that it is less important for life history types that show little overlap in lengths among age classes. This is certainly a reasonable and tenable explanation, but the life-history types in this study had similar characteristics to other studies (especially Ono et al., 2015). Former studies have relied on more realistic sampling schemes, whereas here we used a single sampling scheme and altered the sample size of composition data. Given this systematic test of the importance of composition data and the accompanying results it is therefore plausible that the trend among life-histories is related to an aspect other than overlap in length data among age classes. Here we identify three potential mechanisms that could drive the trend found in the current study. The first is related to parameters driving the recruitment in the OMs. Although capelin had a much higher μ_{BH} than the other life histories, and both cod and flatfish OMs were set up with a μ_{BH} equal to each other, the ratio of μ_{BH}:λ_{BH} was high for capelin, much lower for cod, and lowest for flatfish. While μ_{BH} in the Beverton-Holt stock recruit function drives the overall number of recruits for the most part (Beverton and Holt, 1957), the ratio between the two alters the steepness of the curve. A high ratio will result in a steep stock-recruit curve, which will leave little contrast among different year classes and make overall determination of SSB and number of recruits more difficult. Having a greater sample size of composition data would certainly aid with this. Alternatively, a shallower stock-recruitment curve would yield good contrast among number of recruits, which would likely make determination of year classes more distinguishable thereby reducing the need for higher composition data sample sizes. The reliability of estimates in stock assessment has been shown to depend on contrast in the stock recruit curve (Magnusson and Hilborn, 2007). This trend of increasing importance of composition data with increasing steepness is also somewhat followed by Ono et al. (2015) as flatfish in the OMs presented in that paper had the greatest steepness parameter (h) for the stock-recruit function.

An alternative to importance of the stock recruit curve is natural mortality M, although this seems much less plausible. In the present study, capelin had the highest mortality ($M = 1$), followed by cod ($M = 0.2$), and flatfish ($M = 0.1$). The importance of composition data increasing with M seems implausible due to a similar, but opposite, reason as the effect of $\frac{\mu_{BH}}{\lambda_{BH}}$ on the importance of composition data. That is that increasing M would cause a greater contrast among age groups, whereas a low value for M would create a more homogenous age-structured population leaving little contrast among age groups. To our knowledge no studies have been done which test the effect of different values of M on the importance of composition data. However, it seems likely that other factors such as contrast in biomass via exploitation or through stock recruitment would have a greater effect than M.

Lastly, the differences seen between this study and others could simply be a result of using different modeling platforms for the OMs and EMs. Most studies of this nature have been performed in Stock Synthesis 3 (Methot and Wetzel, 2013; Elvarsson et al., 2018), and while SS3 is similar to Gadget differences do exist (Elvarsson et al., 2018). More generally, different integrated and stock assessment models tend to model stocks with different results (Ralston et al., 2011). As mentioned in the introduction this is the first true simulation study for Gadget, and the trends seen here can only be truly verified with more simulation testing.

The present study is different than that of others which tend to set up scenarios based on more realistic sampling scenarios that represent common collection frequency schemes (i.e. scenarios with annual vs. intermittent sampling, scenarios that increase sample size throughout time, etc. Chen et al., 2003; Wetzel and Punt, 2011; Ono et al., 2015; He et al., 2016). We opted to test the effects of age data sampling size in a more systematic manner by having the same sampling frequency (i.e. beginning in year 70 and sampling twice annually in every year thereafter) across all scenarios while only manipulating sampling sizes. The sampling scheme used in this simulation study is based on fishery-independent surveys performed in Iceland (Hafrannsóknastofnun, 2018), and while this is probably a relatively data-rich sampling scheme in terms of frequency of observations it allows for a clean, systematic test of the performance of different sample sizes. Even though this makes direct comparisons to other studies (i.e. Wetzel and Punt, 2011; Ono et al., 2015) more difficult, it is reassuring that, using this systematic approach, we arrived at many of the results that are consistent with other such studies testing similar aspects of stock assessment (i.e. that increasing composition data beyond a certain point did not drastically improve model performance or fit).

This study carries the same drawback as many of the studies that assess various aspects of assessment models, and that is using the same model for both OM and EM. It is well known that fitting an EM to output from an OM that is a different modeling platform is rife with problems. The difficulty of estimating a model that implements different assumptions is often impractical for the investigation and inspection of certain aspects of assessment models. This drawback is attached to most simulation studies as most are performed using the same model for both OM and EM due to the aforementioned reason. However, readers should not take this lightly as a good deal of knowledge about stock assessment models are based on such studies. Another limitation of this study is the relative paucity of testing performed with our stock assessment modeling software, Gadget. Although Gadget has been fully developed for a number of years, there have not been a great deal of simulation studies testing aspects of the model as there has been with other modeling platforms such as ADMB (Fournier et al., 2012) or Stock Synthesis (Methot and Wetzel, 2013). For example, these software platforms use automatic differentiation to minimize an objective function and are able to estimate uncertainty using MCMC methods, while for this study we used basic nonlinear minimization algorithms of Simulated Annealing and Hooke and Jeeves. However, recently there has been a call for more testing of Gadget (Elvarsson et al., 2018) and there seems to be more synergy for research and development of the Gadget modeling software. Hopefully in the future Gadget will continue to be tested and further developed to keep pace with other technologies and new advances in stock assessment science.

In future studies it would be ideal to implement the sort of systematic framework used here to test stock assessment models in an effort to achieve generalizations about various aspects of and data used by them (e.g. such as the generalization garnered in other studies of the relative triviality of age data used for life-history types exhibiting strong year-class signals). While looking at specific case studies and scenarios offers great insight into a particular stock assessment, species, or sampling regime, generalizations that lead to a broader understanding of stock assessment are difficult to interpret with such studies. For studies that are interested in gaining understanding about broader, more general impacts of various aspects of stock assessment or data that go into such models we recommend that simulation studies be set up in a systematic manner such as this one. Indeed, researchers could also go far beyond the type of setup here and look at, for example, how a single particular aspect of life history might affect stock assessment performance across some facet of data or model misspecification. For example, one could adjust growth rate and variance in length-at-age in the OMs and use different samples sizes of age data to determine how these life-history characteristics impact the effect of age data improving stock assessment performance. Given the complexity of these models (Cotter et al., 2004; Bolker et al., 2013) it would be useful to have some general guidelines as to the different aspects of data or model misspecification that might interact with the various processes in stock assessment models (e.g. such as rapid growth leading to strong year-class signals). Similarly, it would be ideal to establish some general and standardized life-history types that could be used systematically for such studies. Cod, flatfish, and some sort of small pelagic species (i.e. sardine, anchovie, etc.) seem to be fairly common. However, it would also be good to add truly slow-growing, long-lived species such as certain species of redfish (*Sebastes* spp.).

Acknowledgements

Funding for the lead author to undertake his Ph.D. studies was received from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 613571 for the project MareFrame and from the European Commission's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 634495 for the project Science, Technology, and Society Initiative to minimize Unwanted Catches in European Fisheries (Minouw) as well as from the Icelandic Research Fund (Rannís, No. 152039051). PNF is grateful to GS for acquiring said funding.

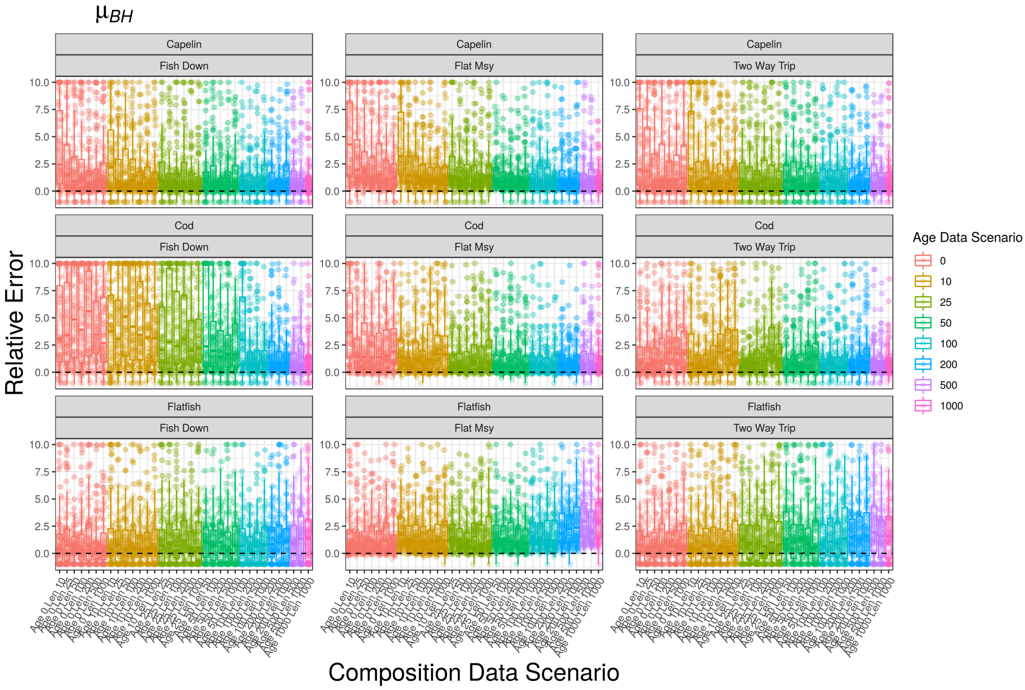
Table 9.10. MARE values for SSB_{100} for all age data, life-history type, and fishery development scenarios.

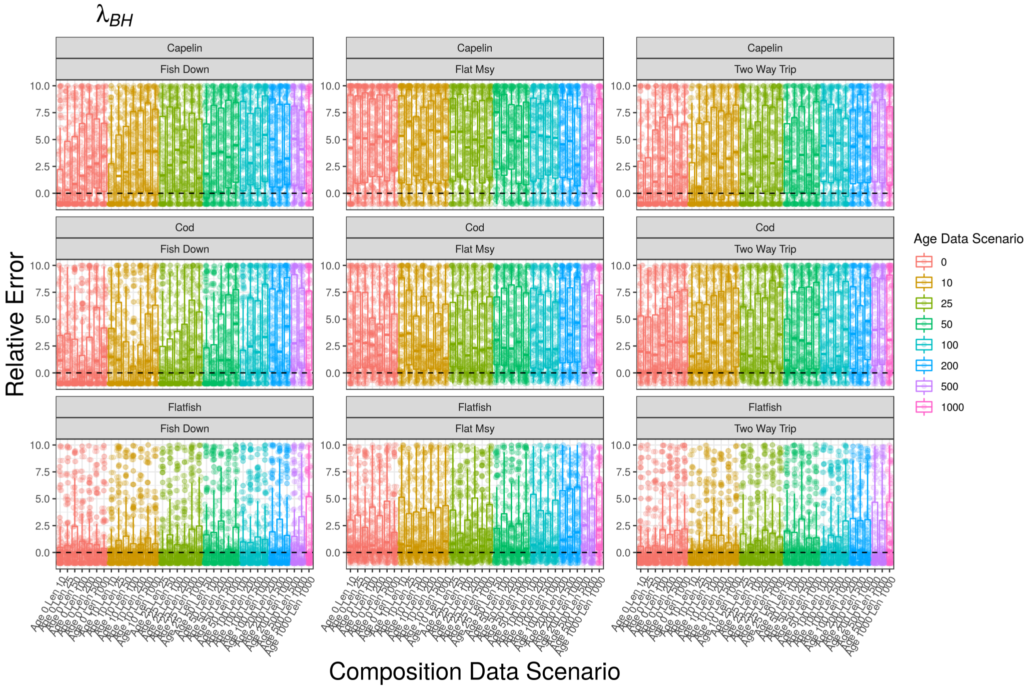
	Capelin			Cod			Flatfish		
Age Data Scenario	Fish Down	Flat MSY	Two Way Trip	Fish Down	Flat MSY	Two Way Trip	Fish Down	Flat MSY	Two Way Trip
Age 0 Len 10	1003.20	18.24	27.92	3.58	5.14	0.99	0.99	0.72	1.37
Age 0 Len 25	225.87	10.08	14.23	7.27	1.64	1.20	1.00	0.57	1.00
Age 0 Len 50	175.66	6.49	8.46	10.24	1.38	1.27	0.90	0.56	0.98
Age 0 Len 100	117.85	5.45	9.03	6.16	1.47	1.60	0.96	0.58	0.80
Age 0 Len 200	145.77	4.99	7.12	13.69	1.14	1.75	1.00	0.50	0.86
Age 0 Len 500	125.27	4.62	7.90	5.70	1.44	1.90	0.99	0.64	0.99
Age 0 Len 1000	122.79	3.95	7.80	5.12	1.71	2.05	0.99	0.57	1.00
Age 10 Len 10	422.99	13.80	38.68	4.41	2.37	1.16	1.00	1.29	1.29
Age 10 Len 25	186.10	6.16	8.84	5.48	0.89	1.25	1.00	0.60	1.00
Age 10 Len 50	161.24	6.72	9.57	11.16	1.16	1.22	1.00	0.60	1.00
Age 10 Len 100	149.51	6.48	7.20	14.29	1.06	1.29	1.00	0.66	1.00
Age 10 Len 200	158.13	4.63	6.76	16.19	1.04	2.16	1.00	0.51	0.95
Age 10 Len 500	130.90	4.53	6.25	3.69	1.87	1.64	0.96	0.58	0.94
Age 10 Len 1000	112.36	4.63	6.71	4.00	1.35	1.84	0.84	0.54	1.00
Age 25 Len 25	175.93	6.27	7.70	5.41	0.98	1.14	1.00	0.76	1.20
Age 25 Len 50	153.43	4.67	7.03	5.89	0.83	1.21	1.00	0.39	1.00
Age 25 Len 100	161.21	5.01	7.20	4.89	0.55	1.30	1.00	0.48	1.00
Age 25 Len 200	125.82	4.01	7.90	7.66	0.91	1.30	1.00	0.62	0.93
Age 25 Len 500	117.16	4.16	6.36	3.30	0.67	2.03	0.87	0.59	1.00
Age 25 Len 1000	121.78	4.16	7.16	4.54	1.17	1.57	0.90	0.65	1.00
Age 50 Len 50	150.76	4.93	6.26	13.25	0.75	1.18	0.97	0.61	1.00
Age 50 Len 100	123.96	4.23	6.99	5.67	0.82	1.40	1.00	0.37	1.00
Age 50 Len 200	129.22	3.32	6.37	4.92	0.62	1.21	1.00	0.47	1.00
Age 50 Len 500	91.90	4.06	5.31	3.68	0.68	1.49	1.00	0.46	0.84
Age 50 Len 1000	106.03	3.77	7.09	3.52	0.59	1.39	0.90	0.49	1.00
Age 100 Len 100	110.63	3.91	6.75	22.21	0.89	1.05	0.98	0.57	0.99
Age 100 Len 200	112.33	3.78	6.37	2.32	0.62	1.17	1.00	0.41	1.00
Age 100 Len 500	99.42	3.73	5.83	1.56	0.40	1.00	0.96	0.46	1.00
Age 100 Len 1000	102.37	4.10	7.32	1.42	0.54	1.20	1.00	0.45	1.00
Age 200 Len 200	105.30	4.04	5.44	1.89	0.71	1.21	1.00	0.72	1.00
Age 200 Len 500	103.66	3.61	6.06	1.39	0.47	1.02	1.00	0.71	1.00
Age 200 Len 1000	91.85	2.95	5.76	1.75	0.53	1.33	1.00	0.47	0.99
Age 500 Len 500	49.77	3.32	5.79	1.51	0.61	1.03	1.00	0.82	1.00
Age 500 Len 1000	76.37	3.58	5.88	1.21	0.57	1.00	1.00	0.88	1.00
Age 1000 Len 1000	93.27	3.47	5.44	1.46	0.49	1.00	1.00	0.83	1.00

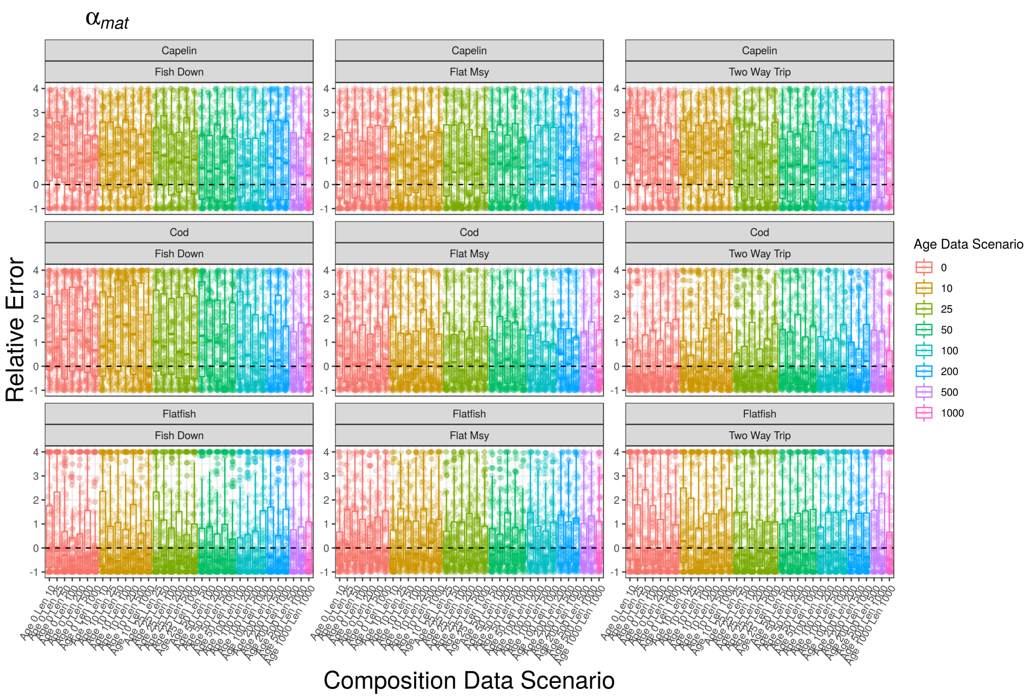
Table 9.11. MARE values for depletion for all age data, life-history type, and fishery development scenarios.

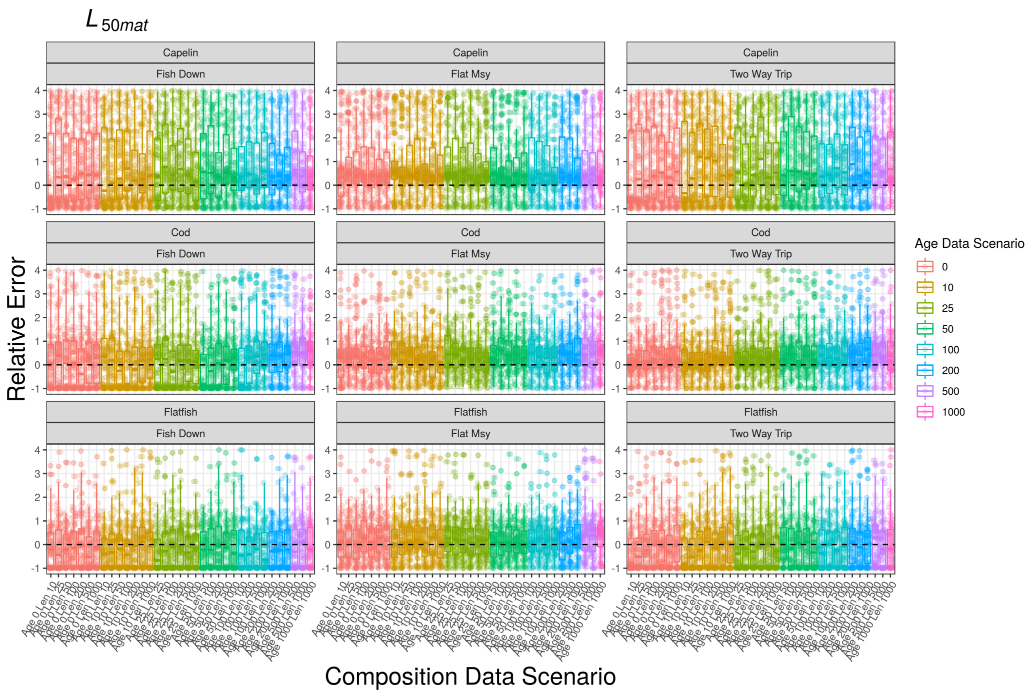
	Capelin			Cod			Flatfish		
Age Data Scenario	Fish Down	Flat MSY	Two Way Trip	Fish Down	Flat MSY	Two Way Trip	Fish Down	Flat MSY	Two Way Trip
Age 0 Len 10	162.66	4.27	8.89	2.61	2.29	1.31	1.00	0.68	1.82
Age 0 Len 25	133.41	2.83	6.91	4.87	1.27	1.33	1.00	0.56	1.23
Age 0 Len 50	120.42	2.21	4.96	8.41	1.19	1.58	1.00	0.55	1.02
Age 0 Len 100	108.35	2.00	4.52	3.77	1.28	1.73	1.00	0.65	1.00
Age 0 Len 200	114.67	1.93	4.02	8.64	1.05	3.49	1.00	0.47	1.00
Age 0 Len 500	110.59	1.83	4.49	3.87	1.18	2.99	1.00	0.66	1.00
Age 0 Len 1000	106.64	1.63	4.27	4.32	1.13	2.93	1.00	0.56	1.04
Age 10 Len 10	150.47	3.29	9.01	2.92	1.63	1.43	1.00	0.99	2.14
Age 10 Len 25	128.18	2.11	5.75	3.41	0.92	1.40	1.00	0.59	1.47
Age 10 Len 50	119.51	2.30	4.72	9.71	1.17	1.38	1.21	0.49	1.11
Age 10 Len 100	113.73	2.19	4.28	8.86	1.01	1.60	1.00	0.64	1.30
Age 10 Len 200	115.34	1.77	4.01	9.49	1.06	3.33	1.00	0.51	1.42
Age 10 Len 500	108.66	1.82	4.34	2.36	1.53	2.64	1.00	0.55	1.00
Age 10 Len 1000	106.26	1.86	3.96	2.35	1.21	2.72	0.91	0.64	1.00
Age 25 Len 25	121.43	2.23	4.86	5.14	0.91	1.33	1.00	0.77	1.81
Age 25 Len 50	116.88	1.84	4.26	3.15	0.88	1.38	1.36	0.35	2.67
Age 25 Len 100	115.39	1.89	4.27	3.44	0.67	1.73	1.00	0.42	1.70
Age 25 Len 200	109.82	1.64	4.53	4.94	1.00	1.56	1.00	0.66	2.45
Age 25 Len 500	104.50	1.73	4.01	2.34	0.75	3.27	0.99	0.60	1.26
Age 25 Len 1000	109.37	1.74	4.33	3.10	0.94	2.28	1.00	0.65	1.00
Age 50 Len 50	113.70	1.82	4.28	8.07	0.78	1.24	1.31	0.50	1.42
Age 50 Len 100	106.67	1.75	4.11	4.02	0.75	1.63	1.00	0.40	1.62
Age 50 Len 200	105.52	1.57	3.81	3.30	0.65	1.36	1.03	0.45	1.17
Age 50 Len 500	77.60	1.70	3.81	2.30	0.78	2.27	1.00	0.51	1.00
Age 50 Len 1000	103.59	1.58	4.32	3.15	0.58	1.75	1.00	0.50	1.63
Age 100 Len 100	96.63	1.73	4.23	10.03	0.88	1.35	1.00	0.60	1.11
Age 100 Len 200	102.13	1.54	4.01	1.79	0.63	1.57	1.23	0.55	2.08
Age 100 Len 500	85.09	1.62	3.86	1.33	0.43	1.18	1.00	0.46	1.88
Age 100 Len 1000	95.01	1.63	4.33	1.17	0.54	1.46	1.70	0.50	1.67
Age 200 Len 200	100.65	1.65	3.85	1.50	0.75	1.56	1.04	0.79	2.01
Age 200 Len 500	98.89	1.50	3.81	1.30	0.61	1.36	1.70	0.82	2.46
Age 200 Len 1000	97.58	1.40	3.71	1.44	0.52	1.64	1.43	0.60	1.30
Age 500 Len 500	32.12	1.30	3.90	1.33	0.73	1.30	1.32	0.96	3.65
Age 500 Len 1000	84.66	1.28	3.96	1.16	0.58	1.11	2.57	0.86	1.18
Age 1000 Len 1000	96.14	1.49	3.84	1.25	0.39	1.19	1.47	1.10	3.35

9.5 Appendix – Parameter RE values




λ_{BH}
Capelin
Fish Down
Flat Msy
Two Way Trip
Cod
Flatfish
Relative Error
Composition Data Scenario
Age Data Scenario
0
10
25
50
100
200
500
1000


α_{mat}
Capelin
Fish Down
Flat Msy
Two Way Trip
Cod
Flatfish
Relative Error
Composition Data Scenario
Age Data Scenario
0
10
25
50
100
200
500
1000


L_{50mat}
Capelin
Fish Down
Flat Msy
Two Way Trip
Cod
Flatfish
Relative Error
Composition Data Scenario
Age Data Scenario
0
10
25
50
100
200
500
1000

10 Paper V - Effects of reduced age data on uncertainty estimates in an integrated assessment model

Paul N. Frater[1] Gunnar Stefansson[1]

[1]University of Iceland, Dunhaga 5, 107 Reykjavík, Iceland

Abstract

Age data has been shown to greatly improve point estimates of parameters and estimated biomass of integrated assessment models in fisheries. The effects of age data on uncertainty estimates of parameters and model output is less well-studied, though. We perform a simulation study as well as look at an example from Atlantic cod *Gadus morhua* to determine the effect of age data on uncertainty estimates using a spatial resampling bootstrap approach. In general, age data narrows the confidence intervals associated with bootstrapped models; however, the effect is more noticeable for some life-history types and fishery development scenarios than others. Model output for cod did not see this same trend, though, as confidence intervals for model output either remained the same across age data proportions or actually increased.

10.1 Introduction

Age and length data (also known as composition data) are crucial for fitting accurate and dependable integrated assessment models in fisheries (Wetzel and Punt, 2011; Ono et al., 2015; He et al., 2016). They have been shown to be important for determining growth (Francis, 2016; Lorenzen, 2016), estimating recruitment parameters (Magnusson and Hilborn, 2007; Ono et al., 2015), and essential for any hope of estimating natural mortality (M) in fish (Lee et al., 2011; Sippel et al., 2017). Much attention has been appropriately devoted to the importance of composition data for accurately and precisely determining model parameters as well as estimated stock biomass. While it is mostly unambiguous that composition data aid in the fitting of stock assessments, it is also generally agreed upon that an overabundance of composition data is not needed to improve model fit (Wetzel and Punt, 2011; Ono et al., 2015; He et al., 2016). Many studies that assess the effects of composition data show that at least some fishery-independent age and length data are needed to reduce bias in parameter estimates and accurately fit stock biomass, but beyond a certain point more data do not necessarily decrease bias. This paints a rosy picture for stock assessments especially for stocks that have some composition data but are otherwise relatively data-poor, but the effect of composition data on parameter and biomass uncertainty estimates has not been rigorously tested. While the relationship between more composition data and better model fit is thought to be asymptotical, the effects of composition data sample size on uncertainty estimates has not been thoroughly explored.

Several different methods are used to estimate uncertainty (e.g. 95% confidence or credible intervals) in integrated assessment models (Magnusson, 2016). These methods are generally Markov-chain monte carlo (MCMC), the Delta method, and the bootstrap. While each of these methods has its strengths and weaknesses all of them rely on available data to estimate uncertainty intervals. MCMC explores an area around the minimum of the objective function to determine credible intervals (Gelman et al., 2014), and the Delta method uses the inverse-Hessian matrix of partial derivatives from the objective function to determine confidence intervals for parameters (Fournier et al., 2012). There are several bootstrap methods all of which utilize a resampling of the data a number of times and re-calculating the model to gain an understanding of how wide or narrow a range of results would be achieved (Efron and Tibshirani, 1993). One way this is commonly done is by resampling a number of times the estimated distributions of various data sources that go into the model (Magnusson et al., 2013). Another more recently introduced method is the spatial-resampling bootstrap method of Elvarsson et al. (2014). This method takes advantage of spatial disaggregations in the data (i.e. data taken from different spatial areas) and randomly resamples data from these different areas with replacement. Given that composition data feed into all of these uncertainty methods it seems a reasonable argument that data of greater quality and quantity would result in greater confidence in the model assuming that the model is correctly specified.

In this study we use the spatial resampling bootstrap method of Elvarsson et al. (2014) to assess how decreased composition data (esp. age data) impact confidence intervals. This method works by taking spatially disaggregated data (i.e. data sampled in spatially discrete regions), randomly selecting available areas (with replacement), and using this resampled data from these areas to fit a model. This procedure is then performed some number of times and the middle $x\%$ of estimated parameter values from the n iterations are used to estimate the $x\%$ confidence interval. For example, let both the fishery-independent and non-fishery-independent data come from 100 spatially discrete areas. At each iteration in the spatial resampling bootstrap method 100 of these units would be randomly sampled with replacement, and the data from these randomly sampled areas would be used to fit a model (with data from some areas being duplicated and data from others not used at all). This procedure would then be repeated some number of times (typically either 100 or 1000) and the resulting range of parameter and biomass estimates used as a confidence interval on the estimates.

This method is wholly different than the typically used bootstrap method introduced by Efron and Tibshirani (1993) and tested by Magnusson et al. (2013). However, the resulting confidence intervals very much rely upon the appropriate data going into the model (i.e. better data and more data should lead to greater confidence in the proposed model). Decreasing composition data is likely to result in a lower confidence of the representation of the underlying processes that created the data, which we fully expect to see. The objective of this study is to determine how confidence intervals become tighter as composition data increases. Therefore, the main objective of this study is to determine if the width of confidence intervals continues to decrease with increasing composition data or if the level of confidence asymptotes with increasing composition data similar to the way that bias in parameter estimates does. Additionally, a secondary objective is to determine how confidence interval changes with composition data differ among life-history types. We compare different life-history types using simulations for life-histories that are based on a cod-like stock, a flatfish-like stock, and a small pelagic stock. We also compute confidence intervals on a stock assessment using data from Atlantic cod in Iceland whereby composition data are sequentially removed and bootstrap confidence intervals estimated. This is used as an approach to explore how confidence intervals change in a real-world example as data are removed. We use the integrated assessment modeling platform Gadget (Begley, 2012) to test the effects of reduced composition data on estimate of uncertainty in stock assessment.

10.2 Methods

To test the effects of reduced composition data on uncertainty estimates we created three operating models based on simulated life-histories and one operating model based on a real stock (Atlantic cod) from Icelandic waters. The three simulated operating models were based all based on fish that are present in Icelandic waters: Atlantic cod (*Gadus morhua* L.), European plaice (*Pleuronectes platessa*), and capelin (*Mallotus villosus*). To differentiate the simulated cod operating model from the stock assessment model for Icelandic cod we will refer to the former as cod_{sim}, and the latter as cod_{ice}. The other simulated operating models will be referred to as flatfish and capelin, respectively. Gadget was used as both an operating and estimation model in order to remove any potential biases that might exist among different types of models, which makes interpretation of the effects being tested more straightforward. For both OM and EM we used Gadget 2.2.00 (Begley, 2012). Additionally, we developed an R package, *gadgetSim* (Frater, 2018a), that easily connects R and Gadget for simple OM and EM setup as well as reading results using the R statistical software environment (R Core Team, 2017).

10.2.1 Operating Models

Operating models were based on the three life-history types described above. Parameters for these models were loosely based on data and assessments for the respective species from the Icelandic Marine and Freshwater Research Institute (Hafrannsóknastofnun, 2018). Parameter values needed that were not available were supplemented either from the literature or using FishBase (Froese and Pauly, 2017b) in cases where the literature was not informative.

Gadget is an age- and size-structured model where fish in each cohort are reduced according to fishing and natural mortality.

$$N_{a,y} = \begin{cases} N_{a-1,y-1} e^{-F_{y-1}-M_a}, & a < a_{max} \\ N_{a-1,y-1} e^{-F_{y-1}-M_a} + N_{a,y-1} e^{-F_{y-1}-M_a}, & a = a_{max} \end{cases}$$

where N_{ay} are numbers at age a in year y, M_a is the age-specific natural mortality, and F is asymptotic length selective fishing mortality.

Spawning in the OMs for each year was based on the Beverton-Holt spawner-recruit function (Beverton and Holt, 1957) but included multiplicative error as follows to create time-varying recruitment:

$$R_y = \frac{\mu S_y}{\lambda + S_y} \cdot e^{\mathrm{N}(0,\sigma_R)}$$

where R_y is the mean recruitment in year y, S_y is SSB in year y, μ and λ are species specific parameters controlling the steepness and maximum level of R, and σ_R is a species specific error term used to introduce time-varying recruitment. The value of σ_R for each species was chosen to induce enough process error to be clearly evident, but not so much that it masked the trajectory of each respective fishery development scenario ($\sigma_R = 0.02$ for cod and flatfish; $\sigma_R = 0.04$ for capelin). The resulting values were used to fix a recruitment value for each year within the OMs.

Growth was implemented as a variant of the von Bertalanffy growth curve:

$$\Delta L = (L_\infty - L_i)(1 - e^{-k\Delta t})$$

where L_∞ is the mean length L as time t approaches infinity, and k is the usual Brody growth coefficient. Variation in mean size-at-age was implemented using a beta-binomial distribution, and weight was based on the length-weight relationship: $W = \alpha L^\beta$. Initial numbers of fish for each age class at year 1 was given as

$N_{a,0} = 1\text{e}08e^{-M \times a}$. For further details on how populations are structured in Gadget, see Begley (2012).

For each life history type and fishery development scenario we simulated a model that ran for 120 years with the first 20 years serving as a burn-in period to allow OMs and EMs to stabilize. For ease of interpretation the burn-in period was discarded and hereafter each year is referred to as the respective year minus 20 (i.e. the final year is 100 and the beginning year at the end of the burn-in period is 0.

The fishery development scenarios were the commonly used fishing trajectories two-way-trip, fish down, and flat MSY. These trajectories were created specific to each life-history type by calculating the fishing mortality at maximum sustainable yield (F_{MSY}) and setting up the development scenarios in relation to that. We calculated F_{MSY} by computing yield curves across values of F and defined F_{MSY} as the value for F that resulted in the maximum yield. Other reference points used to determine fishery development scenarios will be referred to as F_{high} and F_{low} which were defined as the values for F that resulted in yield being 0.8 times the yield at F_{MSY}, where $F_{low} < F_{MSY} < F_{high}$ (Table 10.13). Fishing for all fishery development scenarios began in year 20. For the fish down scenario, the fishery began with an F of 0 and increased linearly to a high of F_{high} in year 100. The two-way-trip scenario began similarly, but reached an F of F_{high} in year 68 and then decreased to F_{MSY} in year 100. The flat MSY scenario began fishing in year 20 at F_{MSY} and remained there for the duration of the model.

10.2.2 Sampling the OM

We used output from the OMs as “data" on which to minimize the objective function in the EMs. To estimate confidence intervals we used the spatial resampling bootstrap method of (Elvarsson et al., 2014). This method randomly samples spatially defined sampling units with replacement and uses the samples from these spatial units (with

Table 10.12. Parameters used in the operating model for each life-history type. Dots for the $Length_{SD_{init}}$ parameter indicates a repitition of the final value until maximum age of the stock. Parameters for the growth, maturity, spawning, and fleet selectivity parameter types refer to the Gadget functions "lengthvbsimple", "exponential" length selection in the spawning file, Beverton-Holt spawner-recruit process, and "exponentiall50" suitability curve, respectively.

Parameter	Parameter Type	Cod	Capelin	Flatfish
Min. Age	Stock	1	1	1
Max. Age	Stock	20	5	30
Min. Length	Stock	1	1	1
Max. Length	Stock	150	25	100
$Length_{SD_{init}}$	Stock	2, 4, 6, 8, 10,...	1, 1.5, 2,...	2, 4, 6,...
M	Mortality	0.2	1	0.1
L_∞	Growth	125	20	55
k	Growth	0.15	0.6	0.14
t_0	Growth	-0.35	0	-1.3
α_{LW}	L-W Rel.	6.7e-06	7.1e-06	9.6e-06
β_{LW}	L-W Rel.	3.1	3.15	3
α_{mat}	Maturity	-0.2	-2	-0.25
$L50_{mat}$	Maturity	64	12.5	30
λ_{BH}	Spawning	1.07e+08	5e+07	1.07e+08
μ_{BH}	Spawning	3e+08	8e+10	2e+08
α_{fleet}	Fleet selectivity	0.1	1	0.2
$L50_{fleet}$	Fleet selectivity	45	7	30

replacement) to calculate a confidence interval on model parameters. Since our OMs were defined as single area models we created pseudo-areas in the OM output by defining 100 false units and dividing the number of fish from each year, step, and age into these false units using a Poisson distribution to calculate the proportion of each fish that should be divided into each unit.

$$N_{yta'l} = N_{ytl} \frac{\rho_{a'}}{\sum_{a'} \rho_{a'}}, \quad \rho_{a'} \sim \text{Pois}(3)$$

where N_{ytal} is the number at each stage of the model as before and $\rho_{a'}$ is the Poisson distribution modeled with $\delta = 3$.

Each bootstrap replicate sampled the OM with asymptotic selectivity similar to the

Table 10.13. Fishing mortality (F) values used to determine fishery development scenarios. Both F_{low} and F_{high} are the F values that produce 80% yield compared to F_{MSY}.

Species	F_{MSY}	F_{low}	F_{high}
Cod	0.22	0.11	0.45
Flatfish	0.12	0.06	0.22
Capelin	1.08	0.53	1.82

fishery selectivity, but with an inflection point scaled lower in order to sample a greater number of smaller-sized fish. Starting in year 70 we sampled OM output biannually by applying asymptotic selectivity on absolute numbers at year, timestep, age, and length combinations up to a maximum of 0.001% and with multiplicative error as follows:

$$I_{ytal} = S_l N_{ytal} e^{\mathrm{N}(\mu,\sigma)}, \qquad \mu = 0, \sigma = 0.2$$

where I_{ytal} is the index of abundance at each year (y), timestep (t), age (a), and length (l), S_l is the selectivity at length up to a maximum of 1% of total numbers N.

We used these as abundance data in our EMs. Composition data was sampled as a subset of the abundance index data. Length data were randomly sampled from the abundance indices at either 50 or 200 samples per year and step, and age samples were randomly selected from the length samples at the following rates for each scenario: 1 (i.e. all length data also aged), 0.75, 0.5, 0.25, 0.1, or 0 (i.e. length data only). This resulted in a total of 12 different models for each of the 3 life-history types and 3 fishery development scenarios for a total of 108 different scenarios.

Catches from fleets were recorded without error, but commercial sampling of composition data was bootstrapped and sampled similarly to survey sampling for composition data, but a size-selective curve was not applied as the fishery already implements length-based selectivity. Rather 0.001% of fishery output were sampled and used as length data regardless of size. Sampling these values to use as age data in the EMs was based on the appropriate scenario in accordance with the fishery-independent sampling described above. For each of the 108 distinct scenarios we performed the spatial resampling bootstrap method from above 100 times.

10.2.3 Estimation Models

EMs were set up for each independent data set under each scenario and fit using the iterative re-weighting procedure described by Stefansson (2003) and implemented by Taylor et al. (2007) and Elvarsson (2015). This procedure consists of performing a separate optimization run for each likelihood component (e.g. indices of abundance, each composition data set, etc.) using a heavy weight on that component. Variance for that component optimization is estimated as the overall likelihood score (SS_c) divided by degrees of freedom (df_c) $\hat{\sigma}_c^2 = \frac{SS_c}{df_c^*}$ where the number of non-zero data points (df_c^*) serves as a proxy for df_c (Elvarsson, 2015). After all component-weighted optimization runs have finished a final optimization run is performed using the inverse variance from above as the weight on that respective component ($w_c = 1/\hat{\sigma}_c^2$). We used the R package Rgadget (Elvarsson and Lentin, 2018) to automate this procedure. The likelihood that was used for composition data was the sum-of-squares function in Gadget. Abundance indices were fitted using a log-linear regression of estimated to observed values with a slope fixed at 1 and intercept estimated internally.

All model parameters except for M_a were estimated. These include parameters for growth (simplified von Bertalanffy function), growth implementation (beta-binomial), initial starting values, Beverton-Holt spawner-recruit parameters, spawning maturity, and selectivity parameters for surveys and fleets. It should be noted that spawning maturity was implemented as the proportion function in the Gadget spawning file and not as a separate stock that matures into a spawning stock.

For each data scenario, life-history type, and fishery development scenario we optimized the 100 bootstrap replicates each consisting of a different randomized data set. All models were assessed for convergence by ensuring that the convergence criteria in the optimization procedure was met.

10.2.4 Performance Assessment

We assessed the performance of EMs under each scenario using relative error (*RE*) and median absolute relative error (*MARE*) of estimated parameters and both spawning stock biomass in the terminal year (SSB_{100}) as well as depletion (DEP = SSB_0/ SSB_{100}).

$$RE = (\hat{\theta} - \theta)/\theta$$

$$MARE = \text{median}(|\hat{\theta} - \theta)|/\theta)$$

10.2.5 Cod Assessment Example

In addition to the above simulations we also performed a similar procedure on the stock of Atlantic cod *Gadus morhua* from the waters around Iceland. An integrated stock assessment for cod was created similar to that of Taylor et al. (2007), but updated to include the years 1985–2015. This stock assessment was then bootstrapped 100 times at each of the same age data proportions as the simulation models above (i.e. 1, 0.75, 0.5, 0.25, 0.1, and 0 – length only). 95% CIs were calculated for estimated parameters and model output and compared across the different age data sample size scenarios to assess how age data affects CIs in this particular stock.

10.3 Results

10.3.1 Simulation Scenarios

Parameter estimate bias and confidence intervals of the parameter estimates both decreased with increasing age data. In addition, confidence intervals were tighter in scenarios with greater quantities of length data. These results were most notable for the μ_{BH} parameter (Fig. 10.39). Other parameters showed results either similar to μ_{BH} or no differences across the age and length scenarios. The results for μ_{BH} were especially evident for both cod and capelin. Median estimates of μ_{BH} for flatfish were relatively similar across both age and length scenarios, but confidence intervals were generally tighter for scenarios with at least 25% of the full age data present. Beyond this level of age data it was difficult to distinguish any differences in confidence intervals for flatfish.

The trend of lower bias and tighter confidence intervals was perhaps even more notable in measures of both depletion and terminal biomass (Figs. 10.40–10.41). Depletion estimates and confidence intervals of those estimates decreased with increasing proportion of age data samples for both cod and capelin. Bias also seemed to decrease and confidence increase with greater quantities of age data to a greater degree in the scenarios with greater length data sample sizes. That is, in the scenarios with only 50 length samples per year increased age data samples did not seem to decrease bias or increase confidence of depletion estimates as much as the scenarios with 200 length samples per year although this pattern seemed to hold more for depletion estimates than for measures of terminal biomass. Estimates of biomass in the terminal year seemed to follow the trend of higher confidence with greater quantities of age data for capelin, but not nearly as much for cod and especially flatfish (Fig. 10.41).

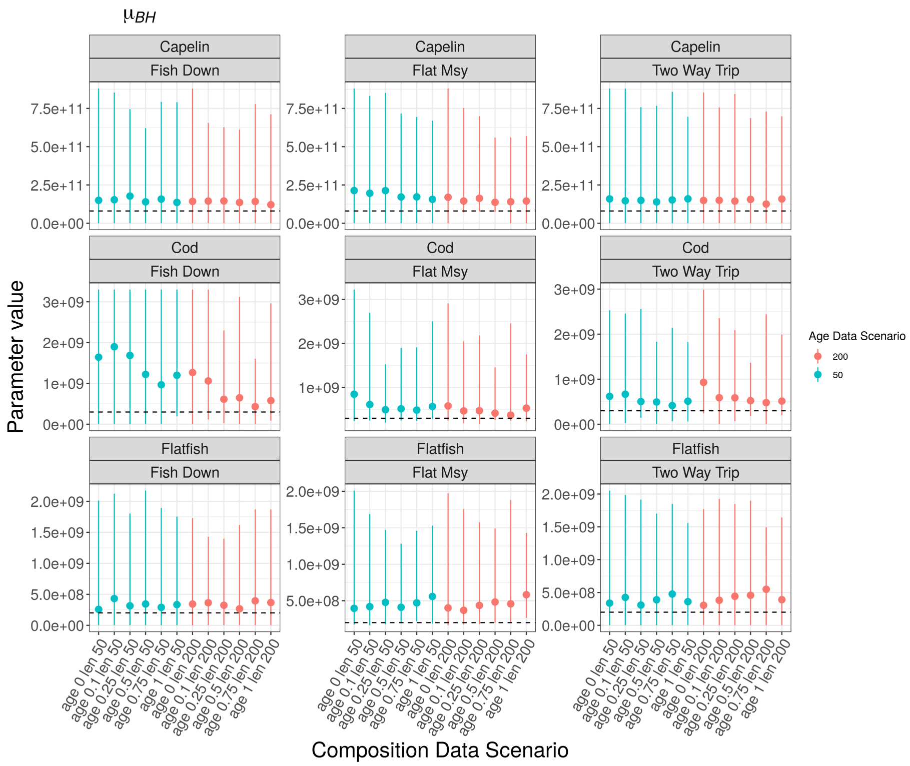


Figure 10.39. Median parameter estimates and 95% confidence interval from bootstrapped models for six different age proportion scenarios and two length scenarios for the μ parameter of the Beverton-Holt stock recruit equation.

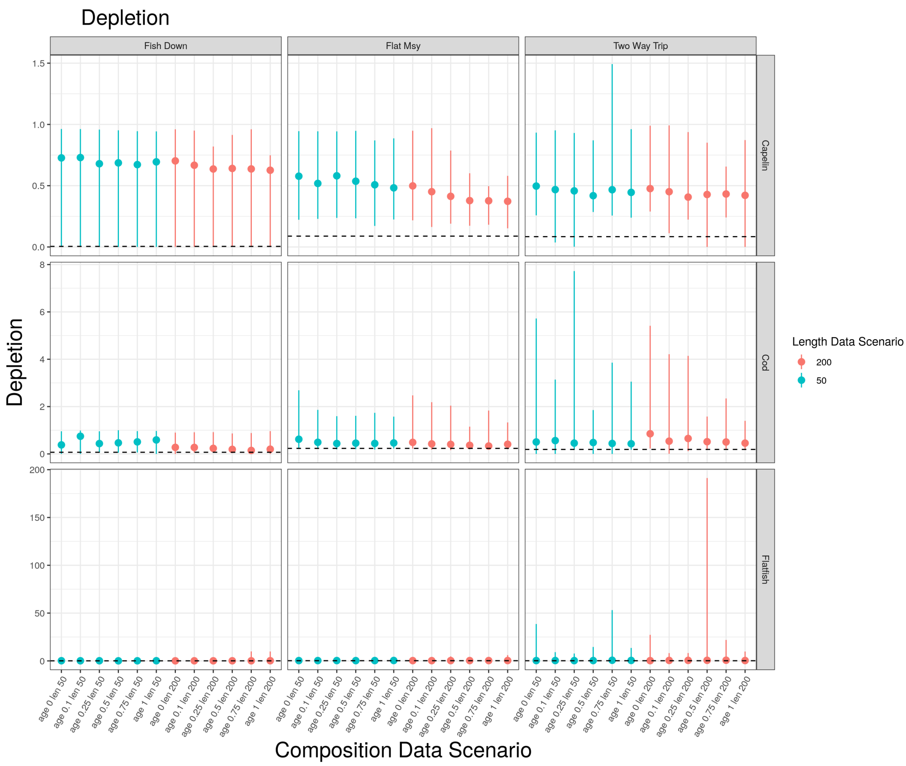


Figure 10.40. Estimates an 95% confidence intervals of stock depletion in bootstrapped models.

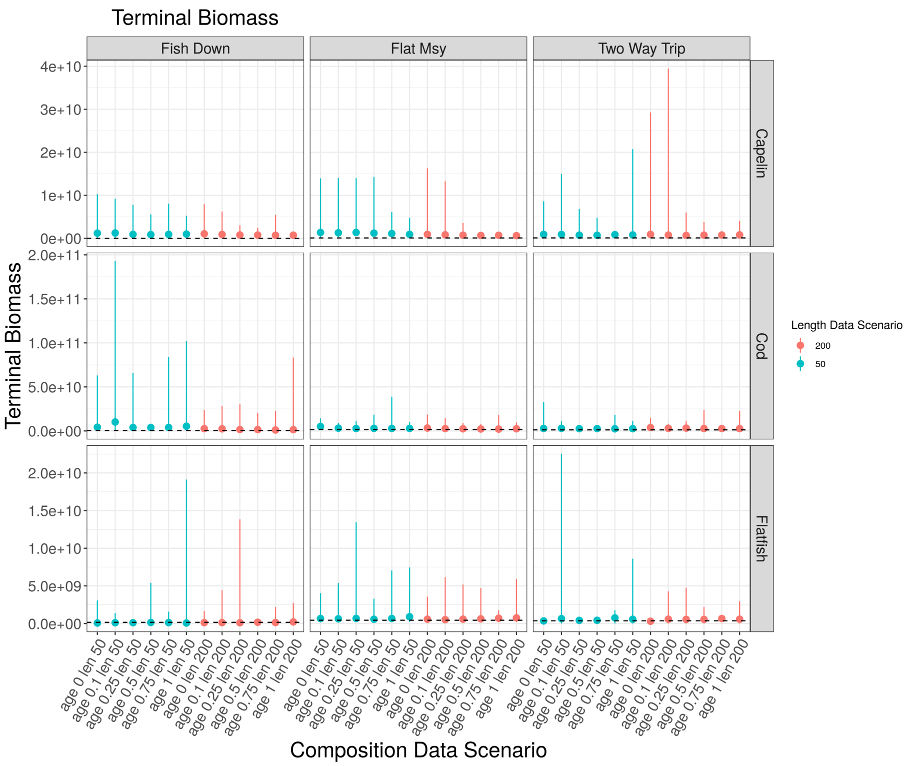


Figure 10.41. Estimates and 95% confidence intervals of terminal biomass in bootstrapped models.

10.3.2 Cod Assessment Scenarios

Decreasing age data samples sizes proportionally to the full data available from surveys generally increased the estimated number and biomass of fish in the Icelandic cod stock (Fig. 10.42–10.45). The 95% CI were not largely affected by decreasing the age data sample sizes in proportion to the full data available. 95% CI for numbers perhaps decreased slightly with greater availability of age data (Fig. 10.42), but actually increased for biomass with greater sample sizes of age data. This is reflected in the parameter estimates. The 95% CI of recruitment parameters decreased with greater age data sample sizes (i.e. higher confidence with more age data–Fig. 10.46) whereas the 95% CI of the both the L_∞ and to some degree the k parameters decreased with increasing age data sample sizes (i.e. lower confidence with more age data–Fig. 10.48). The 95% CI on selectivity parameters seemed to be relatively unaffected by the proportion of age data samples in the model.

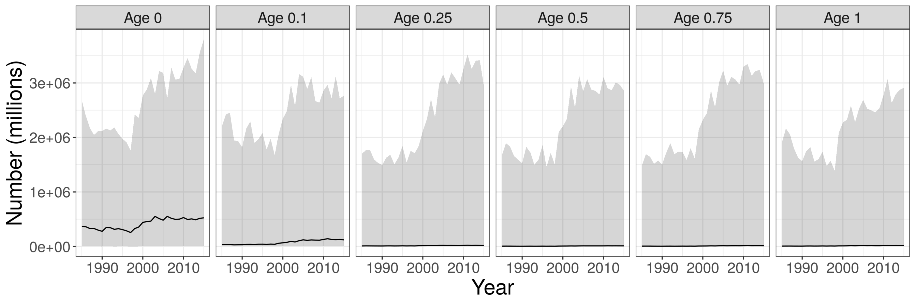


Figure 10.42. Number of fish and 95% confidence interval of bootstrapped cod assessment models across age data sample size proportions.

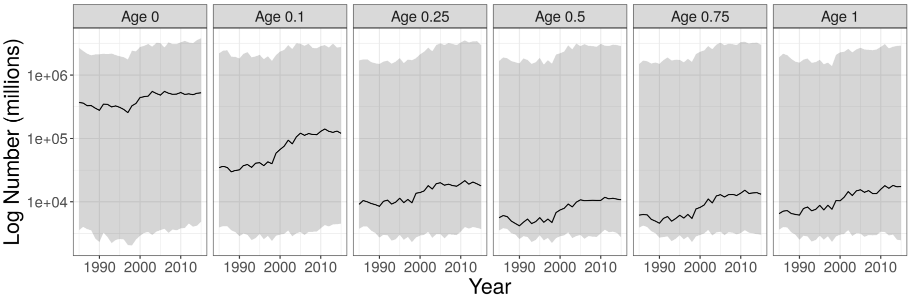


Figure 10.43. Same as Fig. 10.42, but with $\log_{10}$ *scale on the y-axis.*

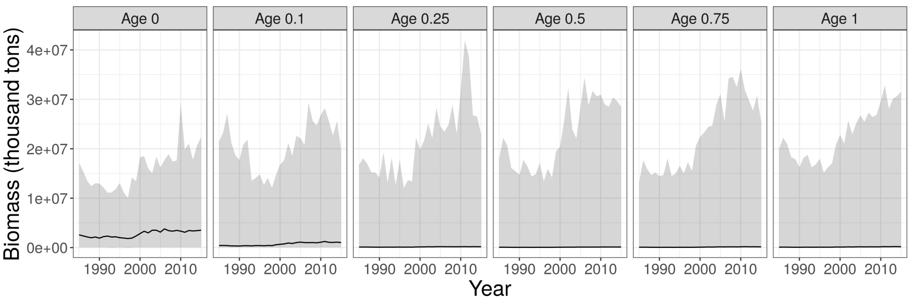


Figure 10.44. Number of fish and 95% confidence interval of bootstrapped cod assessment models across age data sample size proportions.

10.4 Discussion

The decreased bias and tighter CIs in a number of parameters and biomass estimates associated with proportionally higher sample sizes of age data is unsurprising. There is a rich field of study within statistics that is dedicated to the relationship between sample size and model certainty (i.e. power – Cohen, 2013; Kraemer and Blasey, 2015; Aberson, 2019). Additionally, and specific to fisheries stock assessment, a number of studies have shown decreased bias with higher sample sizes of composition data (Wetzel and Punt, 2011; Ono et al., 2015; He et al., 2016). The results from our simulation study are novel as they show the importance of composition data to uncertainty estimates in integrated assessment models. The effects of sample size on confidence intervals in

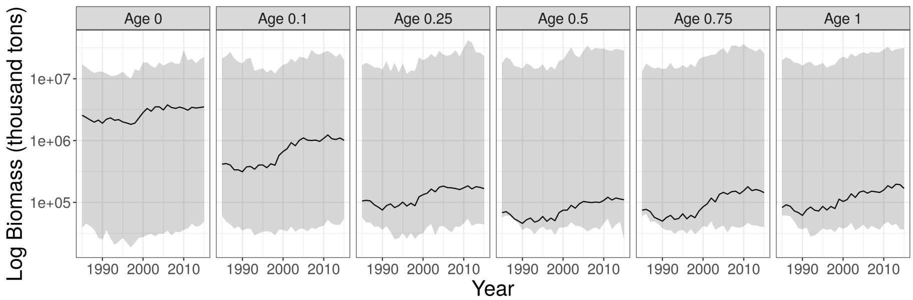


Figure 10.45. Same as Fig. 10.44, but with $\log_{10}$ *scale on the y-axis.*

nonlinear models has been explored more generally (Htjet et al., 1990), and increased sample size has been shown to decrease confidence interval width to a certain point (Royston, 2007). While methods of uncertainty estimates have been previously studied in fisheries models (Magnusson et al., 2013; Magnusson, 2016), the effects of sample size (esp. composition data) on uncertainty estimates in fisheries models is not well studied.

A result from this study that are rather more surprising is the uniformity in confidence intervals across age data sample sizes for estimates of total numbers in the cod assessment and wider confidence intervals with increased age data for biomass. The scenario with no age data whatsoever in the cod assessment had the narrowest CIs for biomass, whereas those with full age data had the widest. It seems that even just 25% of the full available age data is enough to stabilize estimates in the median values of both numbers and biomass in the the cod assessment, while models using 50% of available age data were almost indistinguishable from those with higher age data sample sizes (in terms of median value estimates for both numbers and biomass). The CIs, however, did not follow the same trend as the median point estimates of numbers and biomass. The 95% CI for numbers in the cod assessment model with full age data were almost just as wide as those with no age data, and the width of the 95% CI for biomass just decreased as age data was removed.

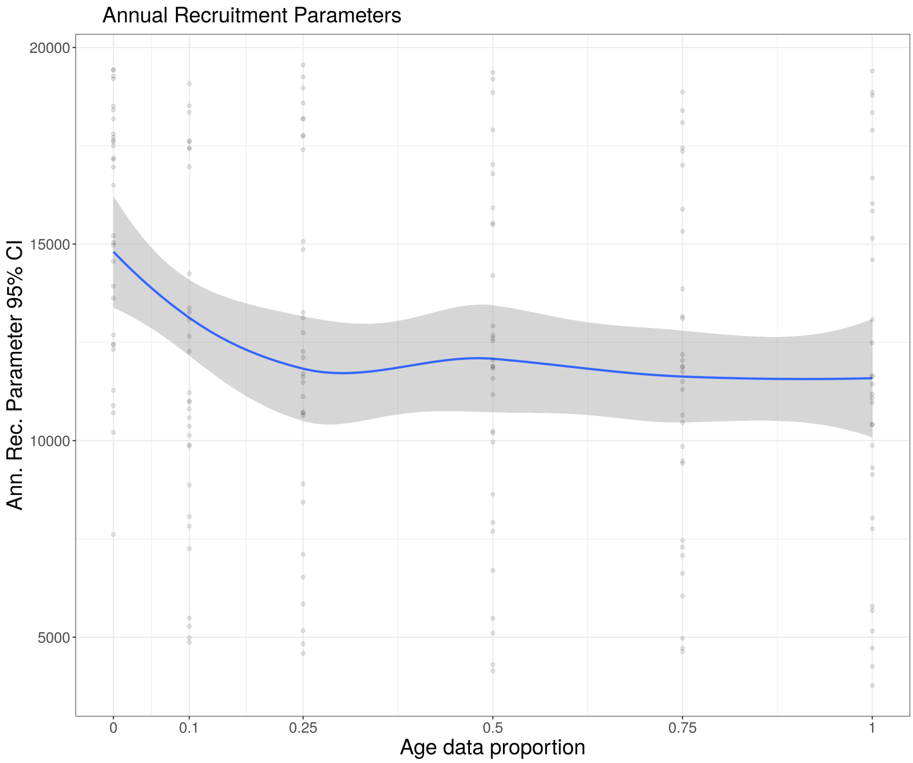


Figure 10.46. 95% CI range of annual recruitment parameters across age data sample size proportions. Each dot represents the 95% range for the recruitment parameter of a particular year from 1985–2015. The blue line is a loess-smoothed line across age data proportions.

The static width of the CIs across models for numbers of fish in the cod assessment was especially surprising given the increased confidence in recruitment parameters with increasing age data. It would make sense that if recruitment parameter CIs were tighter with increasing age then so would be the confidence on estimates of numbers. He and Field (2018) showed that greater recruitment variability can lead to high uncertainty in estimation of recruitment parameters, so perhaps the recruitment variability in the example stock of Icelandic cod used here is high enough that uncertainty in estimates of numbers are doomed to be high no matter how much age data is present.

Our results show that, at least in certain scenarios, fitting integrated assessment models

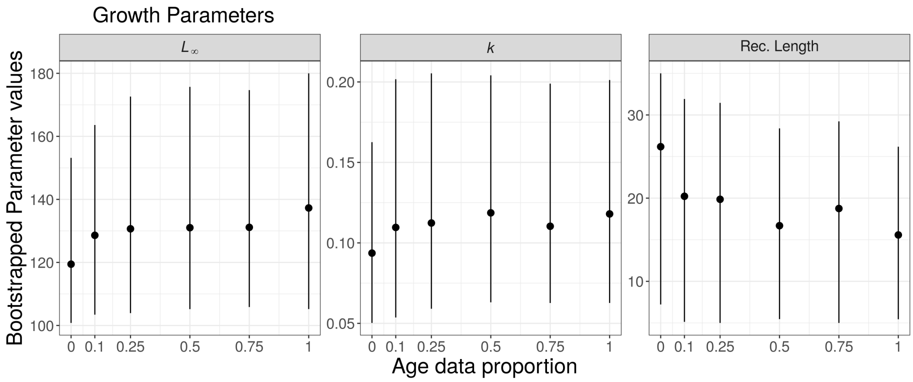


Figure 10.47. Estimates and 95% confidence intervals growth parameters across age data sample size proportions.

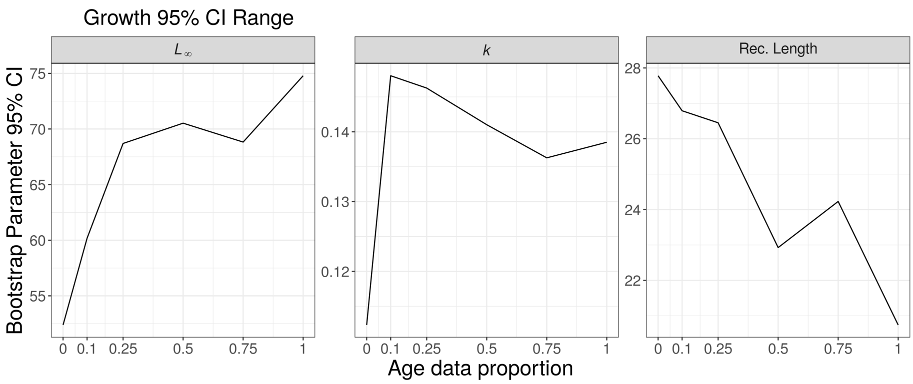


Figure 10.48. Range of 95% confidence intervals for growth parameters across models.

using higher sample sizes of age data leads to narrower confidence intervals. However, in other scenarios confidence intervals were relatively unchanged across sample size proportions of age data. This study represents an initial attempt at trying to understand the effect of age data, and composition data more generally, on confidence intervals in integrated assessment. This study should be extended using other integrated assessment models (i.e. Stock Synthesis 3, etc.) and other methods of uncertainty estimation.

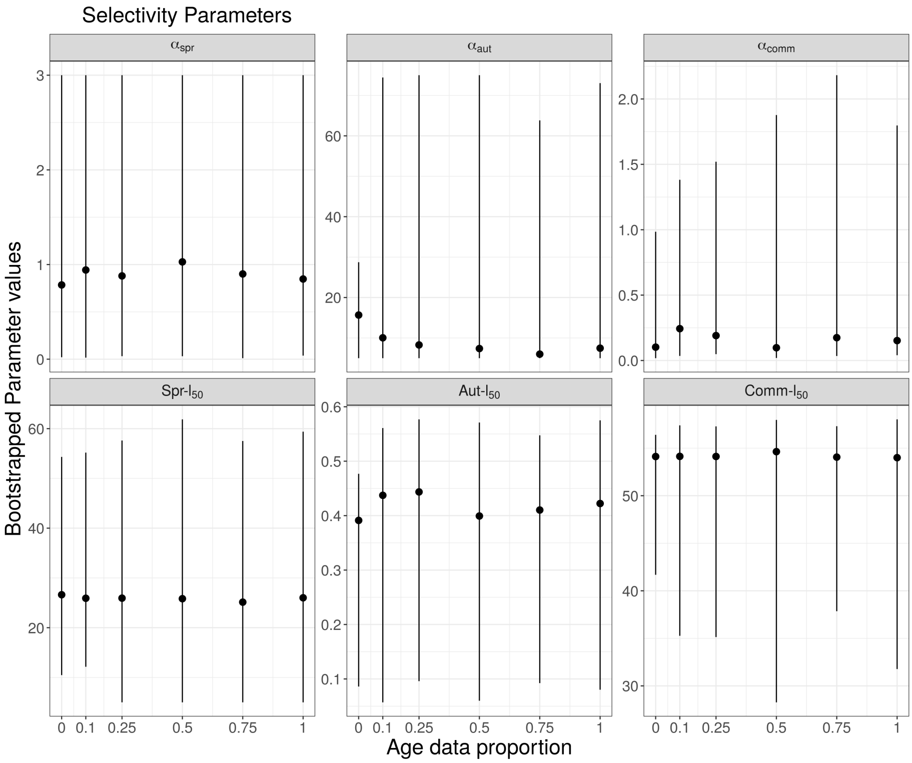


Figure 10.49. Estimates and 95% confidence intervals selectivty parameters across age data sample size proportions.

Acknowledgements

Funding for the lead author to undertake his Ph.D. studies was received from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 613571 for the project MareFrame and from the European Commission's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 634495 for the project Science, Technology, and Society Initiative to minimize Unwanted Catches in European Fisheries (Minouw) as well as from the Icelandic Research Fund (Rannís, No. 152039051). PNF is grateful to GS for acquiring said funding.

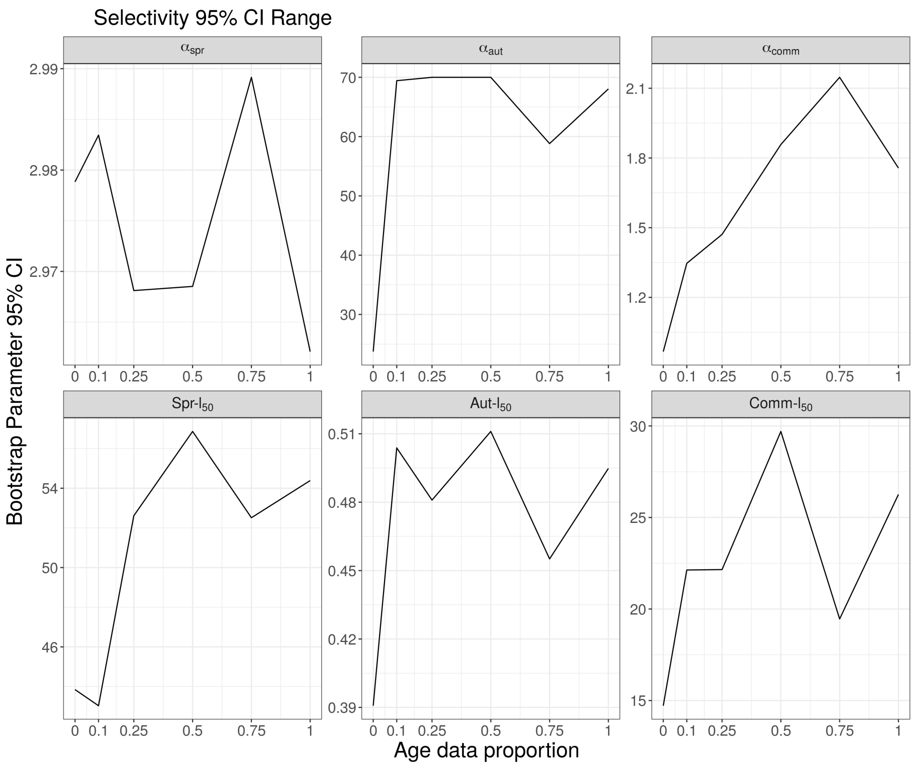


Figure 10.50. Range of 95% confidence intervals for selectivity parameters across models.

References

Aberson, C. L. (2019). *Applied power analysis for the behavioral sciences*. Routledge.

Alós, J., Palmer, M., Balle, S., Grau, A. M., and Morales-Nin, B. (2009). Individual growth pattern and variability in serranus scriba: a bayesian analysis. *ICES Journal of Marine Science: Journal du Conseil*, 67(3):502–512.

Andersen, K. H. and Brander, K. (2009). Expected rate of fisheries-induced evolution is slow. *Proceedings of the National Academy of Sciences*, 106(28):11657–11660.

Anderson, N., Ottersen, G., and Swain, D. (2002). Report of the ices/globec workshop on the dynamics of growth in cod. Technical report.

Bal, G., Rivot, E., Prévost, E., Piou, C., and Baglinière, J.-L. (2011). Effect of water temperature and density of juvenile salmonids on growth of young-of-the-year atlantic salmon *Salmo salar*. *Journal of Fish Biology*, 78(4):1002–1022.

Baranov, F. I. (1918). K voprosu o biologicheskikh osnovaniyakh rylmogo khozyaistva. [On the question of the biological basis of fisheries.]. *Izvestiya otdela rybovodstva i nauchno-promyslovykh issledovanii*, 1:81–128.

Barneche, D. R. and Allen, A. P. (2018). The energetics of fish growth and how it constrains food-web trophic structure. *Ecology Letters*, 21(6):836–844.

Barneche, D. R., Robertson, D. R., White, C. R., and Marshall, D. J. (2018). Fish reproductive-energy output increases disproportionately with body size. *Science*, 360(6389):642–645.

Basson, M., Rosenberg, A., and Beddington, J. (1988). The accuracy and reliability of two new methods for estimating growth parameters from length-frequency data.

ICES Journal of Marine Science: Journal du Conseil, 44(3):277–285.

Begley, J. (2012). *Gadget User Guide*. Marine Research Institute Report Series. Available at http://www.hafro.is/gadget/files/userguide.pdf.

Béné, C., Macfadyen, G., and Allison, E. H. (2007). Increasing the contribution of small-scale fisheries to poverty alleviation and food security. Technical Report 481, FAO, Rome.

Beverton, R. J. and Holt, S. J. (1957). *On the dynamics of exploited fish populations*. Springer Science & Business Media.

Björnsson, B. and Steinarsson, A. (2002). The food-unlimited growth rate of atlantic cod (gadus morhua). *Canadian Journal of Fisheries and Aquatic Sciences*, 59(3):494–502.

Björnsson, B., Steinarsson, A., and Árnason, T. (2007). Growth model for atlantic cod (gadus morhua): Effects of temperature and body weight on growth rate. *Aquaculture*, 271(1-4):216–226.

Bogstad, B. and Gjøsæter, H. (2001). Predation by cod (gadus morhua) on capelin (mallotus villosus) in the barents sea: implications for capelin stock assessment. *Fisheries Research*, 53(2):197–209.

Bolker, B. M., Gardner, B., Maunder, M., Berg, C. W., Brooks, M., Comita, L., et al. (2013). Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. *Methods in Ecology and Evolution*, 4(6):501–512.

Borah, M. and Mahanta, D. J. (2018). An approach to estimate the parameters of von bertalanffy growth models in forestry. *Indian Forester*, 144(1):41–53.

Bowker, D. (1995). Modelling the patterns of dispersion of length at age in teleost fishes. *Journal of Fish Biology*, 46(3):469–484.

Brander, K. (1995). The effect of temperature on growth of atlantic cod (*Gadus morhua* l.). *ICES Journal of Marine Science: Journal du Conseil*, 52(1):1–10.

Brander, K. (2000). Effects of environmental variability on growth and recruitment in cod (gadus morhua) using a comparative approach. *Oceanologica Acta*, 23(4):485–

496.

Brander, K. M. (2007). The role of growth changes in the decline and recovery of north atlantic cod stocks since 1970. *ICES Journal of Marine Science: Journal du Conseil*, 64(2).

Catalano, M. J. and Allen, M. S. (2010). A size-and age-structured model to estimate fish recruitment, growth, mortality, and gear selectivity. *Fisheries Research*, 105(1):38–45.

Chadwick-Furman, N. E., Goffredo, S., and Loya, Y. (2000). Growth and population dynamic model of the reef coral fungia granulosa klunzinger, 1879 at eilat, northern red sea. *Journal of Experimental Marine Biology and Ecology*, 249(2):199–218.

Chen, Y., Chen, L., and Stergiou, K. I. (2003). Impacts of data quantity on fisheries stock assessment. 65(1):92–98.

Chen, Y., Jackson, D., and Harvey, H. (1992). A comparison of von bertalanffy and polynomial functions in modelling fish growth data. *Canadian Journal of Fisheries and Aquatic Sciences*, 49(6):1228–1235.

Christensen, V. and Walters, C. J. (2004). Ecopath with ecosim: methods, capabilities and limitations. *Ecological Modelling*, 172(2-4):109–139.

Clark, W. G. and Hare, S. R. (2002). Effects of climate and stock size on recruitment and growth of pacific halibut. *North American Journal of Fisheries Management*, 22(3):852–862.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Contreras-Reyes, J. E., Quintero, F. O. L., and Wiff, R. (2018). Bayesian modeling of individual growth variability using back-calculation: Application to pink cusk-eel (genypterus blacodes) off chile. *Ecological Modelling*, 385:145–153.

Cook, R., Sinclair, A., and Stefansson, G. (1997). Potential collapse of north sea cod stocks. *Nature*, 385(6616):521.

Cope, J. M. and Punt, A. E. (2007). Admitting ageing error when fitting growth curves: an example using the von bertalanffy growth function with random effects. *Canadian*

Journal of Fisheries and Aquatic Sciences, 64(2):205–218.

Costello, C., Ovando, D., Hilborn, R., Gaines, S. D., Deschenes, O., and Lester, S. E. (2012). Status and solutions for the world's unassessed fisheries. *Science*, 338(6106):517–520.

Cotter, A., Burt, L., Paxton, C., Fernandez, C., Buckland, S., and Pan, J.-X. (2004). Are stock assessment methods too complicated? *Fish and Fisheries*, 5(3):235–254.

Crone, P. R. and Valero, J. L. (2014). Evaluation of length-vs. age-composition data and associated selectivity assumptions used in stock assessments based on robustness of derived management quantities. *Fisheries Research*, 158:165–171.

Crowder, L. B., Hazen, E. L., Avissar, N., Bjorkland, R., Latanich, C., and Ogburn, M. B. (2008). The impacts of fisheries on marine ecosystems and the transition to ecosystem-based management. *Annual Review of Ecology, Evolution, and Systematics*, 39:259–278.

Dalpadado, P. and Bogstad, B. (2004). Diet of juvenile cod (age 0–2) in the barents sea in relation to food availability and cod growth. *Polar biology*, 27(3):140–154.

de Zárate, V. O. and Babcock, E. (2016). Estimating individual growth variability in albacore (thunnus alalunga) from the north atlantic stock: Aging for assessment purposes. *Fisheries Research*, 180:54–66.

Deriso, R. B. (1987). Optimal f 0.1 criteria and their relationship to maximum sustainable yield. *Canadian Journal of Fisheries and Aquatic Sciences*, 44(S2):s339–s348.

Deroba, J., Butterworth, D. S., Methot, R., De Oliveira, J., Fernandez, C., Nielsen, A., Cadrin, S., Dickey-Collas, M., Legault, C., Ianelli, J., et al. (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: Journal du Conseil*, 72(1):19–30.

Dortel, E., Massiot-Granier, F., Rivot, E., Million, J., Hallier, J.-P., Morize, E., Munaron, J.-M., Bousquet, N., and Chassot, E. (2013). Accounting for age uncertainty in growth modeling, the case study of yellowfin tuna (*Thunnus albacares*) of the Indian ocean. *PloS one*, 8(4):e60886.

Dowling, N. A., Smith, A. D., Smith, D. C., Parma, A. M., Dichmont, C. M., Sainsbury, K., Wilson, J. R., Dougherty, D. T., and Cope, J. M. (2019). Generic solutions for data-limited fishery assessments are not so simple. 20(1):174–188.

Dyck, A. J. and Sumaila, U. R. (2010). Economic impact of ocean fish populations in the global fishery. *Journal of Bioeconomics*, 12(3):227–243.

Efron, B. and Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman and Hall, New York, NY.

Elvarsson, B., Taylor, L., Trenkel, V., Kupca, V., and Stefansson, G. (2014). A bootstrap method for estimating bias and variance in statistical fisheries modelling frameworks using highly disparate datasets. *African Journal of Marine Science*, 36(1):99–110.

Elvarsson, B. T. (2015). *Statistical models of marine multispecies ecosystems*. PhD thesis, University of Iceland.

Elvarsson, B. T. and Lentin, J. (2018). Rgadget: R package for Gadget models. R package Version 0.5.

Elvarsson, B. T., Woods, P. J., Björnsson, H., Lentin, J., and Thordarson, G. (2018). Pushing the limits of a data challenged stock: A size-and age-structured assessment of ling (molva molva) in icelandic waters using gadget. *Fisheries Research*, 207:95–109.

Enberg, K., Jørgensen, C., Dunlop, E. S., Varpe, Ø., Boukal, D. S., Baulier, L., Eliassen, S., and Heino, M. (2012). Fishing-induced evolution of growth: concepts, mechanisms and the empirical evidence. *Marine Ecology*, 33(1):1–25.

Fabens, A. J. et al. (1965). Properties and fitting of the von bertalanffy growth curve. *Growth*, 29:265–289.

FAO (2018). The state of world fisheries and aquaculture 2018 - meeting the sustainable development goals. Technical report, Food and Agricultural Organization - United Nations, Rome. Licence: CC BY-NC-SA 3.0 IGO.

Finkbeiner, E. M., Bennett, N. J., Frawley, T. H., Mason, J. G., Briscoe, D. K., Brooks, C. M., Ng, C. A., Ourens, R., Seto, K., Switzer Swanson, S., et al. (2017). Reconstructing overfishing: Moving beyond malthus for effective and equitable solutions. *Fish and Fisheries*, 18(6):1180–1191.

Fishbase (2017). Species population growth list - Gadus morhua. http://www.fishbase.se/popdyn/PopGrowthList.php?ID=69&GenusName=Gadus&SpeciesName=morhua&fc=183. Accessed 10 December 2017.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., and Sibert, J. (2012). Ad model builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, 27(2):233–249.

Francis, R. C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(6):1124–1138.

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

Francis, R. C. (2017). Revisiting data weighting in fisheries stock assessment models. *Fisheries Research*, 192:5–15.

Francis, R. C., Aires-da Silva, A. M., Maunder, M. N., Schaefer, K. M., and Fuller, D. W. (2016). Estimating fish growth for stock assessments using both age–length and tagging-increment data. *Fisheries Research*, 180:113–118.

Frater, P. (2018a). gadgetsim: A package for creating simulated datasets from gadget output. R package - 0.0.2.

Frater, P. N. (2018b). gadgetSim: R package for creating Gadget simulation studies. R package Version 0.0.2. https://github.com/inspktrGadget/gadgetSim.

Frater, P. N. and Stefansson, G. (2019). Aspects of both growth and selectivity affect growth parameter estimation bias. *Fisheries Research*, 212:154–161.

Frater, P. N. and Stefansson, G. (2020). Comparison and evaluation of approaches aimed at correcting or reducing selectivity bias in growth parameter estimates for fishes. *Fisheries Research*, 225.

Froese, R. and Pauly, D. (2017a). Fishbase. http://www.fishbase.org/Images/Glospic/G_Fig13a6181_FL.jpg. version (01/2019).

Froese, R. and Pauly, D. (2017b). Fishbase. www.fishbase.org. version (06/2017).

Froese, R., Winker, H., Coro, G., Demirel, N., Tsikliras, A. C., Dimarchopoulou, D., Scarcella, G., Probst, W. N., Dureuil, M., Pauly, D., et al. (2018). A new approach for estimating stock status from length frequency data. *ICES Journal of Marine Science: Journal du Conseil*, 75(6):2004–2015.

Fry, F. E. (1949). Statistics of a lake trout fishery. *Biometrics*, 5(1):27–67.

Fulton, E. A., Link, J. S., Kaplan, I. C., Savina-Rolland, M., Johnson, P., Ainsworth, C., Horne, P., Gorton, R., Gamble, R. J., Smith, A. D., et al. (2011). Lessons in modelling and management of marine ecosystems: the atlantis experience. *Fish and Fisheries*, 12(2):171–188.

Fulton, E. A., Punt, A. E., Dichmont, C. M., Harvey, C. J., and Gorton, R. (2019). Ecosystems say good management pays off. *Fish and Fisheries*, 20(1):66–96.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 2. CRC press Boca Raton, FL.

Gelman, A., Rubin, D. B., et al. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Geromont, H. F. and Butterworth, D. S. (2015). Complex assessments or simple management procedures for efficient fisheries management: a comparative study. *ICES Journal of Marine Science: Journal du Conseil*, 72(1):262–274.

Gislason, H., Daan, N., Rice, J. C., and Pope, J. G. (2010). Size, growth, temperature and the natural mortality of marine fish. *Fish and Fisheries*, 11(2):149–158.

Gjøsæter, H., Dalpadado, P., and Hassel, A. (2002). Growth of barents sea capelin (*Mallotus villosus*) in relation to zooplankton abundance. *ICES Journal of Marine Science: Journal du Conseil*, 59(5):959–967.

Goodyear, C. P. (1995). Mean size at age: an evaluation of sampling strategies with simulated red grouper data. *Transactions of the American Fisheries Society*, 124(5):746–755.

Gwinn, D. C., Allen, M. S., and Rogers, M. W. (2010). Evaluation of procedures to reduce bias in fish growth parameter estimates resulting from size-selective sampling. *Fisheries Research*, 105(2):75–79.

Hafrannsóknastofnun (2018). Icelandic groundfish survey. Hafrannsoknastofnun - The Icelandic Marine and Freshwater Research Institute – http://www.hafro.is/index_eng.php. Fish surveys sampled bi-annually via standardized bottom trawl.

Hafrannsóknastofnun (2019a). Environmental conditions. data retrieved from https://data.hafro.is/environ/readme.html – accessed Feb. 25, 2019.

Hafrannsóknastofnun (2019b). MFRI | Marine and Freshwater Research Institute. https://data.hafro.is/en – accessed Feb. 25, 2019.

Handegard, N. O. and Tjøstheim, D. (2005). When fish meet a trawling vessel: examining the behaviour of gadoids using a free-floating buoy and acoustic split-beam tracking. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(10):2409–2422.

Harris, J. E., Newlon, C., Howell, P. J., Koch, R. C., and Haeseker, S. L. (2018). Modelling individual variability in growth of bull trout in the walla walla river basin using a hierarchical von bertalanffy growth model. 27(1):103–115.

He, X. and Field, J. C. (2018). Effects of recruitment variability and fishing history on estimation of stock-recruitment relationships: Two case studies from us west coast fisheries. *Fisheries Research*.

He, X., Field, J. C., Pearson, D. E., and Lefebvre, L. S. (2016). Age sample sizes and their effects on growth estimation and stock assessment outputs: three case studies from us west coast fisheries. *Fisheries Research*, 180:92–102.

Heino, M., Diaz Pauli, B., and Dieckmann, U. (2015). Fisheries-induced evolution. *Annual Review of Ecology, Evolution, and Systematics*, 46:461–480.

Hidalgo, M., Olsen, E. M., Ohlberger, J., Saborido-Rey, F., Murua, H., Piñeiro, C., and Stenseth, N. C. (2014). Contrasting evolutionary demography induced by fishing: the role of adaptive phenotypic plasticity. *Ecological Applications*, 24(5):1101–1114.

Hsieh, C.-h., Yamauchi, A., Nakazawa, T., and Wang, W.-F. (2010). Fishing effects on age and spatial structures undermine population stability of fishes. *Aquatic Sciences*, 72(2):165–178.

Htjet, S., Jolivet, E., and Messean, A. (1990). Some simulations results about confidence intervals and bootstrap methods in nonlinear regression. *Statistics: A Journal of Theoretical and Applied Statistics*, 21(3):369–432.

Hulson, P.-J. F. and Hanselman, D. H. (2014). Tradeoffs between bias, robustness, and common sense when choosing selectivity forms. *Fisheries Research*, 158:63–73.

Ichinokawa, M., Okamura, H., and Takeuchi, Y. (2014). Data conflict caused by model mis-specification of selectivity in an integrated stock assessment model and its potential effects on stock status estimation. *Fisheries Research*, 158:147–157.

Jensen, A. (1996). Beverton and holt life history invariants result from optimal trade-off of reproduction and survival. *Canadian Journal of Fisheries and Aquatic Sciences*, 53(4):820–822.

Jiao, Y., Rogers-Bennett, L., Taniguchi, I., Butler, J., and Crone, P. (2010). Incorporating temporal variation in the growth of red abalone (haliotis rufescens) using hierarchical bayesian growth models. *Canadian Journal of Fisheries and Aquatic Sciences*, 67(4):730–742.

Kimura, D. K. (1980). Likelihood methods for the von bertalanffy growth curve. *Fishery Bulletin*, 77(4):765–776.

Kraemer, H. C. and Blasey, C. (2015). *How many subjects?: Statistical power analysis in research*. Sage Publications.

Kraus, G. and Diekmann, R. (2018). Impact of fishing activities on marine life. In Salomon, M. and Markus, T., editors, *Handbook on Marine Environment Protection*, pages 79–96. Springer.

Kuriyama, P. T., Ono, K., Hurtado-Ferro, F., Hicks, A. C., Taylor, I. G., Licandeo, R. R., Johnson, K. F., Anderson, S. C., Monnahan, C. C., Rudd, M. B., et al. (2016). An empirical weight-at-age approach reduces estimation bias compared to modeling parametric growth in integrated, statistical stock assessment models when growth is time varying. 180:119–127.

Lee, H.-H., Maunder, M. N., Piner, K. R., and Methot, R. D. (2011). Estimating natural mortality within a fisheries stock assessment model: an evaluation using simulation analysis based on twelve stock assessments. *Fisheries Research*, 109(1):89–94.

Lee, H.-H., Piner, K. R., Methot Jr, R. D., and Maunder, M. N. (2014). Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: an example using blue marlin in the pacific ocean. *Fisheries Research*, 158:138–146.

Linde, M., Palmer, M., and Alós, J. (2011). Why protogynous hermaphrodite males are relatively larger than females? testing growth hypotheses in mediterranean rainbow wrasse coris julis (linnaeus, 1758). *Environmental Biology of Fishes*, 92(3):337–349.

Lockwood, S. J. (1974). The use of the von bertalanffy growth equation to describe the seasonal growth of fish. *ICES Journal of Marine Science: Journal du Conseil*, 35(2):175–179.

Lopez Quintero, F. O., Contreras-Reyes, J. E., Wiff, R., and Arellano-Valle, R. B. (2017). Flexible bayesian analysis of the von bertalanffy growth function with the use of a log-skew-t distribution. *Fishery Bulletin*, 115(1).

Lorenzen, K. (2016). Toward a new paradigm for growth modeling in fisheries stock assessments: embracing plasticity and its consequences. *Fisheries Research*, 180:4–22.

Lorenzen, K. and Enberg, K. (2002). Density-dependent growth as a key mechanism in the regulation of fish populations: evidence from among-population comparisons. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1486):49–54.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. (2013). *The BUGS book: A practical introduction to Bayesian analysis*. CRC press.

Magnusson, A. (2016). *Informative Data and Uncertainty in Fisheries Stock Assessment*. PhD thesis, University of Washington.

Magnusson, A. and Hilborn, R. (2007). What makes fisheries data informative? *Fish and Fisheries*, 8(4):337–358.

Magnusson, A., Punt, A. E., and Hilborn, R. (2013). Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and mcmc. *Fish and Fisheries*, 14(3):325–342.

Martell, S. and Stewart, I. (2014). Towards defining good practices for modeling time-varying selectivity. *Fisheries Research*, 158:84–95.

Maunder, M., Crone, P., Valero, J., and Semmens, B. (2014). Selectivity: theory, estimation, and application in fishery stock assessment models. *Fisheries Research*, (158):1–4.

Maunder, M. N., Crone, P. R., Punt, A. E., Valero, J. L., and Semmens, B. X. (2016). Growth: Theory, estimation, and application in fishery stock assessment models. *Fisheries Research*, 180:1–3.

Maunder, M. N., Crone, P. R., Punt, A. E., Valero, J. L., and Semmens, B. X. (2017). Data conflict and weighting, likelihood functions and process error. *Fisheries Research*, (192):1–4.

Maunder, M. N. and Piner, K. R. (2015). Contemporary fisheries stock assessment: many issues still remain. *ICES Journal of Marine Science: Journal du Conseil*, 72(1):7–18.

Maunder, M. N. and Piner, K. R. (2017). Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. *Fisheries Research*, 192:16–27.

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

McAllister, M. K. and Ianelli, J. N. (1997). Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(2):284–300.

McGarvey, R. and Fowler, A. J. (2002). Seasonal growth of king george whiting (sillaginodes punctata) estimated from length-at-age samples of the legal-size harvest. *Fishery Bulletin*, 100(3):545–558.

Mehl, S. and Sunnanå, K. (1991). Changes in growth of northeast arctic cod in relation to food consumption in 1984-1988. *ICES Journal of Marine Science: Journal du Conseil*, 193:109–112.

Methot, R. D. and Wetzel, C. R. (2013). Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, 142:86–99.

Minte-Vera, C. V., Maunder, M. N., Aires-da Silva, A. M., Satoh, K., and Uosaki, K. (2017). Get the biology right, or use size-composition data at your own risk. *Fisheries Research*, 192:114–125.

Monnahan, C. C., Ono, K., Anderson, S. C., Rudd, M. B., Hicks, A. C., Hurtado-Ferro, F., Johnson, K. F., Kuriyama, P. T., Licandeo, R. R., Stawitz, C. C., et al. (2016). The effect of length bin width on growth estimation in integrated age-structured stock assessments. *Fisheries Research*, 180:103–112.

Morgan, M. J., Koen-Alonso, M., Rideout, R. M., Buren, A. D., Maddock Parsons, D., and editor: Emory Anderson, H. (2018). Growth and condition in relation to the lack of recovery of northern cod. *ICES Journal of Marine Science: Journal du Conseil*, 75(2):631–641.

Myers, R. A. and Worm, B. (2003). Rapid worldwide depletion of predatory fish communities. *Nature*, 423(6937):280–283.

NMFS (2018). Fisheries economics of the united states, 2017. Technical report, U.S. Department of Commerce, NOAA Current Fishery Statistics No. 2017. https://www.fisheries.noaa.gov/content/fisheries-united-states-2017, Accessed January 18, 2019.

NOAA ESRL (2019). ERSL:PSD:COBE SST. data retrieved from https://www.esrl.noaa.gov/psd/data/gridded/data.cobe.html – accessed Feb. 11, 2019.

Ono, K., Licandeo, R., Muradian, M. L., Cunningham, C. J., Anderson, S. C., Hurtado-Ferro, F., Johnson, K. F., McGilliard, C. R., Monnahan, C. C., Szuwalski, C. S., et al. (2015). The importance of length and age composition data in statistical age-structured models for marine species. *ICES Journal of Marine Science: Journal du Conseil*, 72(1):31–43.

Pardo, S. A., Cooper, A. B., and Dulvy, N. K. (2013). Avoiding fishy growth curves. *Methods in Ecology and Evolution*, 4(4):353–360.

Pauly, D. (1980). On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. *ICES Journal of Marine Science: Journal du Conseil*, 39(2):175–192.

Pella, J. J. and Tomlinson, P. K. (1969). A generalized stock production model. *Inter-American Tropical Tuna Commission Bulletin*, 13(3):416–497.

Perdikaris, S. and McGovern, T. H. (2009). Viking age economics and the origins of commercial cod fisheries in the north atlantic. In Sicking, L. and Abreu-Ferreira, D., editors, *Beyond the catch. Fisheries of the North Atlantic, the North Sea and the Baltic, 900–1850*, chapter 2, pages 61–90. Brill, Leiden, Netherlands.

Pienaar, L. V. and Turnbull, K. J. (1973). The chapman-richards generalization of von bertalanffy's growth model for basal area growth and yield in even-aged stands. *Forest Science*, 19(1):2–22.

Plagányi, É. E. (2007). *Models for an ecosystem approach to fisheries*. Number 477. Food & Agriculture Org.

Prince, J. and Hordyk, A. (2019). What to do when you have almost nothing: A simple quantitative prescription for managing extremely data-poor fisheries.

Punt, A. E. (2017). Some insights into data weighting in integrated stock assessments. *Fisheries Research*, 192:52–65.

Punt, A. E., Akselrud, C. A., and Cronin-Fine, L. (2017). The effects of applying mis-specified age-and size-structured models. *Fisheries Research*, 188:58–73.

Punt, A. E. and Hilborn, R. (1997). Fisheries stock assessment and decision analysis:

the bayesian approach. *Reviews in Fish Biology and Fisheries*, 7(1):35–63.

Punt, A. E., Huang, T., and Maunder, M. N. (2013). Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. *ICES Journal of Marine Science: Journal du Conseil*, 70(1):16–33.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ralston, S. and Ianelli, J. (1998). When lengths are better than ages: The complex case of bocaccio. In *Int. Symp. on Fishery Stock Assessment Models for the 21st Century, Anchorage, Alaska, EEUU. 8-11 Oct 1997.*

Ralston, S., Punt, A. E., Hamel, O. S., DeVore, J. D., and Conser, R. J. (2011). A meta-analytic approach to quantifying scientific uncertainty in stock assessments. *Fishery Bulletin*, 109(2):217–232.

Ratkowsky, D. (1986). Statistical properties of alternative parameterizations of the von bertalanffy growth curve. *Canadian Journal of Fisheries and Aquatic Sciences*, 43(4):742–747.

Reeves, S. A. (2003). A simulation study of the implications of age-reading errors for stock assessment and management advice. *ICES Journal of Marine Science: Journal du Conseil*, 60(2):314–328.

Rice, J. (2011). Managing fisheries well: delivering the promises of an ecosystem approach. *Fish and Fisheries*, 12(2):209–231.

Ricker, W. (1969). Effects of size-selective mortality and sampling bias on estimates of growth, mortality, production, and yield. *Journal of the Fisheries Board of Canada*, 26(3):479–541.

Ricker, W. E. (1954). Stock and recruitment. *Journal of the Fisheries Board of Canada*, 11(5):559–623.

Ricklefs, R. E. (1968). Patterns of growth in birds. *Ibis*, 110(4):419–451.

Rochet, M.-J. and Rice, J. C. (2009). Simulation-based management strategy evaluation: ignorance disguised as mathematics? *ICES Journal of Marine Science: Journal du*

Conseil, 66(4):754–762.

Roff, D. A. (1984). The evolution of life history parameters in teleosts. *Canadian Journal of Fisheries and Aquatic Sciences*, 41(6):989–1000.

Rose, G. and O'driscoll, R. (2002). Capelin are good for cod: can the northern stock rebuild without them? *ICES Journal of marine Science*, 59(5):1018–1026.

Royston, P. (2007). Profile likelihood for estimation and confidence intervals. *Stata*, 7(3):376–387.

Ryer, C. H. (2008). A review of flatfish behavior relative to trawls. *Fisheries Research*, 90(1-3):138–146.

Sainsbury, K. (1980). Effect of individual variability on the von bertalanffy growth equation. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(2):241–247.

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

Sampson, D. B. and Scott, R. D. (2011). A spatial model for fishery age-selection at the population level. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(6):1077–1086.

Sampson, D. B. and Scott, R. D. (2012). An exploration of the shapes and stability of population–selection curves. *Fish and Fisheries*, 13(1):89–104.

Schnute, J. (1981). A versatile growth model with statistically stable parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, 38(9):1128–1140.

Schueller, A. M., Williams, E. H., and Cheshire, R. T. (2014). A proposed, tested, and applied adjustment to account for bias in growth parameter estimates due to selectivity. *Fisheries Research*, 158:26–39.

Seidov, D., Mishonov, A., Reagan, J., and Parsons, R. (2017). Multidecadal variability and climate shift in the north atlantic ocean. *Geophysical Research Letters*.

Seung, C. K. and Waters, E. C. (2006). A review of regional economic models for fisheries management in the us. *Marine Resource Economics*, 21(1):101–124.

Shin, Y.-J. and Rochet, M.-J. (1998). A model for the phenotypic plasticity of north sea herring growth in relation to trophic conditions. *Aquatic Living Resources*, 11(5):315–324.

Shine, R. and Charnov, E. L. (1992). Patterns of survival, growth, and maturation in snakes and lizards. *American Naturalist*, 139(6):1257–1269.

Sigfússon, Þ., Gestsson, H. M., et al. (2012). Iceland´ s ocean economy, the economic impact and performance of the ocean cluster in 2011. Technical report, University of Iceland. http://www.sjavarklasinn.is/wp-content/uploads/2012/12/IcelandsOceanEconomy2011.pdf, Accessed January 18, 2019.

Sinclair, A. and Swain, D. (1996). Comment: Spatial implications of a temperature-based growth model for atlantic cod (gadus morhua) off the eastern coast of canada. *Canadian Journal of Fisheries and Aquatic Sciences*, 53(12):2909–2911.

Sinclair, A., Swain, D., and Hanson, J. (2002a). Measuring changes in the direction and magnitude of size-selective mortality in a commercial fish population. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(2):361–371.

Sinclair, A. F., Swain, D. P., and Hanson, J. M. (2002b). Disentangling the effects of size-selective mortality, density, and temperature on length-at-age. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(2):372–382.

Sippel, T., Lee, H. H., Piner, K., and Teo, S. L. (2017). Searching for m: Is there more information about natural mortality in stock assessments than we realize? *Fisheries Research*, 192:135–140.

Somerton, D. A., Williams, K., von Szalay, P. G., and Rose, C. S. (2011). Using acoustics to estimate the fish-length selectivity of trawl mesh. *ICES Journal of Marine Science: Journal du Conseil*, 68(7):1558–1565.

Sparre, P. and Venema, S. C. (1998). *Introduction to tropical fish stock assessment-Part 1: Manual*. FAO.

Stawitz, C. C., Haltuch, M. A., and Johnson, K. F. (2019). How does growth misspecification affect management advice derived from an integrated fisheries stock assessment model? 213:12–21.

Stefansson, G. (2003). Issues in multispecies models. *Natural Resource Modeling*, 16(4):415–437.

Stefánsson, G., Pálsson, Ó., et al. (1997). Bormicon: a boreal migration and consumption model. Technical report.

Stewart, I. J. and Monnahan, C. C. (2017). Implications of process error in selectivity for approaches to weighting compositional data in fisheries stock assessments. *Fisheries Research*, 192:126–134.

Stokes, M. (1996). Larval settlement, post-settlement growth and secondary production of the florida lancelet (= amphioxus) branchiostoma floridae. *Marine Ecology Progress Series*, pages 71–84.

Su, Y.-S. and Yajima, M. (2015). *R2jags: Using R to Run 'JAGS'*. R package version 0.5-7.

Swain, D. P., Sinclair, A. F., Castonguay, M., Chouinard, G., Drinkwater, K. F., Fanning, L., and Clark, D. S. (2003). Density-versus temperature-dependent growth of atlantic cod (gadus morhua) in the gulf of st. lawrence and on the scotian shelf. *Fisheries Research*, 59(3):327–341.

Szuwalski, C. (2016). Biases in biomass estimates: the effect of bin width in size-structured stock assessment methods. *Fisheries Research*, 180:169–176.

Tang, M., Jiao, Y., and Jones, J. W. (2014). A hierarchical bayesian approach for estimating freshwater mussel growth based on tag-recapture data. *Fisheries Research*, 149:24–32.

Taylor, I. G. and Methot Jr, R. D. (2013). Hiding or dead? a computationally efficient model of selective fisheries mortality. *Fisheries Research*, 142:75–85.

Taylor, L., Begley, J., Kupca, V., and Stefansson, G. (2007). A simple implementation of the statistical modelling framework Gadget for cod in Icelandic waters. *African Journal of Marine Science*, 29(2):223–245.

Taylor, L. and Stefansson, G. (2004). Gadget models of cod-capelin-shrimp interactions in icelandic waters. *ICES CM*.

Taylor, N. G., Walters, C. J., and Martell, S. J. (2005). A new likelihood for simultaneously estimating von bertalanffy growth parameters, gear selectivity, and natural and fishing mortality. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(1):215–223.

Thorson, J. T. and Minte-Vera, C. V. (2016). Relative magnitude of cohort, age, and year effects on size at age of exploited marine fishes. *Fisheries Research*, 180:45–53.

Thorson, J. T. and Simpfendorfer, C. A. (2009). Gear selectivity and sample size effects on growth curve selection in shark age and growth studies. *Fisheries Research*, 98(1-3):75–84.

Thorson, J. T. and Taylor, I. G. (2014). A comparison of parametric, semi-parametric, and non-parametric approaches to selectivity in age-structured assessment models. *Fisheries Research*, 158:74–83.

Troynikov, V. S. (1999). Use of bayes theorem to correct size-specific sampling bias in growth data. *Bulletin of Mathematical Biology*, 61(2):355–363.

Troynikov, V. S. and Koopman, M. T. (2009). The effect of danish seine selectivity and retention on growth estimates of tiger flathead platycephalus richardsoni. *Fisheries Science*, 75(4):833–838.

Vilhjálmsson, H. (2002). Capelin (*Mallotus villosus*) in the Iceland–East Greenland–Jan Mayen ecosystem. *ICES Journal of Marine Science: Journal du Conseil*, 59(5):870–883.

von Bertalanffy, L. (1938). A quantitative theory of organic growth (inquiries on growth laws. ii). 10(2):181–213.

Wallis, P. and Flaaten, O. (2001). Fisheries management costs: Concepts and studies. In *Microbehavior and Macroresults*. International Institute of Fisheries Economics and Trade.

Wang, S.-P. and Maunder, M. N. (2017). Is down-weighting composition data adequate for dealing with model misspecification, or do we need to fix the model? *Fisheries Research*, 192:41–51.

Weatherley, A. (1990). Approaches to understanding fish growth. *Transactions of the American Fisheries Society*, 119(4):662–672.

Webber, D. N. and Thorson, J. T. (2016). Variation in growth among individuals and over time: A case study and simulation experiment involving tagged antarctic toothfish. *Fisheries Research*, 180:67–76.

Wetzel, C. R. and Punt, A. E. (2011). Performance of a fisheries catch-at-age model (stock synthesis) in data-limited situations. *Marine and Freshwater Research*, 62(8):927–936.

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

Wijffels, S., Roemmich, D., Monselesan, D., Church, J., and Gilson, J. (2016). Ocean temperatures chronicle the ongoing warming of earth. *Nature Climate Change*, 6(2):116.

Wilson, K. L., Matthias, B. G., Barbour, A. B., Ahrens, R. N., Tuten, T., and Allen, M. S. (2015). Combining samples from multiple gears helps to avoid fishy growth curves. *North American Journal of Fisheries Management*, 35(6):1121–1131.

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., Jackson, J. B., Lotze, H. K., Micheli, F., Palumbi, S. R., et al. (2006). Impacts of biodiversity loss on ocean ecosystem services. *Science*, 314(5800):787–790.

Yaragina, N. and Marshall, C. (2000). Trophic influences on interannual and seasonal variation in the liver condition index of northeast arctic cod (gadus morhua). *ICES Journal of Marine Science*, 57(1):42–55.

Zhang, Z., Lessard, J., and Campbell, A. (2009). Use of bayesian hierarchical models to estimate northern abalone, haliotis kamtschatkana, growth parameters from tag-recapture data. *Fisheries Research*, 95(2-3):289–295.