Modelling the Growth of Northern Cod

by

Zakiyah Mohammed

A practicum submitted to the School of Graduate Studies

in partial fulfillment of the requirement for the degree of

Master of Applied Statistics

Department of Mathematics and Statistics

Memorial University

St John's Newfoundland, Canada December 2015

Abstract

We develop a body size growth model of Northern cod (*Gadus morhua*) in Northwest Atlantic Fisheries Organization (NAFO) Divisions 2J3KL during 2009-2013. We use individual length-at-age data from the bottom trawl survey in these divisions during 2009–2013. We use the Von Bertalanffy (VonB) model extended to account for between-individual variations in growth, and variations that may be caused by the methods which fish are caught and sampled for length and age measurements. We assume between-individual variation in growth appears because individuals grow at a different rate (k), and they achieve different maximum sizes (l_∞). We also included measurement error in length and age in our model since ignoring these errors can lead to biased estimates of the growth parameters. We use the structural errors-in-variables (SEV) approach to estimate individual variation in growth, ageing error variation, and the true age distribution of the fish. Our results shows the existence of individual variation in growth and ME in age. According to the negative log likelihood ratio (NLLR) test, the best model indicated: 1) different growth patterns across divisions and years. 2) Between individual variation in growth is the same for the same division across years. 3) The ME in age and true age distribution are different for each year and division.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr Noel Cadigan, for his patience, invaluable assistance, constructive criticisms and commitment which brought my practicum to a successful completion. I gratefully acknowledge the funding received towards my MAS from School of Graduate Studies, the Department of Mathematics & Statistics , and my supervisor in the form of graduate fellowships and teaching assistantships.

I also appreciate the assistance I received from the faculty members and staff of our department over the last two years, I am very grateful. My warmest gratitude goes to my family for their love, encouragement and constant support. I extend my sincere thanks to Nadarajah Tharshanna and Patricia Judd for reading through my work. Finally, I would like to thank Dr, Joanne Morgan of the Department of Fisheries Oceans (DFO) Canada in St. John's NL for providing data for the analysis and some comments on earlier versions of some chapters.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.0.1 Fish growth

The study of fish growth has become increasingly important. This is because the sizes of fish caught by fishers (for recreational and commercial purposes) have a great impact on the population dynamics of the fish stock, the potential yield that the stock can sustain (Alos et al., 2010), and the ocean ecosystems. Additionally, fishery managers rely on fish growth models to predict fishery trends, in order to set fishing quotas, size limits, and gear restrictions to maintain a healthy and sustainable fish population and fishing industry while still preserving ocean ecosystems. Thus, the knowledge of fish growth is necessary for stock assessments and for development of management or conservation plans.

Growth models are often fitted to data obtained by sampling fish and measuring their length and age. The length of a fish is obtained by laying the fish straight along a tape measure on a horizontal surface. In temperate areas, the age of a fish is determined by counting growth rings on the otolith (fish ear bone). Rings are formed due to strong fluctuations in environmental conditions from summer to winter and vice versa (Sparre and Venema, 1998). Fish growth may be measured in terms of the length or weight of the fish; in this study we measure fish growth in terms of length.

Various models have been used for the growth of fish. The theory has been reviewed by a number of researchers, such as Beverton and Holt (1957), Ursin (1968), Ricker (1975), Gulland (1983), Pauly (1984), and Pauly & Morgan (1987). However, the growth model developed by Von Bertalanffy (VonB) is the most commonly used model in fisheries science (Quist et al., 2012) because it conforms to the observed growth of many fish species. The theory of the VonB model is based on the assumption that the change in length per unit time ($dy(t)/dt$) declines with size. If $y(t)$ denotes the length at time (t), then the growth rate model is

$$\frac{dy(t)}{dt} = k\{l_\infty - y(t)\}, \tag{1.1}$$

and $dy(t)/dt = 0$ when $y(t) = l_\infty$. Thus, the growth rate of fish will get smaller and eventually becomes zero as a fish nears its maximum possible size (l_∞). The parameter l_∞ is the asymptotic length at which the growth rate is zero and k is the growth rate parameter. Assuming that $y(t) = 0$ when $t = 0$, the solution of Eq. 1.1 is

$$y(t) = l_\infty \left(1 - e^{-kt}\right). \tag{1.2}$$

We illustrate this model in Figure 1.1 (top panel) when $k = 0.15$ and $l_\infty = 120$. Generally the length of the fish during its first year (age zero) is not zero; i.e., $y(t) > 0$ at $t = 0$. To account for this, we use the following form of the VonB growth model,

$$y(t) = l_\infty \left(1 - e^{-k(t-t_o)}\right), \tag{1.3}$$

where $t_o(< 0)$ is the theoretical age at which the fish would have had zero length. In practical terms age cannot be negative, but if $y(t) > 0$ at age $t = 0$ and we extrapolate the growth curve back to when $y(t) = 0$, we obtain a negative age (see Figure 1.2). Alternatively, if $y(0)$ is the length of the fish at time $t = 0$ such that $\rho_0 = y(0)/l_\infty$, then the length at time t is

$$\begin{aligned} y(t) &= l_\infty - l_\infty (1 - \rho_0) e^{-kt} \\ &= l_\infty \left(1 - [1 - \rho_0] e^{-kt}\right) \end{aligned} \tag{1.4}$$

In this case, at time $t = 0$ the length of the fish is

$$\begin{aligned} y(0) &= l_\infty - l_\infty \left(1 - \frac{y(0)}{l_\infty}\right) e^{-k*0} \\ &= l_\infty - [l_\infty - y(0)] \\ &= y(0). \end{aligned} \tag{1.5}$$

Thus $e^{kt_o} = 1 - \rho_0$ (compare Eq. 1.3 and 1.4). Hence $t_o = log(1 - \rho_0)/k$. This is another interpretation of ρ_0. We compare the VonB models (Eq. 1.2 and 1.3) in Figure 1.1 when $k = 0.15$ and $l_\infty = 120$cm.

The VonB model (Eqn. 1.3) is used to describe the mean growth of a population where l_∞, k and t_o are the population mean growth parameters.

Remark 1. We use upper case notation to denote random variables. For example K_i is the random value for the i^{th} fish.

The VonB model (Eqn. 1.3) assumes that every individual randomly selected from the population has the same growth parameters (l_∞ and k). However, growth parameters differ from species to species, stock to stock within the same species, and for individuals within the same stock. That is, the i^{th} individual in the population grows according to the VonB model with growth parameters $L_{\infty i}$ and K_i. Nonetheless some authors (e.g. Beverton & Holt, 1957) overlook this individual variation in growth and fit the VonB (Eqn. 1.3) to individual length and age data. In such a case, they are describing individual growth through the population mean growth parameters (Alos et al., 2010). This leads to bias in estimating the population mean growth parameters and length at age as noted by Sainsbury (1980) and Wang et al. (1995).

We illustrate this bias when l_∞ and t_o are constant parameters and there is only between-individual variation in K. We denote the mean and variance of K as k and σ_k^2, respectively. We assume $K \sim Gamma(\alpha, \beta)$, where $\alpha = k^2/\sigma_k^2 = (CV)^{-2}$, $\beta = \sigma_k^2/k = (CV)^2 k$, and CV is the coefficient of variation. The random length Y of a fish at age t is

$$Y(t) = l_\infty \left(1 - e^{-K(t-t_o)}\right). \tag{1.6}$$

The expected length is given by

$$\begin{aligned} E(Y(t)|t) &= l_\infty \left\{1 - E[e^{-K(t-t_o)}]\right\} \\ &= l_\infty \left\{1 - [1 + \beta(t - t_o)]^{-\alpha}\right\} \\ &= l_\infty \left\{1 - [1 + (CV)^2 k(t - t_o)]^{-(CV)^{-2}}\right\}. \end{aligned} \tag{1.7}$$

The expected length when variability in K is included in the VonB model (Eqn. 1.7) is not equal to the length that the VonB model (Eqn. 1.3) predicts. This causes the bias in the growth parameters when individual variation in growth is ignored. Figure 1.3 shows the effect of variability in K on the expected length of a fish, where the CV of K is 0.08, 0.3 and 0.5. The lengths at ages estimated from Eqn. 1.7 were smaller than those from Eqn. 1.3 (compare VonB with CV=0.3 and CV=0.5 curve in Figure 1.3). The effect of fitting Eqn. 1.3 to individual length and age data is to underestimate the population mean growth rate (k) and length at age. To see this, note that Eqn. 1.7 $\Longrightarrow$ Eqn. 1.3 as CV $\Longrightarrow 0$, but Eqn. 1.7 becomes smaller progressively as the CV of K increases. Thus the magnitude of underestimation of k depend on the CV of K.

Sainsbury (1980) used simulations to examine the effect of individual variability on the length at age estimates when the VonB model is not extended to account for this variability. He found that when individual data are analyzed using Eqn. 1.3 (thus, ignoring variation K), the estimate of k obtained will be an underestimate. Shackell et al. (1997) also found large variation in the individual growth of cod and advised that when applying the VonB model to individual growth data, the growth model should be extended to include individual variation.

1.0.2 Variation in fish growth.

Individual variations in the growth of fish can be attributed to the following:

- between-individual variation (BT),
- within-individual (WI) and

- a mixture of these.

There is also measurement error (ME), which mostly occurs in age reading but also in length measurements. Between-individual variation in growth can have substantial evolutionary consequences; that is, size selective fishing mortality can change productivity dynamics of the stock by giving slow growers an evolutionary selective advantage. ME has no evolutionary consequence and is not clear about within-individual variation.

Although ME has no evolutionary consequence, it can lead to biased estimates of the growth parameters (Heifetz et al., 1998; Kimura, 1990) if the ME is ignored. It is well known that ME in covariates (age in this case) may produce a bias in parameter estimation for statistical models (Carroll et al., 2006). Thus, not accounting for ME in age may result in biased estimates of VonB parameters, and the effect of this will lead to under-estimating l_∞ and over-estimating k (see Chapter 3). In this case the maximum sustainable yield harvest rate will seem higher than it is when ME bias is absent because the growth rate of the population will be estimated to be higher than it is, and the maximum size will be estimated to be lower. Hence, the productivity of the stock will be over-estimated when fishery exploitation rates are high and under-estimated when exploitation rates are low. As such, not accounting for this bias when it is large may result in fisheries harvest advice that is sub-optimal and possibly unsustainable. In this study, we only focus on BT and ME in length and age variations.

The VonB model we described earlier (Eqn. 1.3) assumes that all individuals in the population have the same growth parameters and does not account for the variability among individuals. As a result, different studies have been carried out to extend this model to account for individual variation in growth using different approaches. Sainsbury (1980) extended the VonB model to account for individual variation in growth by varying the VonB growth parameters (L_∞ and K) among individuals. He assumed that:

- L_∞ is normally distributed with mean l_∞ and variance $\sigma^2_{l_\infty}$.
- K follows gamma distribution with parameters α and β (mean $k = \alpha/\beta$ and variance $\sigma^2_k = \alpha/\beta^2$).

Some authors including Shelton & Mangel (2012) argue that a better way to formulate the VonB model is not to use the growth parameter L_∞ but use a different constant parameter w such that $L_\infty = w/K$. They varied only K to account for individual variation in growth. Note that in this approach the growth parameters K and L_∞ are not independent and they are negatively correlated. That is, a slow grower will have a higher L_∞ and vice versa. However, in practice this may not be the case.

Faster growing fish may have a higher L_∞. Deciding if this approach is the best way to model individual variability in growth is beyond the scope of this research.

In this practicum we assume between-individual variation in growth appears because individuals grow at a different rate (K), and they achieve different maximum sizes (L_∞). We assume that L_∞ and K are independent (i.e., corr $(L_\infty, K) = 0$). We use a similar approach to Sainsbury (1980) to extend the VonB to incorporate individual variation in fish growth. However, we use a lognormal distribution to model the random variation in L_∞ and K. Thus, for $log(L_{\infty i}) \sim N(log(l_\infty), \sigma^2_\infty)$ and $log(K_i) \sim N(log(k), \sigma^2_k)$, the length of the ith individual is given by

$$Y_i(t) = L_{\infty i}\{1 - e^{-K_i(t-t_o)}\}, \tag{1.8}$$

where σ_∞ and σ_k describe the BT variations.

While within-individual (WI) variation in growth is not discussed in this study, Filipe et al. (2010) addressed this using a stochastic differential equation model (SDE). They described the growth of an individual in a random environment as $dY_t = \beta(\alpha - Y_t)dt + \sigma dW_t$, where Y_t is the size of the individual at time (t), α is the average asymptotic size; and $\beta(> 0)$ is the relative growth rate of Y_t. The parameter σ describes WI variability, and W_t is the standard Wiener process. They emphasized that the SDE models are quite appropriate for including the effect of random environmental fluctuations that affect individual growth rate.

In this practicum we develop a model for estimating growth of northern cod in NAFO Divisions 2J and 3KL (see Figure 1.4) during 2009-2013. In developing this model, we extend the VonB to include ME in length and age, take into account BT variation; and variations that may be caused by the way fish are caught and sampled for length and age measurements (Goodyear, 1995; Candy, 2005). We present this project in six chapters. It is important to account for ME in length and age so that these sources of variation are not included in BT variance parameter estimates. It is also important to account for the sampling designs used to collect growth data because ignoring this can lead to biased growth parameter estimates, as we will show in this report.

In Chapter 2, we extend the VonB model to include ME in length and take into account BT variation. We first study the commonly used maximum likelihood estimator (MLE) that assumes only ME in length (MEL) and does not include BT variation in L_∞ and K. The second estimation procedure we study includes ME in length and BT variation in VonB parameters (MELBT; see Chapter 2 for model equations). We use simulation studies to examine the performance of the two estimators (MLE of MEL and MELBT) in estimating the VonB growth parameters, and in separating the ME in length and BT variation (i.e., estimate the standard deviation of ME in length, L_∞, and K). We use the Template Model Builder (TMB) package in R (https://github.com/kaskr/adcomp; Kristensen et al., 2015) for the MLE of MELBT. TMB is a new package for fast fitting of complex statistical models that may include random effects. Thus, the MLE of MEL estimates the variance of ME in length, and the mean parameters l_∞, and k, while the MLE of MELBT estimates the mean parameters l_∞, and k, and the variance of L_∞, K, and ME in length. Unfortunately but not surprisingly, the three sources of variation are somewhat confounded and difficult to estimate reliably. We address this problem in Chapter 3.

In Chapter 3, we extend the growth model in Chapter 2 to include ME in age. The ME variance in age is confounded with the ME variance in length and variation in L_∞ and K. However, for the data we examine it is reasonable to assume that the ME errors in length are small because length measurements can be reliably obtained by trained specialists. Measuring the lengths of fish is much less prone to error than measuring the age - which involves counting growth rings in small ear bones using a microscope. Hence, we assume the ME variance in length is a small but known value. The ageing error is modelled such that the coefficient of variation (CV) is a linear function of the true age of the fish. This is because as fish get older, rings on their otolith tend to shrink, which makes them more difficult to count; thus, ageing errors are larger for older fish but not necessarily with a constant CV.

This source of variation is still confounded with between-individual variation in growth rates. We address this confounding using a spatial model and assuming that there is no between-individual variation in L_∞ and K for fish caught close together in space. Variation in growth for such fish is attributed to ME in length and age. The rationale for this assumption is that fish from the same spawning schools are genetically more similar and have experienced and grown in more similar physical environments than fish in different spawning schools. We recognize that this assumption will not be exactly met in practise, however we suggest that this is a pragmatic approach to infer ME in age when there is no age-validation data available, as is the case for the cod data we analyze. For simplicity we use a spatially stratified model and assume that σ_∞ and σ_k are negligible for fish caught in the same year and stratum.

We study two ME models described by Carroll et al. (2006): the functional errors-in-variables (FEV) and the structural errors-in-variables (SEV) models. The FEV model considers the true age of the ith individual to be an unknown parameter, and the SEV approach considers the true age of the ith individual to be a random variable with some distribution (see Chapter 3 for model equations). We use simulation studies and the TMB package in R to examine the performance of the two estimators (FEV and SEV) in estimating the VonB growth parameters and separating the ME in age and the BT variation. The SEV approach estimates the mean parameters l_∞, k, the variance of L_∞, and K, the ME variance in age and the true age distribution. The FEV approach estimates the mean parameters l_∞, k, and the standard deviations for L_∞, K, and the unknown ages that are measured with error for each fish in the sample.

In Chapter 4, we extend the growth model in Chapter 3 to account for variations that may be caused by the methods which fish are caught and sampled for length and age measurements. In Chapters 2 and 3, we assumed that fish are randomly sampled from the population, even though in practice fish are caught using sampling trawls that target and capture fish by size and species during the fishing operation. Thus, all sampling trawls are size and species-selective to varying degrees (Walsh, 1996). In the data we examine, all fish in research survey trawl catches were first measured for length and then classified based on length classes (3cm length class). A fixed number of fish were selected from the length classes for age determination (i.e., size-stratified age-sampling). The second sample provides the length-at-age data that is used to fit the VonB growth model. This is a type of size-biased (aka length-biased) sampling design (Hu, 2013). Thus, in Chapter 4, the model in Chapter 3 is extended to take into account the size selectivity of the fishing gear used to catch the fish, and size-stratified age sampling. We use simulation studies to examine the performance of the SEV approach in estimating the mean parameters l_∞, k, the variance of L_∞, and K, the ME variance in age and the true age distribution.

In Chapter 5, we apply the growth model developed in Chapter 4 to the length-at-age cod data collected in Northwest Atlantic Fisheries Organization (NAFO) divisions 2J, 3K, and 3L (see Figure 1.4) during 2009 to 2013. Finally, Chapter 6 will conclude with a summary and a discussion of all our results.

1.1 Figures

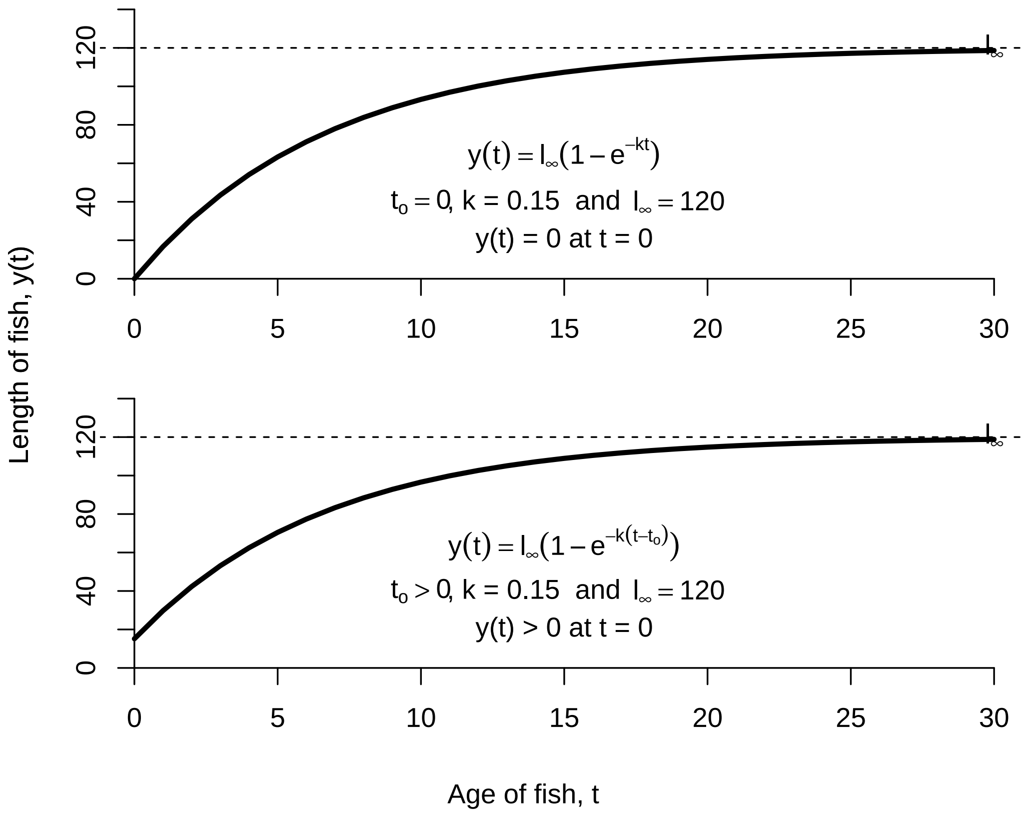

Figure 1.1: Application of the von Bertalanffy (VonB) model for $l_\infty = 120$cm, $k = 0.15$, which assumes that fish grow most quickly when they are young. Growth slows gradually as the individual fish ages, and eventually stops growing at length, $l_\infty = 120$cm. The first growth curve is for when $t_o = 0$ and second is for when $t_o > 0$.

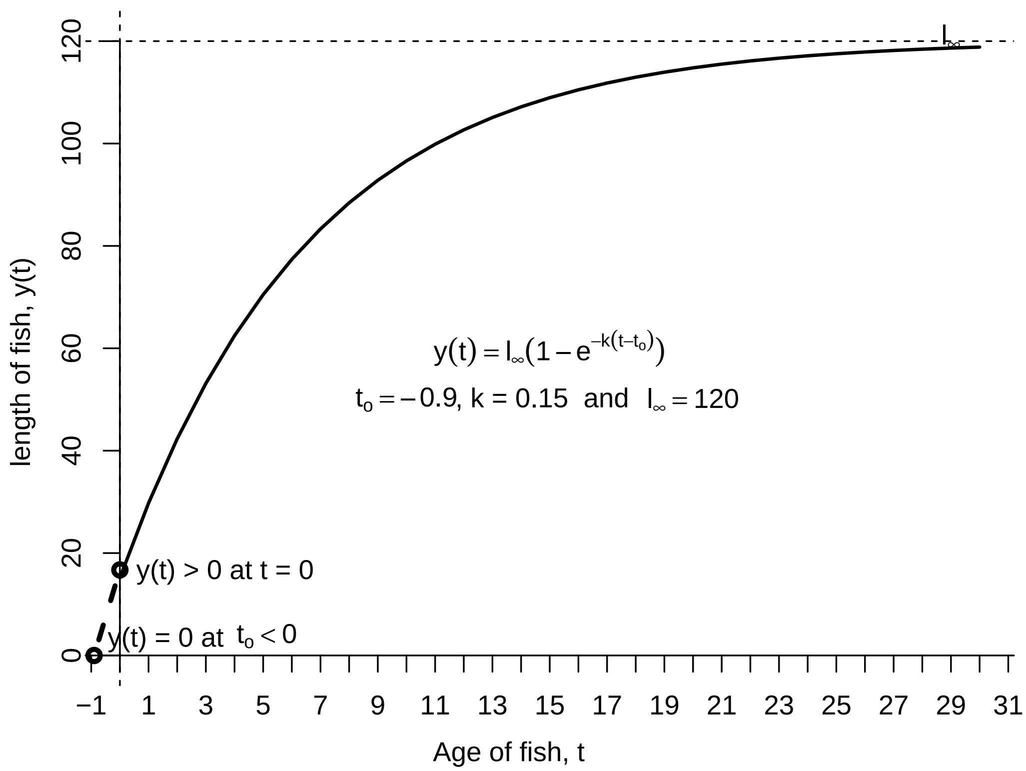

Figure 1.2: A case where the length ($y(t)$) of the fish at birth ($t = 0$) is greater than zero, i.e., $y(t) > 0$ at $t = 0$. If we extrapolate the curve back, the age ($t = t_o$) at which the fish have a length ($y(t)$) equal to 0 is negative.

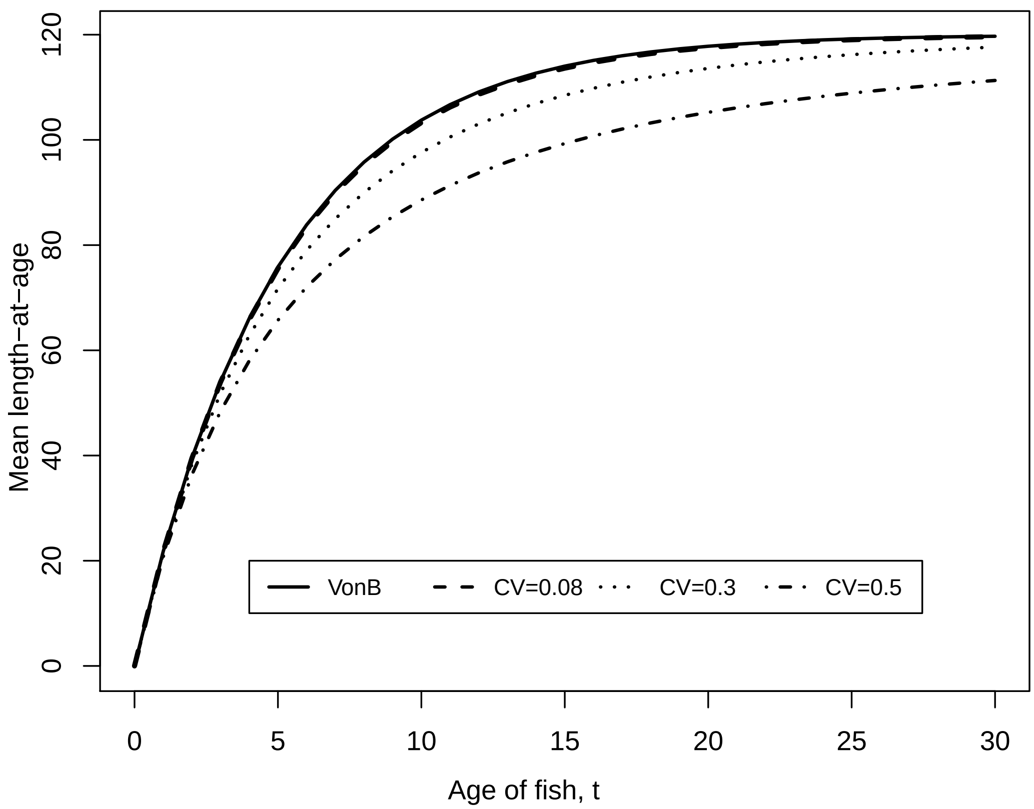

Figure 1.3: The effects of variability in K upon the mean of length at given age. von Bertalanffy (VonB) model with variation in K, $l_\infty = 120$cm , $k = 0.3$: (a) Coefficient of variation $CV = \sigma_k/k = 0.05$. (b) $CV = 0.3$. (c) $CV = 0.5$. The difference between the VonB (solid line) and the VonB with variation in K (- -, ... , and .-.- lines) mean length at age increases as the CV of K increases.

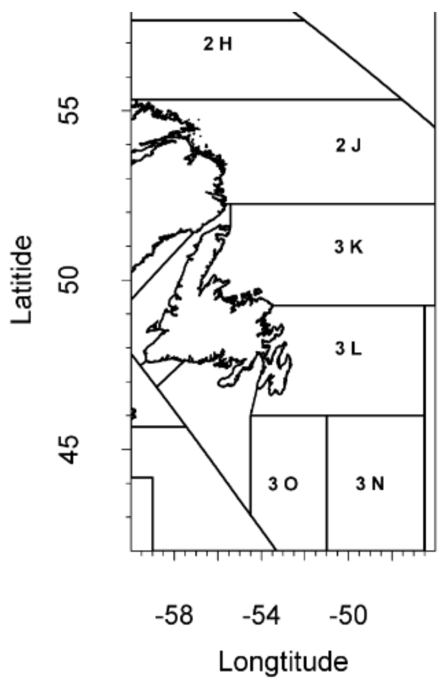

Figure 1.4: Map of Northwest Atlantic Fisheries Organization (NAFO): divisions 2J, 3K, and 3L.

Chapter 2

Von Bertalanffy (VonB) model with measurement error (ME) in length.

2.0.1 Growth model

In the previous chapter we discussed that individual fish do not grow at the same rate. There is individual variation which is not accounted for in the commonly used VonB model (Eqn. 1.3). Sainsbury (1980), Smith et al. (1997), and others have expressed concern about estimation using Eqn. 1.3 with individual length and age data and the effect of individual variation in the VonB parameters. Additionally, measurement error (ME) in length is not included in the VonB model (Eqn. 1.3).

In this Chapter, we extend the VonB model to include ME in length and take into account individual variability in growth. We account for individual variability in growth by assuming that each fish has its own VonB growth curve with parameters $L_{\infty i}$ and K_i, where $L_{\infty i}$ and K_i are modelled as

$$log(L_{\infty i}) \sim N(log(l_\infty), \sigma_\infty^2) \tag{2.1}$$

and

$$log(K_i) \sim N(log(k), \sigma_k^2). \tag{2.2}$$

Thus we expressed the VonB model that accounts for BT variation as

$$Y_i(t) = L_{\infty i}\{1 - e^{-K_i(t-t_o)}\}, \tag{2.3}$$

where σ_∞ and σ_k describe the BT variations.

The growth model with only ME in length is

$$y(t) = l_\infty \left(1 - e^{-k(t-t_o)}\right) + \varepsilon_t \tag{2.4}$$

where $\varepsilon_t \sim logN(0, \sigma_\varepsilon^2)$ and σ_ε describes the ME in length variation. We use additive ME in length, so that the errors are independent of the size of the fish, which is reasonable in practise.

The growth model with ME in length that also accounts for BT variability is

$$Y(t) = L_{\infty i}\left(1 - e^{-K_i(t-t_o)}\right) + \varepsilon_t. \tag{2.5}$$

We use simulations to study two methods of estimation. The first method is the commonly used maximum likelihood estimator (MLE) that assumes only ME in length and does not include BT variation in VonB parameters (L_∞ and K). The second method includes ME in length and BT variation in L_∞ and K. Our intent is to find the best estimator of the VonB parameters, and the BT and ME in length variation parameters.

2.1 Estimation Method 1: Maximum likelihood estimator (MLE) including only ME in length (MEL).

The probability density function (pdf) of the i^{th} length Y_i in the sample (i = 1,...., n) given the growth parameters l_∞, k, t_o and the ME in length variance parameter σ_ε is

$$f(Y_i = y_i|l_\infty, k, t_o, \sigma_\varepsilon) = \frac{1}{\sigma_\varepsilon\sqrt{2\pi}} \exp\left[-\frac{\{y_i - y(t_i)\}^2}{2\sigma_\varepsilon^2}\right], \tag{2.6}$$

where

$$y(t_i) = l_\infty \left(1 - e^{-k(t_i - t_o)}\right),$$

$\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and $y_i = y(t_i) + \varepsilon$ is the observation of $y(t)$ measured with error at $t = t_i$. The likelihood for $l_\infty, k, \sigma_\varepsilon, t_o$ for the entire sample is

$$L(l_\infty, k, t_o, \sigma_\varepsilon | y_1, ..., y_n) = \prod_i^n \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left[-\frac{(y_i - y(t_i))^2}{2\sigma_\varepsilon^2}\right], \tag{2.7}$$

The only variation included in this model is the ME in length; this model does not include BT variation in L_∞ and K. This model is perfectly specified in data with only ME in length variation; it is totally misspecified in data:

- With only BT source of variability.
- With BT and ME in length variations.

2.2 Estimation Method 2: Marginal maximum likelihood estimator (MLE) including ME in length and BT variation (MELBT).

The conditional pdf of the i^{th} random length Y_i in the sample $(i = 1, ..., n)$, given the random effects $L_{\infty i}$ and K_i is

$$f(Y_i = y_i \mid L_{\infty i}, K_i, t_o, \sigma_\varepsilon) = \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left[- \frac{\{y_i - y(t_i)\}^2}{2\sigma_\varepsilon^2} \right], \tag{2.8}$$

where

$$y(t_i) = L_{\infty i} \left(1 - e^{-K_i(t_i - t_o)}\right),$$

and $y_i = y(t_i) + \varepsilon$ is the observation of $y(t)$ measured with error at $t = t_i$, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

The random effect $L_{\infty i}$ is log-normally distributed with mean $l_\infty e^{\frac{\sigma_\infty^2}{2}}$ and variance $l_\infty^2 e^{\sigma_\infty^2}(e^{\sigma_\infty^2} - 1)$. The random effect K_i is log-normally distributed with mean $ke^{\frac{\sigma_k^2}{2}}$ and variance $k^2 e^{\sigma_k^2}(e^{\sigma_k^2} - 1)$. The pdf of $log(L_{\infty i})$ and $log(K_i)$ are

$$f(log\{L_{\infty i}\} = x \mid l_\infty, \sigma_\infty) = \frac{1}{\sigma_\infty \sqrt{2\pi}} exp\left[- \frac{1}{2} \left\{ \frac{x - log(l_\infty)}{\sigma_\infty} \right\}^2 \right] \tag{2.9}$$

and

$$f(log\{K_i\} = x \mid k, \sigma_k) = \frac{1}{\sigma_k \sqrt{2\pi}} exp\left[- \frac{1}{2} \left\{ \frac{x - log(k)}{\sigma_k} \right\}^2 \right] \tag{2.10}$$

We did not account for the log-transformation bias which will be small when σ is small.

The joint ME and BT density of the observation and associated random effects is

$$f_\theta(Y_i = y_i, L_{\infty i} = u, K_i = v | l_\infty, k, t_o, \sigma_\varepsilon, \sigma_\infty, \sigma_k) =$$

$$f(Y_i = y_i | L_{\infty i}, K_i, t_o, \sigma_\varepsilon) f(log(L_{\infty i}) = log(u) | l_\infty, \sigma_\infty) f(log(K_i) = log(v) | k, \sigma_k).$$

Let $\theta = (l_\infty, k, t_o, \sigma_\varepsilon, \sigma_\infty, \sigma_k)$. The marginal density of Y_i is

$$f_\theta\{Y_i = y_i \mid \theta\} = \iint f_\theta\{Y_i = y_i \mid L_{\infty i} = u, K_i = v\} f_\theta\{log(L_{\infty i}) = log(u)\}$$

$$f_\theta\{log(K_i) = log(v)\} \partial u \partial v. \quad (2.11)$$

The marginal likelihood for θ for the entire sample is

$$L(\theta) = \prod_{i=1}^{n} f_\theta(Y_i = y_i \mid \theta). \quad (2.12)$$

Let λ denote the vector of all random effects $(L_{\infty 1}, K_1), ..., (L_{\infty n}, K_n)$ and

$$f(\theta, \lambda) = \prod_{i=1}^{n} f_\theta(Y_i = y_i \mid L_{\infty i}, K_i, \theta) f_\theta(log(L_{\infty i}) = log(u)) f_\theta(log(K_i) = log(v)),$$

$$(2.13)$$

be the joint density for all n observations. The marginal likelihood function is

$$L(\theta) = \int f(\theta, \lambda) \partial \lambda = \int \exp\{h(\theta, \lambda)\} \partial \lambda, \quad (2.14)$$

where $h(\cdot)$ is the joint log-likelihood function of θ and λ. The main computational challenge is evaluating the marginal likelihood in Eqn. 2.14, because there is no analytical solution to this equation. We use the Template Model Builder (TMB) package in R to solve this problem.

2.2.1 Template Model Builder (TMB).

TMB is a free and open source new R package designed for fast fitting of complex nonlinear statistical models that may include random effects. The user must only define the joint log-likelihood function of the data conditional on the random effects and the log-likelihood for the random effects as a C++ template function. Other operations are done in R. TMB uses the Laplace approximation (Fournier et al., 2012) to Eqn. 2.14, which yields the marginal likelihood approximation as

$$L^*(\theta) = \det\{H(\theta)\}^{-1/2} \exp[h\{\hat{\lambda}(\theta), \theta\}], \tag{2.15}$$

where $\hat{\lambda}(\theta) = \text{argmax}_\lambda : h\{\lambda(\theta), \theta\}$ and $H(\theta) = -\frac{\partial^2}{\partial \lambda^2} h(\lambda, \theta)|_{\lambda=\hat{\lambda}(\theta)}$. The term $\exp[h\{\hat{\lambda}(\theta), \theta\}]$ in Eqn. 2.15 is a profile likelihood, which treats the random effect λ as nuisance parameters and θ as the parameter of interest.

In TMB, the Hessian, H in Eqn. 2.15 is evaluated by automatic differentiation (Fournier et al., 2012). From a user's perspective, the parameter estimation is fairly simple. We just need to specify the joint log-likelihood function $h(\theta, \lambda)$. TMB provides the gradient function for the marginal likelihood using automatic differentiation (AD) computation. Any gradient-based optimization method can be used to find the MLEs for θ. We use the *nlminb* routine in R (R Core Team, 2014). A major advantage of TMB is that it produces the marginal gradient function automatically after coding the joint log-likelihood function. This greatly improves the speed and accuracy of marginal MLEs for θ. In an R session, we read the data, dynamically link the C++ function template, set up the initial values for θ, specify the random effects, and optimize the objective function. TMB automatically provides a standard error report for $\hat{\theta}$ using the δ-method, and it also provides this for any differentiable function of θ and λ that the user specifies.

2.3 Simulation study of between-individual and measurement error variability in growth

In this section we present a simulation study to examine the performance of two estimators (MLE of MEL and MELBT) of the VonB growth parameters, and in separating the ME in length and BT variation. We randomly generated 2000 data sets as follows:

Step 1. We generated 1000 ages from a Gamma distribution with $\alpha = 3.5$ and $\beta = 2$ so that $E(t) = 7$ and $Var(t) = 14$. The pdf of the age (t) is

$$f(t) = \frac{1}{\Gamma(3.5)(2^3)(5)} t^{2.5} e^{\frac{-t}{2}} \tag{2.16}$$

and the age distribution is shown in Figure 2.1.

Step 2. With $l_\infty = 120$, $k = 0.3$, and $\sigma_\infty = \sigma_k = 0.1$; we randomly generated L_∞'s and K's from $120 \times e^{N(0,0.1)}$ and $0.3 \times e^{N(0,0.1)}$, respectively (i.e., CV= 10%). The distributions of the random L_∞ and K are shown in Figure 3.2.

MEBT data: We generated $N = 1000$ lengths using Eqn. 2.5, the ages (t's) from Step 1 and the random values for L_∞ and K from Step 2, with ME in length from $logN(0, 0.1)$. This produced data with both BT and ME in length variability.

ONBT data: We generated $N = 1000$ lengths using Eqn. 2.3, the ages (t's) from Step 1 and the random values for L_∞ and K from Step 2, with no ME in length. This produced data with only BT variation.

ONME data: We generated $N = 1000$ lengths using Eqn. 2.4, the ages (t's) from Step 1 with ME in length from $logN(0, 0.1)$ but with no BT variation; $l_\infty = 120$cm, and $k = 0.3$. This produced data with only ME in length variability.

In all we randomly generated 2000 ONBT, ONME and BTME data sets consisting of 1000 lengths and ages. A sample size of 1000 is not atypical for fisheries data.

We compare the variability in the ONBT, ONME and BTME data and the VonB curve (green solid line) which assumes no ME in length and no variation in growth parameters in Figure 2.3. The BTME data (bottom panel) is more variable than the ONBT data (middle panel) and the ONME (top panel) as we would expect. This is confirmed in Figure 2.4 with the residual plots.

We used the MLE procedure in R to compute the MLE of MEL parameters and TMB to compute the MLE of the MELBT parameters. Our primary goal is to determine which of the estimating methods is better for separating the sources of variation in the data.

2.3.1 Results

The model performance was measured using root mean square error (RMSE) of the estimates. Both methods (MLE of MEL and MELBT) converged in all 2000 data sets. The simulation averages and the corresponding RMSE's are presented in Table 3.5 for both methods. The results (Table 3.5) shows that the mean population growth parameter estimates (l_∞, k and t_o) are close to their true population values for all methods. However, the RMSE of the MLE of MELBT estimates are lower. The MLE of MEL estimated the ME in length variation (σ_ε) perfectly (estimated value same as true population value) for the ONME data. However, the MLE of MEL estimates for both ONBT and BTME data sets were discouraging, the optimizer did not know what to do when the variation in the data was only BT (see Figure 2.5). This is not surprising because the MLE of MEL is totally misspecified in both ONBT and BTME data sets. We will not report on the MLE of MEL further since the growth data we investigate include more than just ME in length variation.

The MLE of MELBT simulation averages of σ_ε, σ_∞ and σ_k for the ONME and ONBT data sets were surprising. For instance:

a) For the ONME data set, we were expecting estimates of σ_∞ and σ_k to be approximately 0, since the only variation in this data set is ME in length (σ_ε). While σ_k was close to 0, the estimated value of σ_∞ was equal to the estimated value of σ_∞.

b) With ONBT data, the variation in this data is BT (σ_∞, and σ_k) and there is no ME in length (thus $\sigma_\varepsilon = 0$), but the method estimates of σ_ε and σ_∞ were the same. σ_k was underestimated.

Additionally, the MLE of MELBT estimates of σ_∞, σ_k and σ_ε for the BTME data were not good. Again, σ_k was underestimated while the estimates of σ_ε and σ_∞ were close to the true population values. Regardless of the sources of variation in the data, the estimated values of σ_ε and σ_∞ were equivalent. This indicates that σ_ε and σ_∞ are confounded (see Eqn. 2.17 below); making it difficult for the MLE of MELBT to separate the sources of variation. In Figure 2.6 we can see the MLE of MELBT struggling to determine the estimates of σ_∞, σ_k and σ_ε for all data types.

After further consideration we realized that σ_∞ and σ_ε are completely confounded.

To see this, note that if $L_\infty = l_\infty e^{\delta}$, where $\delta \sim N(0, \sigma_\infty)$ then

$$\begin{aligned} Y_i &= L_{\infty i}\{1 - e^{-K_i(t-t_o)}\} \times e^{\varepsilon_i} \\ &= l_\infty e^{\delta_i}\{1 - e^{-K_i(t-t_o)}\} \times e^{\varepsilon_i} \\ &= l_\infty\{1 - e^{-K_i(t-t_o)}\} \times e^{\varepsilon_i+\delta_i} \end{aligned} \tag{2.17}$$

where $\varepsilon_i + \delta_i \sim N(0, \sigma_\varepsilon^2 + \sigma_\infty^2)$ and any combination of values for σ_ε and σ_∞ such that the total variance, $\sigma_\varepsilon^2 + \sigma_\infty^2$ is constant will give the same negative likelihood (nll) value. The result in Figure 2.6 does not show the confounding in the estimates because of starting value. We illustrate this with the ONME data set that produced parameter estimates, $l_\infty = 120.756$, $k = 0.295$, $t_o = 0.008$, $\sigma_\infty = 0.070$ $\sigma_k = 0.001$, and $\sigma_\varepsilon = 0.071$. The nll was -887.4498. If we replace $\sigma_\infty = 0$ and increase $\sigma_\varepsilon = \sqrt{\sigma_\varepsilon^2 + \sigma_\infty^2}$ then we get the exact same fit (nll=-887.4498).

In Chapter 3 we use a spatial model with certain constrains that allow us to identify σ_∞ and σ_ε (this will be explained in Chapter 3).

2.4 Figures

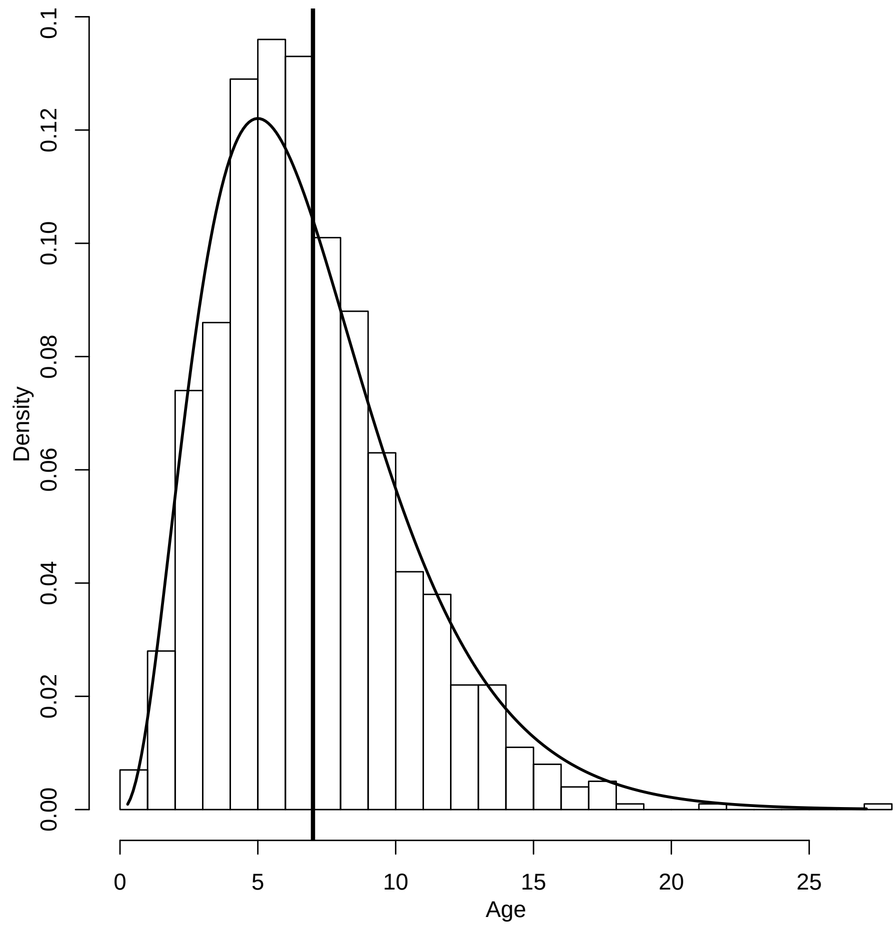

Figure 2.1: Distribution of N = 1000 ages. The solid curve is the Gamma pdf and the vertical line indicates the mean age.

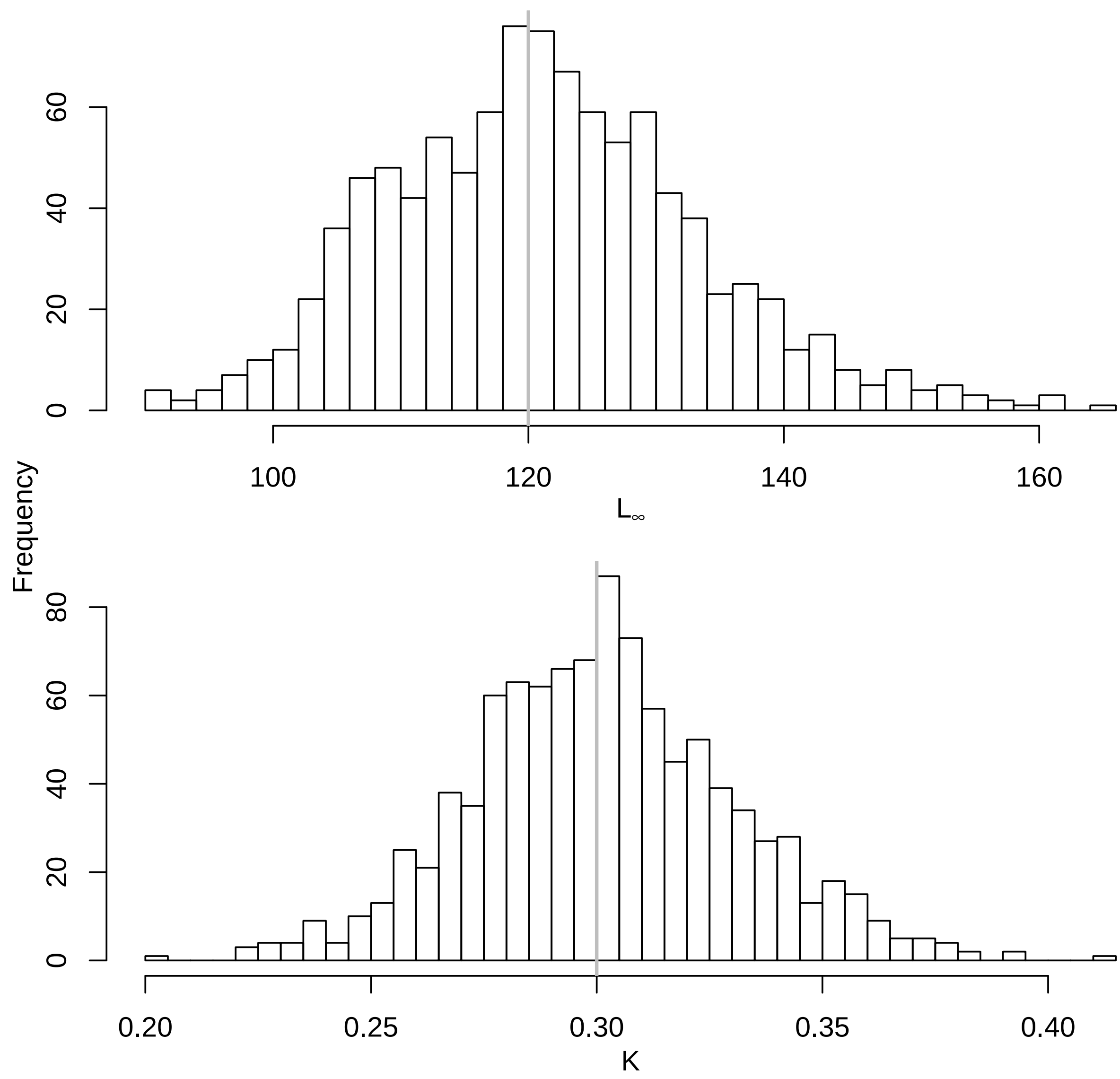

Figure 2.2: Distribution of N=1000 random values of L_∞ and K. The vertical grey lines indicate the populations mean.

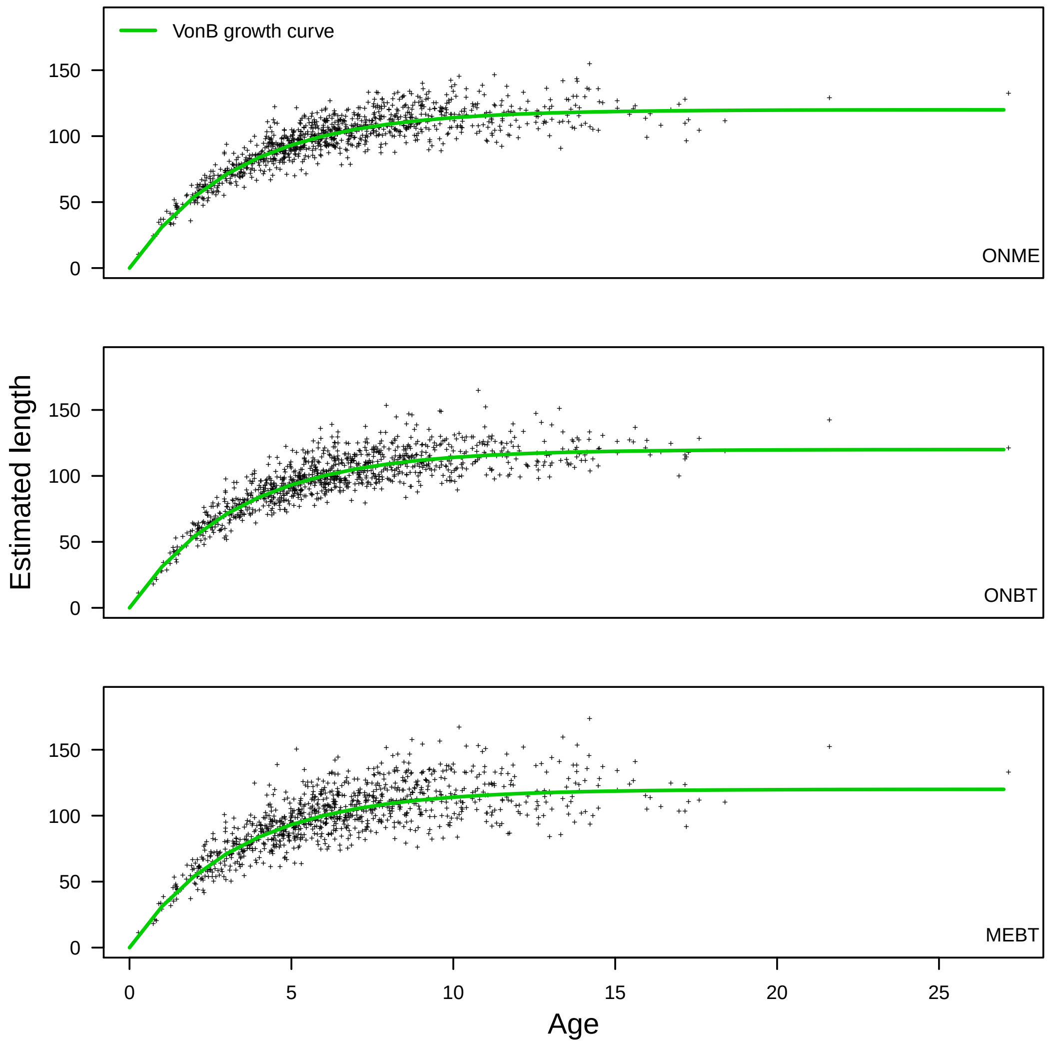

Figure 2.3: Length verses age. The ONME data contains only measurement error (ME) in length variation; ONBT data (middle panel) contains only between-individual (BT) variation; MEBT data contains both ME and BT source of variability.

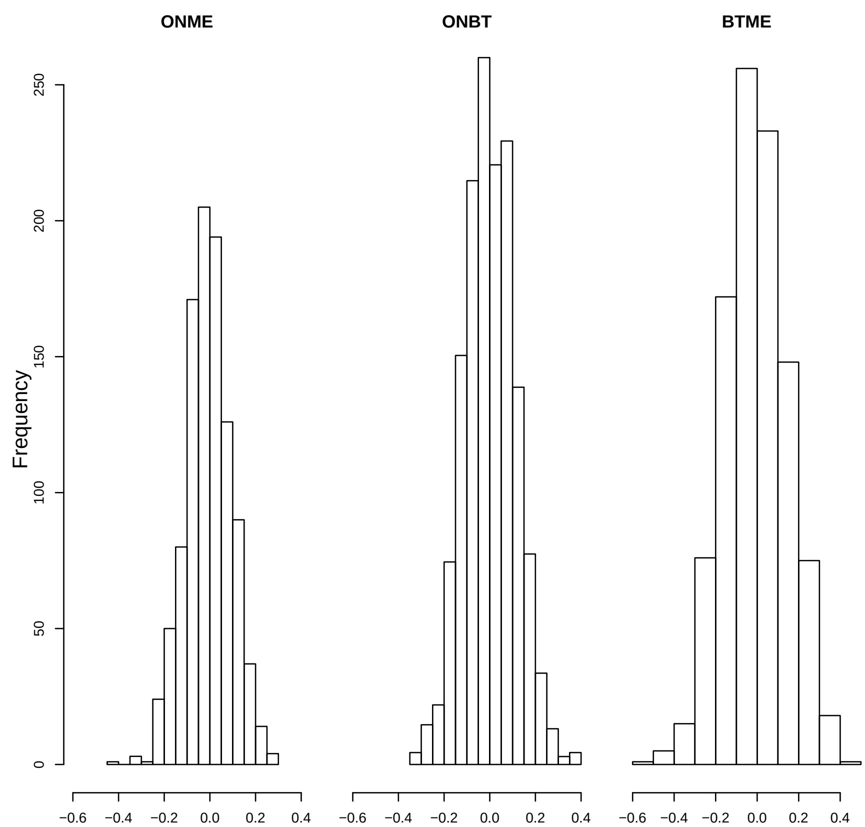

Figure 2.4: Residual plot for the ONME, ONBT and BTME data sets in Figure 2.3.

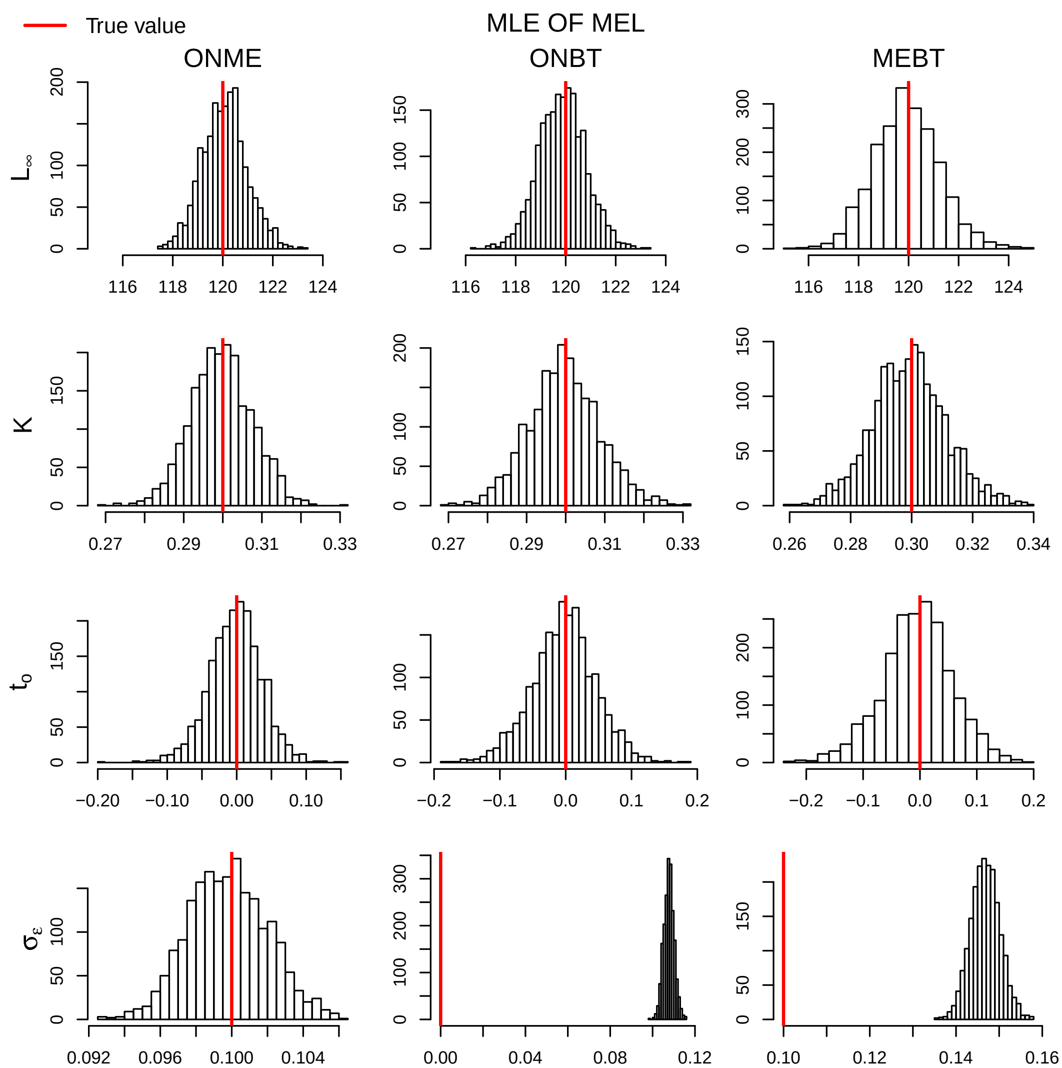

Figure 2.5: Simulation estimates of the mean growth parameter and the BT and ME in length variations.

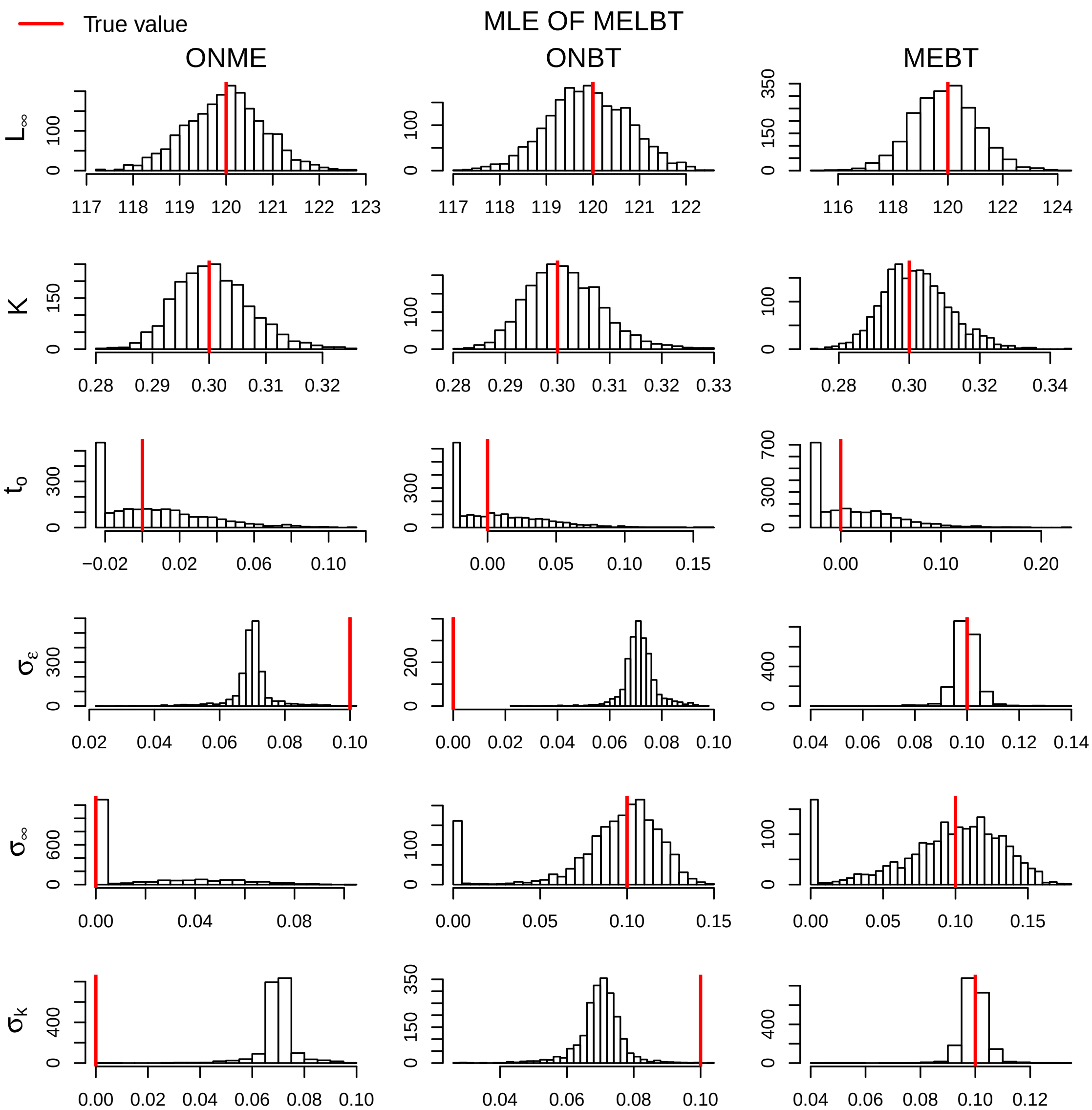

Figure 2.6: Simulation estimates of the mean growth parameter and the BT and ME in length variations.

2.5 Tables

Variation in data	Parameter	Value	MLE of MELBT Estimate	MLE of MELBT RMSE	MLE of MEL Estimate	MLE of MEL RMSE
ONME	l_∞	120.000	119.976	0.835	120.070	0.893
	k	0.300	0.301	0.007	0.300	0.008
	t_o	0.000	0.005	0.028	-0.002	0.038
	σ_ε	0.100	0.069	0.031	0.100	0.002
	σ_∞	0.000	0.069	0.031	-	-
	σ_k	0.000	0.001	0.087	-	-
ONBT	l_∞	120.000	119.900	0.866	119.862	0.949
	k	0.300	0.301	0.007	0.300	0.009
	t_o	0.000	0.007	0.033	-0.003	0.049
	σ_ε	0.000	0.070	0.031	0.108	0.008
	σ_∞	0.100	0.071	0.029	-	-
	σ_k	0.100	0.058	0.034	-	-
BTME	l_∞	120.000	119.885	1.186	119.936	1.305
	k	0.300	0.302	0.010	0.299	0.012
	t_o	0.000	0.012	0.041	-0.005	0.062
	σ_ε	0.100	0.099	0.005	0.147	0.047
	σ_∞	0.100	0.099	0.005	-	-
	σ_k	0.100	0.049	0.041	-	-

Table 2.1: Simulation averages for 2 estimators of von Bertalanffy growth parameters, measurement error in length and between-individual variation.

Chapter 3

Growth model including measurement error in age.

We demonstrated in Chapter 2 that the parameters that describes between-individual (BT) variation (σ_∞ and σ_k), and the standard deviation of ME in length (σ_ε) were confounded. In this Chapter, we fix this confounding using a spatial model and include ageing error to have a more realistic error structure. In practise length can be estimated much more accurately than age. We still expect the ageing error standard deviation (σ_ξ) to be confounded with σ_∞ and σ_k. Ideally if we have replicate measurements of age then we could estimate σ_ξ from the replicates, then estimate σ_∞ and σ_k from the data. However, we do not have replicated age measurements and σ_ξ

will be complicated, as it depends on the true age of a fish and the individual reading the age. The σ_∞ and σ_k are also more complicated than we have so far discussed. There will be a spatial and temporal component of these parameters, since differences in environmental conditions (such as temperature and food resources) will result in growth rate variation.

We will assume that σ_∞ and σ_k are negligible for fish from the same "school". It is not possible to identify schools based on the sampling design used for the data available to us. As a proxy, we assume that σ_∞ and σ_k are negligible for fish caught in the same year and spatial stratum. The data we examine comes from a stratified sampling design in which there are many small spatial strata. Variation in growth for such fish is attributed to ME in length and age (σ_ε and σ_ξ). The rationale for this assumption is that fish from the same spawning schools are genetically more similar and have experienced and grown in more similar physical environments than fish in different spawning schools. Thus, we assume that variation in the growth rate of individuals within a stratum is due to measurement error in length and age (σ_ε and σ_ξ) only, and the value for σ_ε is known. This makes it possible to estimate σ_ξ.

We study two ME models described by Carroll et al. (2006): the functional errors-in-variables (FEV) and the structural errors-in-variables (SEV) models. The FEV model considers the true age of the ith individual to be an unknown parameter, whereas the SEV approach considers the true age of the ith individual to be a random variable with some distribution.

Both approaches have been used in the area of fisheries research. Cope & Punt (2007) studied a related problem using the SEV approach. They used two models for the distribution of the true ages A^T; exponential and gamma distributions. They estimated the ageing error standard deviation ($\sigma_{A_i^T}$) separately based on multiple age readings and by assuming estimated ages are unbiased. However, in the simulation studies they included the situation where $\sigma_{A_i^T}$ increased with age as we would expect. In practice, we would not expect the true age distribution ($Pr(A_i^T)$) to have a simple form like Cope & Punt (2007). This is because the true age distribution of a population is the result of complex temporal variability in previous reproduction and survival rates (e.g., Kitakado, 2001). It is realistic to expect a continuous probability distribution with a multi-modal distribution for temperate and boreal fish species, that reproduce in specific seasons and the subsequent survival rates differ from one individual (fish) to the other.

Kitakado (2001) used simulation studies to investigate a particular FEV approach called the conditional-score function. This approach conditions on sufficient statistics to obtain estimators and it reduced biased caused by nuisance parameters (Carroll et al., 2006). He found that the estimating functions of the conditional score method

(FEV) are unbiased, while those of the commonly used MLE methods were biased. The conditional-score function approach is not pursued in this study; however, we investigate an alternative approach to deal with the unknown age nuisance parameters (see Section 3.3).

We use a similar SEV approach to Cope & Punt (2007) but we modelled the ageing error such that the coefficient of variation (CV) is a linear function of the true age of the fish. This is because, as fish get older, rings on their otoliths tend to shrink, making it more difficult to count. Therefore, ageing errors are larger for older fish but not necessarily with a constant CV. Additionally, we account for spatial and temporal variability by using a spatially stratified model.

We use simulation studies and the TMB package in R to examine the performance of the two estimators (FEV and SEV) in estimating the VonB growth parameters and separating the ME in age and the BT variation. The SEV approach estimates the mean growth parameters l_∞, k, the variance of L_∞, and K, and the ME variance in age. The SEV approach also estimates parameters of the distribution of true ages. The FEV approach estimates the mean parameters (l_∞ and k), and the standard deviations for L_∞ and K, as well as the unknown ages that are measured with error for each fish in the sample and ME variance.

3.1 Model

Let the length of the fish at age A be Y. For this investigation we used the VonB growth model, which is a parametric model of Y as a function of A. Here the observations of Y and A both include measurement error, while Y^T and A^T indicate their true values without measurement error. The VonB growth model is

$$Y^T(A^T, L_\infty, K) = L_\infty\{1 - \exp(-KA^T)\}. \tag{3.1}$$

We account for BT variation using different VonB growth parameters for each strata. Assume the data come from $s = 1, ..., S$ strata. The length at age A for the i^{th} individual in stratum s is $Y(A, L_{\infty s}, K_s)$, where

$$\log(L_{\infty s}) \stackrel{iid}{\sim} N\left\{\log(l_\infty), \sigma^2_\infty\right\}, s = 1, ..., S, \tag{3.2}$$

and

$$\log(K_s) \stackrel{iid}{\sim} N\left\{\log(k), \sigma^2_k\right\}, s = 1, ..., S. \tag{3.3}$$

The length of i^{th} individual including measurement error (ME) is

$$Y_i = Y_i^T + \varepsilon_i, \qquad \text{where} \qquad \varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2_\varepsilon). \tag{3.4}$$

Additive ME in length seems reasonable. ME in age is more difficult since ageing errors are larger for older fish (Morales-Nin, 1992) but not necessarily with a constant

CV. We assume an additive model for ME in age.

$$A_i = A_i^T + \xi_i, \qquad \text{where} \qquad \xi_i \overset{iid}{\sim} N\left\{0, \sigma_\xi^2(A_i^T)\right\} \tag{3.5}$$

and

$$\sigma_\xi(A_i^T) = (\sigma_{o\xi} + \sigma_{1\xi} A_i^T) A_i^T. \tag{3.6}$$

That is, the CV is a linear function of the true age. The equations 3.5 and 3.6 represent the ME in age processes.

3.2 Estimating method 1: Structural Errors in Variables Model (SEV)

In this approach the true age of the ith individual is assumed be a random variable with some distribution like the gamma. The probability that the i^{th} fish in the sample has measured age a and length y, is

$$\Pr(A_i = a, Y_i = y) = \iint \Pr(A_i = a, Y_i = y | A_i^T, Y_i^T) \Pr(A_i^T, Y_i^T) \partial A_i^T \partial Y_i^T. \tag{3.7}$$

We assume for now that ME in A and Y are independent[1], i.e.,

$$\Pr(A_i = a, Y_i = y | A_i^T, Y_i^T) = \Pr(A_i = a | A_i^T) \Pr(Y_i = y | Y_i^T). \tag{3.8}$$

[1]This assumption may need more consideration. If the length of a fish is used to help determining the age when the number of growth rings is not clear, then this may introduce some complex dependencies in this distribution.

Equations (3.4) and (3.5) specify $\Pr(Y_i = y|Y_i^T)$ and $\Pr(A_i = a|A_i^T)$, respectively. The joint distribution of A^T and Y^T can be expressed as

$$\Pr(A_i^T, Y_i^T) = \Pr(Y_i^T|A_i^T)\Pr(A_i^T). \tag{3.9}$$

If there is no between-individual variability in growth rates, then Y^T is a deterministic function of A^T; however, this is not realistic. As a result, we assume between-individual variability is present (i.e., Eqn. 3.3) so that

$$\Pr(Y_i^T|A_i^T) = \iint \Pr(Y_i^T|A_i^T, L_{\infty i}, K_i)\Pr(L_{\infty i}, K_i)\partial L_{\infty i}\partial K_i. \tag{3.10}$$

If we know the true age of the i^{th} fish when measured, and its specific VonB growth parameters ($L_{\infty i}$ and K_i), we can assume that we would know the length of the fish exactly. That is, $Y_i^T|(A_i^T, L_{\infty i}, K_i) = \mu\left(A_i^T, L_{\infty i}, K_i\right)$ is the VonB model value of length for an age A_i^T of fish. Hence,

$$\Pr(Y_i^T|A_i^T, L_{\infty i}, K_i) = \begin{cases} 1, & Y_i^T = \mu\left(A_i^T, L_{\infty i}, K_i\right), \\ 0 & \text{otherwise.} \end{cases} \tag{3.11}$$

We use the gamma distribution proposed by Cope and Punt (2007) for $\Pr(A_i^T)$,

$$\Pr(A_i^T = a) = \frac{a^{\gamma-1}\exp(-a/\beta)}{\beta^{\gamma}\Gamma(\gamma)}. \tag{3.12}$$

This is a right-skewed distribution with parameters γ and β, $E(A_i^T) = \gamma\beta$ and $Var(A_i^T) = \gamma\beta^2$. The parameters of interest are $\theta = (l_\infty, k, \gamma, \beta, \sigma_\infty, \sigma_k, \sigma_{o\xi}, \sigma_{1\xi})$. Note that we assume the ME in length standard deviation (σ_ε) is known. Combining Equations (3.7)-(3.10), the joint probability of A_i and Y_i depends on θ and is

$$\Pr(A_i, Y_i; \theta) = \iiiint \Pr(A_i|A_i^T)\Pr(Y_i|Y_i^T)\Pr(Y_i^T|A_i^T, L_{\infty i}, K_i)$$
$$\Pr(L_{\infty i})\Pr(K_i)\Pr(A_i^T)\partial L_{\infty i}\partial K_i\partial A_i^T\partial Y_i^T, \quad (3.13)$$

and assuming Equation (3.11),

$$\Pr(A_i, Y_i; \theta) = \iiint \Pr(A_i|A_i^T)\Pr\left\{Y_i|Y_i^T = \mu\left(A_i^T, L_{\infty i}, K_i\right)\right\}$$
$$\Pr(L_{\infty i})\Pr(K_i)\Pr(A_i^T)\partial L_{\infty i}\partial K_i\partial A_i^T. \quad (3.14)$$

Let D be the set of observed data, $D = \{(a_1, y_1), ..., (a_n, y_n)\}$. The likelihood for θ,

$$L_{SEV}(\theta; D) = \prod_{i=1}^{n} \Pr(A_i = a_i, Y_i = y_i; \theta), \quad (3.15)$$

is computed using Equation (3.14) based on Equations (3.1)-(3.6) and (3.12).

3.3 Estimating method 2: Functional Errors in Variables (FEV) and restricted MLE

In this approach, we do not make any assumptions about the distribution of age in the population. The parameters of interest $\theta = (l_\infty, k, \sigma_\infty, \sigma_k, \sigma_{o\xi}, \sigma_{1\xi})$ and the n unknown ages, $A_1^T, ..., A_n^T$. The probability distribution of the i^{th} fish in the sampled measured with age a and length y is

$$\Pr(A_i, Y_i; \theta, A_i^T) = \iint \Pr\left\{Y_i | Y_i^T = \mu\left(A_i^T, L_{\infty i}, K_i\right)\right\}$$
$$\Pr(A_i | A_i^T)\Pr(L_{\infty i})\Pr(K_i)\partial L_{\infty i}\partial K_i. \quad (3.16)$$

The likelihood is

$$L(\theta, \{A_1^T, ..., A_n^T\}; D) = \prod\nolimits_{i=1}^{n} \Pr(A_i = a_i, Y_i = y_i; \theta, A_i^T).$$

Note that the number of parameters is $n + 8$ and the number of observations are $2n$.

When the number of nuisance parameters is large relative to n, it is well known that maximum likelihood estimators (MLEs) for some parameters, particularly for variances, can be biased and inefficient (e.g., Barndorff-Nielsen and Cox, 1994; Berger et al., 1999). Finding a conditional distribution or score function that does not depend on nuisance parameters is one approach to deal with this problem. This was the approach used by Kitakado (2001) for growth parameter estimation from capture-recapture data. Various likelihood adjustments have been proposed to correct for the bias in MLEs resulting from many nuisance parameters and many of these are reviewed by Severini (2000) and Cox (2006). Berger et al. (1999) advocated using integrated likelihood methods to eliminate nuisance parameters. This approach is related to restricted maximum likelihood estimation (REML; see Searle et al., 1998), which is a class of methods commonly used to reduce or eliminate bias in MLEs of variance parameters. This is the approach we pursue. The integrated likelihood for θ in which the age parameters are simply integrated out of the likelihood is

$$L_{SEV}(\theta; D) = \prod_{i=1}^{n} \int \Pr(A_i = a_i, Y_i = y_i; \theta, A_i^T) \partial A_i^T. \tag{3.17}$$

This is the same as Equation (3.15) if $\Pr(A_i^T) = 1$ in a Bayesian context. This is the most commonly used default conditional prior and Berger et al. (1999) used it in several examples to improve statistical inference when nuisance parameters were present. There are potential problems with this approach; most notably, the resulting integrated likelihood may not exist; however, it is a simple (apart from the integration) and objective method, and the one we pursue. We use TMB for the calculation of the integrated likelihood.

3.4 Parameter estimation

Let $\lambda = \{A_1^T, ..., A_n^T, L_{\infty 1}, ..., L_{\infty n}, K_1, ..., K_n\}$ be the vector of the random effects in the SEV approach or the vector of fixed age-effects plus random VonB parameters in the FEV approach. The joint density of the observed data (D) is

$$f(D; \theta, \lambda) = \begin{cases} \prod_{i=1}^{n} \Pr\left\{Y_i | Y_i^T = \mu\left(A_i^T, L_{\infty i}, K_i\right)\right\} \Pr(A_i | A_i^T) \Pr(L_{\infty i}) \Pr(K_i) \Pr(A_i^T), & \text{SEV}, \\ \prod_{i=1}^{n} \Pr\left\{Y_i | Y_i^T = \mu\left(A_i^T, L_{\infty i}, K_i\right)\right\} \Pr(A_i | A_i^T) \Pr(L_{\infty i}) \Pr(K_i), & \text{FEV}. \end{cases}$$

The SEV marginal likelihood and the FEV restricted marginal likelihood function can be expressed as

$$L(\theta; D) = \int \cdots \int f(D; \theta, \lambda) d\lambda = \int \cdots \int \exp\{h(D; \theta, \lambda)\} d\lambda. \tag{3.18}$$

The main computational challenge is evaluating the integral in Eqn. (3.18) because there is no analytical solution. TMB uses the Laplace approximation (e.g., Skaug and Fournier, 2006; see Chapter 2) to solve Eqn. 3.18.

3.5 Simulation Studies

We conducted a simulation study to investigate the performance of the FEV and SEV methods in estimating the VonB growth parameters, separating the ME in age, and BT variation. The performance of the SEV and FEV estimators was measured using root mean square error (RMSE) based on 1000 simulated data sets.

We illustrate methods by simulating a single data set consisting of 1000 length and age measurement using the following examples.

Example 1: Age follows gamma distribution

We generated 1000 ages from a gamma distribution with $\alpha = 7$ and $\beta = 1$ so that $E(A) = 7 = Var(A)$. Lengths were generated from the VonB model (3.1), with $l_\infty = 120$, $\sigma_\infty = 0.1$, $k = 0.2$, and $\sigma_k = 0.1$ in Equation (3.3). We set the age $t_o = -0.5$ when the length is zero. We set $\sigma_\varepsilon = 0.05$, and for age ME we used $\sigma_{o\xi} = 0.05$ and $\sigma_{1\xi} = 0.0075$. We fixed the number of strata in our simulation to be 250 and we randomly assigned the length samples to one of the 250 strata, and assumed that L_{∞_i} and K_i are the same for all fish sampled within a stratum; that is, there is no BT variation in fish from the same strata. The number of samples per stratum ranged from 1 to 9 (Table 3.1). In this table, 15 strata had only one age-length sample, 32 strata has 2 samples, etc. The distribution of true ages is shown in Figure 3.1, and the distributions of the random L_∞ and K are shown in Figure 3.2. The range of ME in ages appears realistic (Figure 3.3), as does the variation in length at age (Figure 3.4). The SEV model is exactly specified in this example because the same gamma distribution for true ages was used when generating simulated data and when estimating the VonB parameters.

Example 2: Age follows a mixture of gamma distribution.

In this example, 1000 ages were generated from a mixture of gamma distributions (Figure 3.6), but the other simulation parameters were the same as for Example 1. Such an age distribution will often be unrealistic but the purpose of this example was to investigate the efficacy of the SEV model when the true age distribution is substantially different from the assumed gamma distribution. However, the range of ME in ages (Figure 3.7) and the variation in length at age (Figure 3.8) are realistic.

Example 3: Age follows a mixture of normal distribution.

In this example, we used a different true age distribution, based on a mixture of normal distributions (Figure 3.10). Other simulation parameters were the same as for Example 1 (Figures 3.11-3.12).

These three example describe three simulation models, each with a different distribution for the true age. We repeated this procedure 1000 times to generate simulated data sets consisting of $N = 1000$ length and age measurements.

3.5.1 Results for single data set

Example 1 result (Table 3.2): The SEV estimates of the VonB growth parameters (l_∞ and k), and their standard deviations (σ_∞, and σ_k, respectively), are fairly close to their true values. The 95% CIs of these parameters cover the true values, however, the CIs for σ_∞ and σ_k are fairly wide, which indicates the difficulty in separating the sources of variation. The FEV estimates are less accurate and their CIs for l_∞ and k do not contain the true values.

This may be related to how the unknown true ages are modelled in the SEV and FEV models. We can predict the values for these ages; TMB provides these automatically using the method outlined in Skaug and Fournier (2006). The total variation in these age predictions, as a percentage of the total variation in the observed ages with ME, is higher in the FEV model than the SEV model (Figure 3.5) and higher still than the total variation in true ages. This indicates that the FEV predicted true ages have higher total variation than the observed ages with ME. Even if we do not know the true distribution of ages, it is reasonable to expect that their total variation should be less than the total variation of the observed ages, which are simply the true ages plus measurement error; hence the FEV age predictions are not reasonable. There is a systematic bias in both the SEV and FEV predicted ages (middle panel), particularly for older ages, which explains part of the reason why the VonB parameters differ from their true values.

Example 2 result (Table 3.3): The SEV parameter estimates are almost as reliable as in Example 1 and the CIs for l_∞ and k contain the true values. Similar to Example 1, the SEV model provided estimates that are closer to their true population values than the FEV model. The results are somewhat surprising because in this example the SEV model is mis-specified because the SEV estimation is based on a simple gamma distribution assumption for true ages whereas the true ages were generated from a more complicated gamma mixture distribution (see Figure 3.6). The FEV model is not mis-specified because it makes no assumption about the distribution of true ages. Predicted values for the true ages (Figure 3.9) have a similar pattern compared to Figure 3.5, but the predicted ages appear to be closer to the true ages in Figure 3.9.

Example 3 result (Table 3.4): Again the SEV model provided estimates that are closer to their true values when compared to the FEV model. The FEV model predictions of the true ages (Figure 3.13) were somewhat closer to the true values than predicted by the SEV model but still the parameter estimates were less accurate.

3.5.2 Results for 1000 data sets

The SEV and FEV converged in all 1000 data sets. The mean of the 1000 estimated parameters and their root mean square error (RMSE) is shown in Tables 3.5, 3.6 and 3.7 for examples 1, 2, and 3, respectively. In all three examples, the SEV mean estimates of the VonB parameters and their standard deviations (σ_∞ and σ_k) are fairly close to their true population values compared to the FEV estimates. The FEV estimates have higher RMSE than the SEV estimates (i.e., FEV model estimates are less accurate). In the FEV approach, the true ages are not treated as random and would require bias corrections for estimating equations. Additionally, the SEV model is reliable even when the true age distribution is different from a gamma (compare the SEV estimates in Tables 3.1, 3.3, and 3.4). We conclude that the SEV method is more reliable than the REML FEV method we investigated and the SEV is the estimating method we pursue in the remainder of this report.

In the next Chapter, we extend the growth model to account for variations that may be caused by the way fish are caught, and sampled for length and age measurements.

3.6 Figures

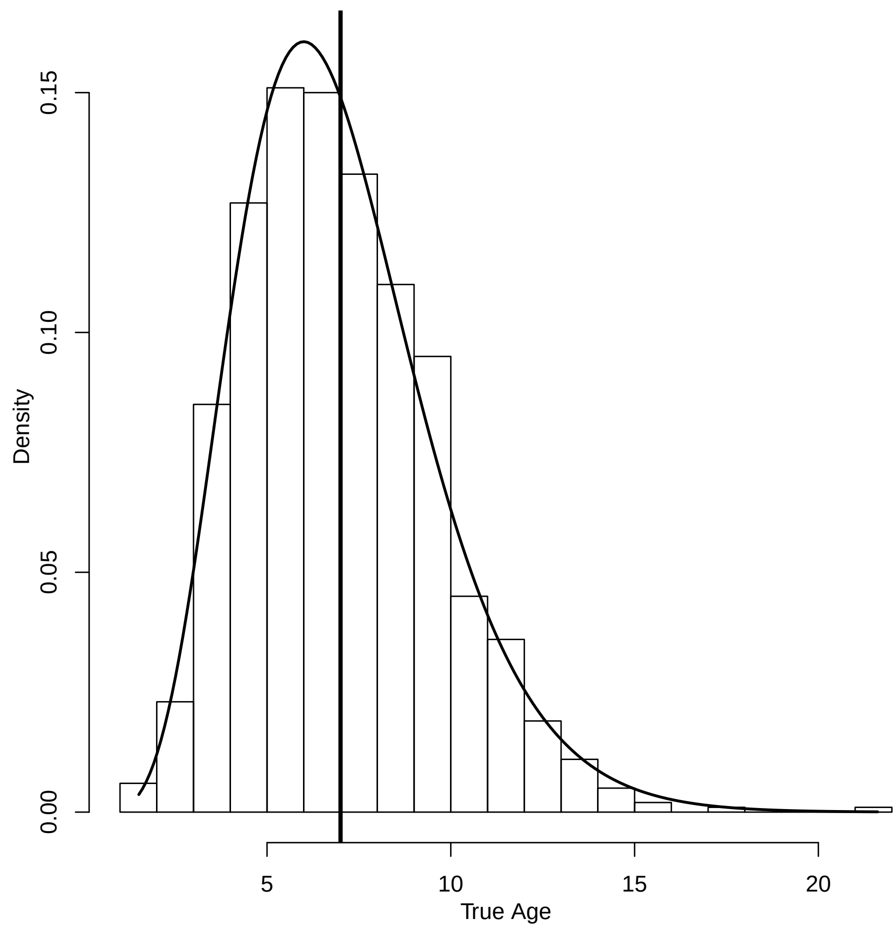

Figure 3.1: Distribution of $N = 1000$ true ages for Example 1.

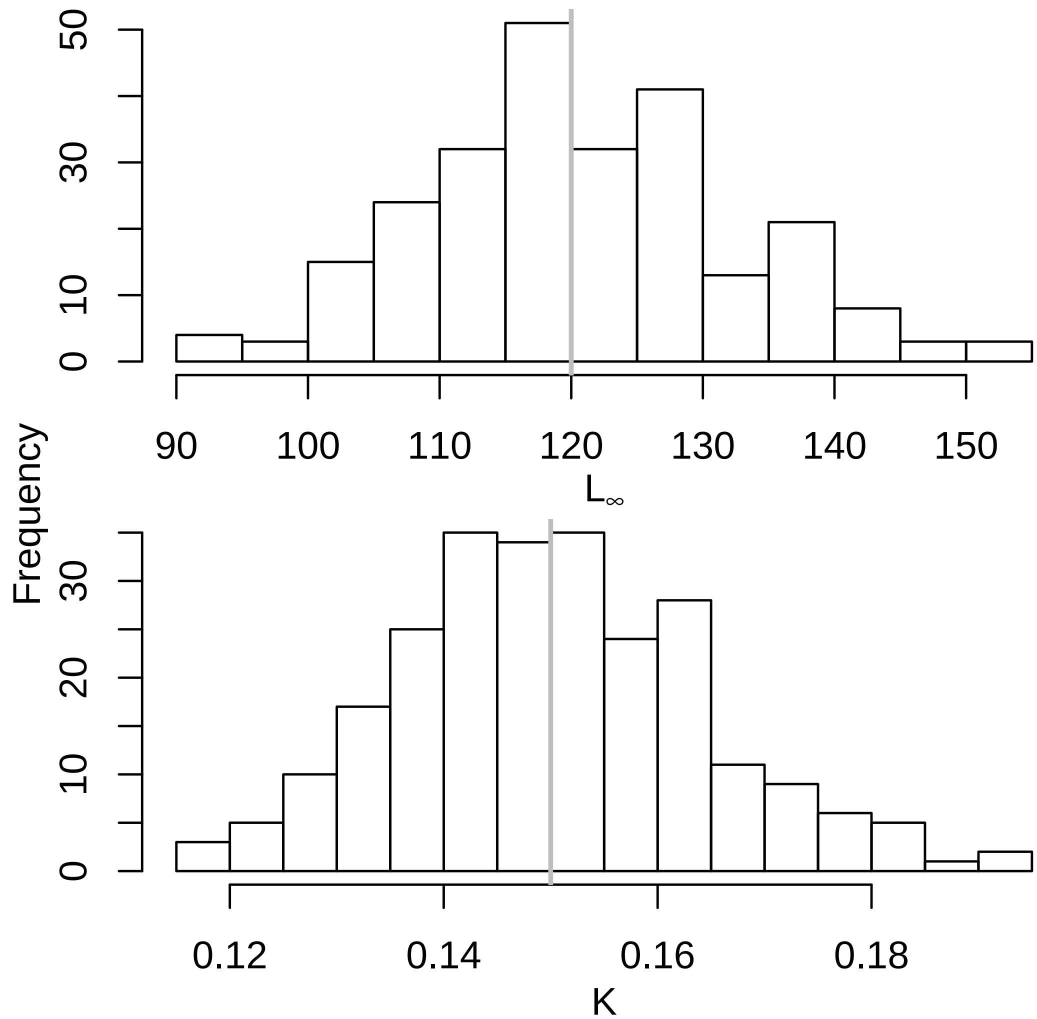

Figure 3.2: Distribution of $S = 250$ with random values of L_∞ and K. The vertical lines indicate the populations mean.

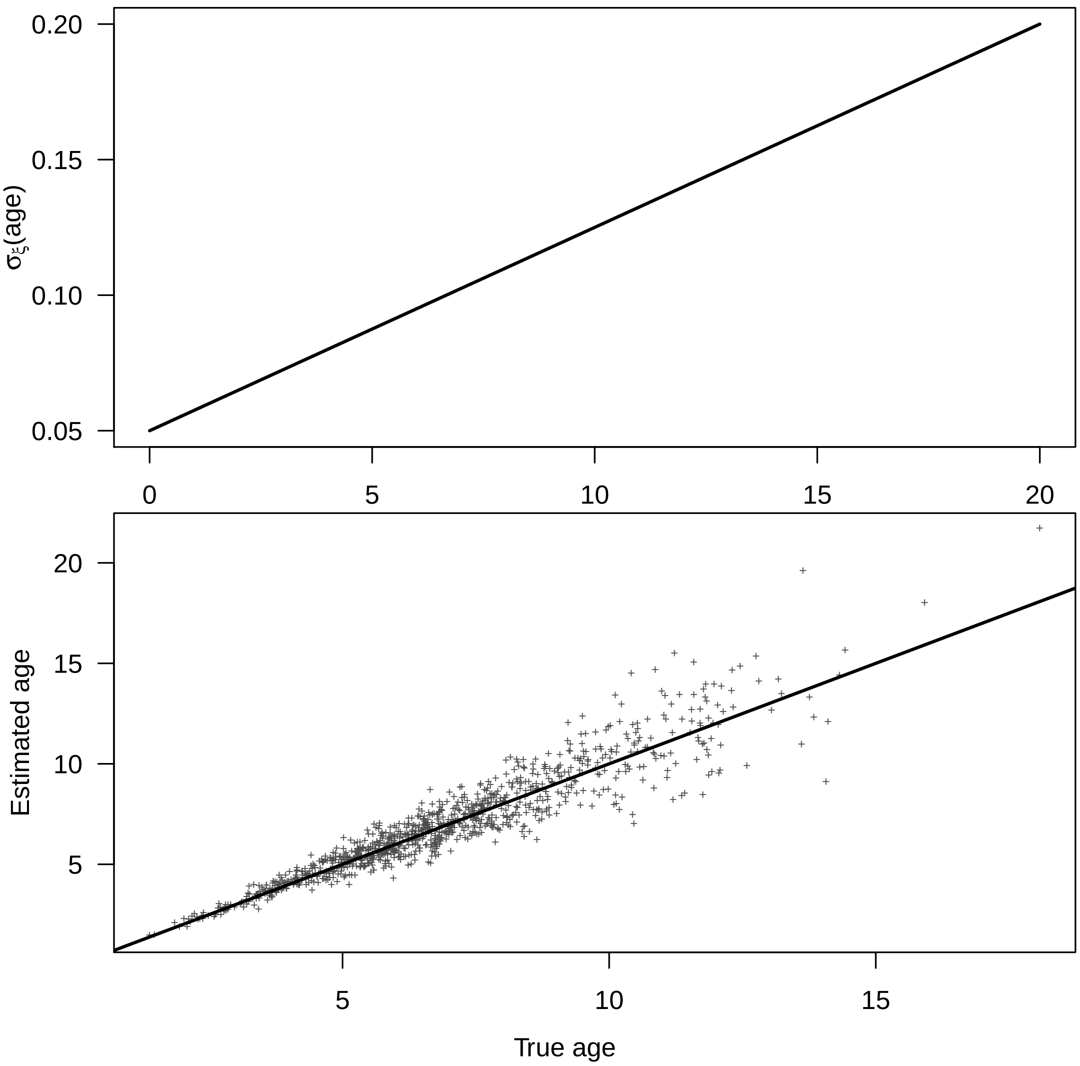

Figure 3.3: Top panel: Age ME coefficient of variation as a function of true age. Bottom panel: Estimated ages versus true ages and the 1-1 reference line for Example 1.

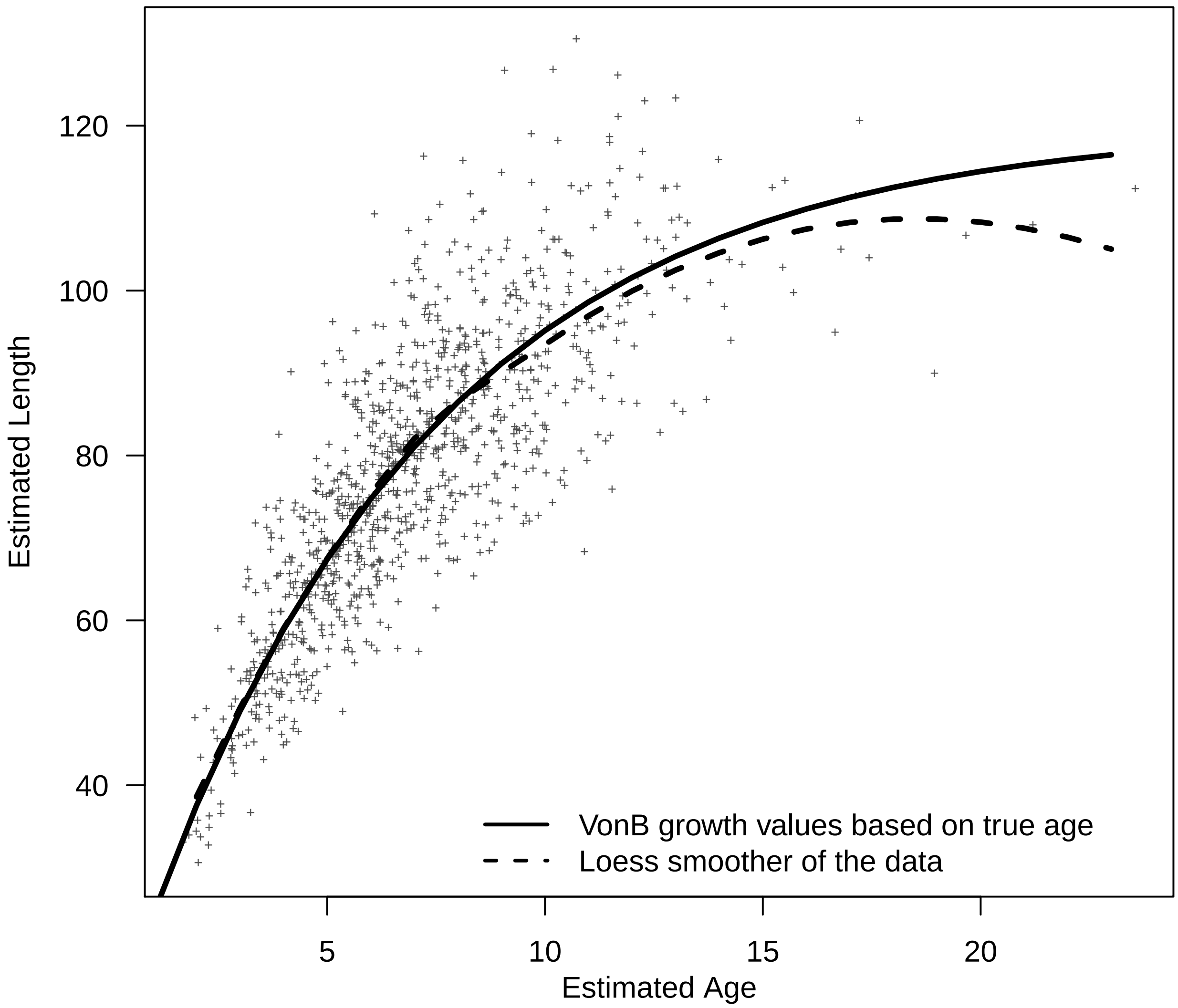

Figure 3.4: Length versus age for Example 1. Both variables are measured with error.

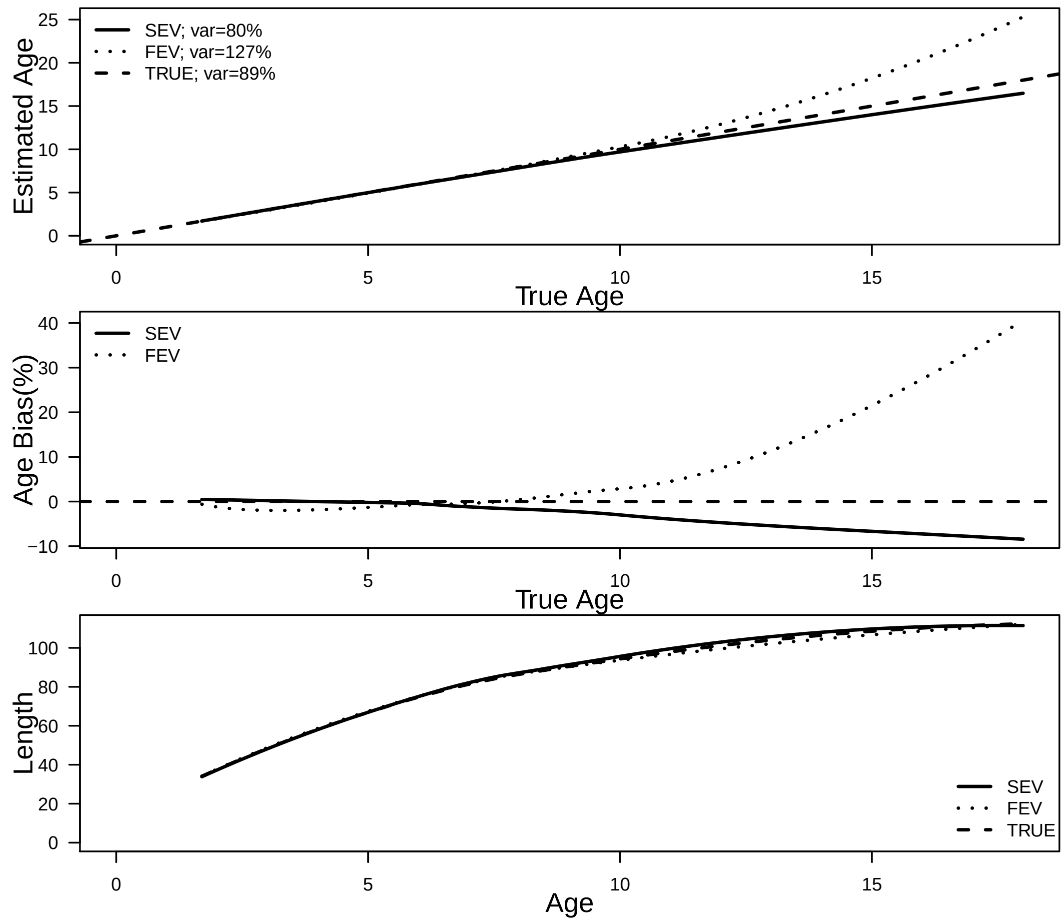

Figure 3.5: Some model diagnostics from Example 1. Top panel: Smoothed model predicted ages versus true age (simulation) values. Middle panel: Bias in model predicted ages, in percent of true ages. Bottom panel: Smoothed lengths with ME as functions of true and model predicted ages.

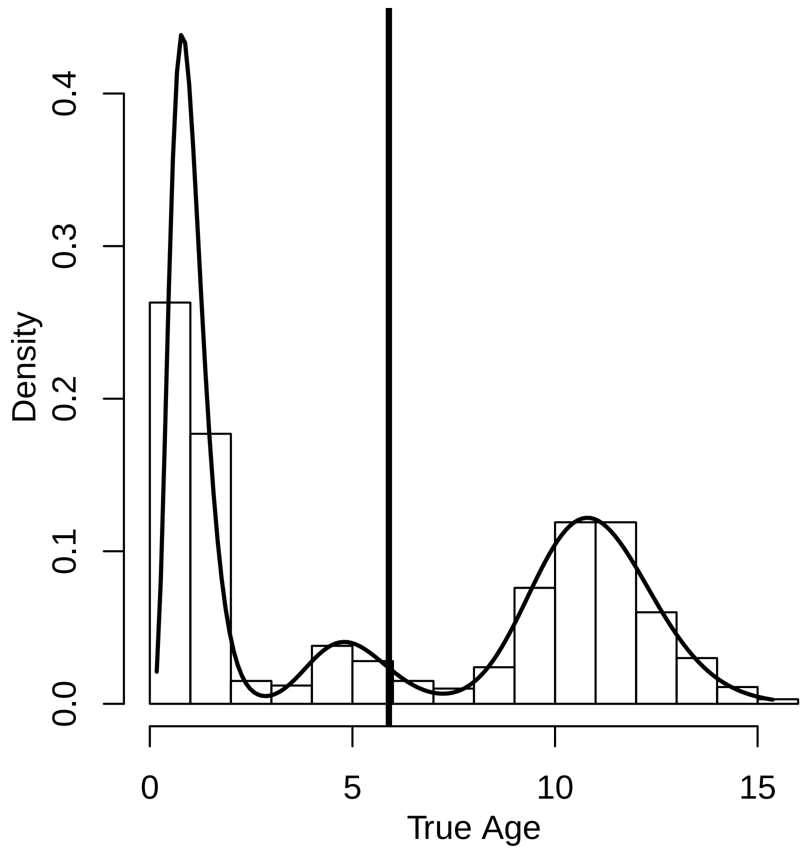

Figure 3.6: Distribution of $N = 1000$ true ages for Example 2. The solid curve is the Gamma mixture pdf and the vertical line indicates the mean age.

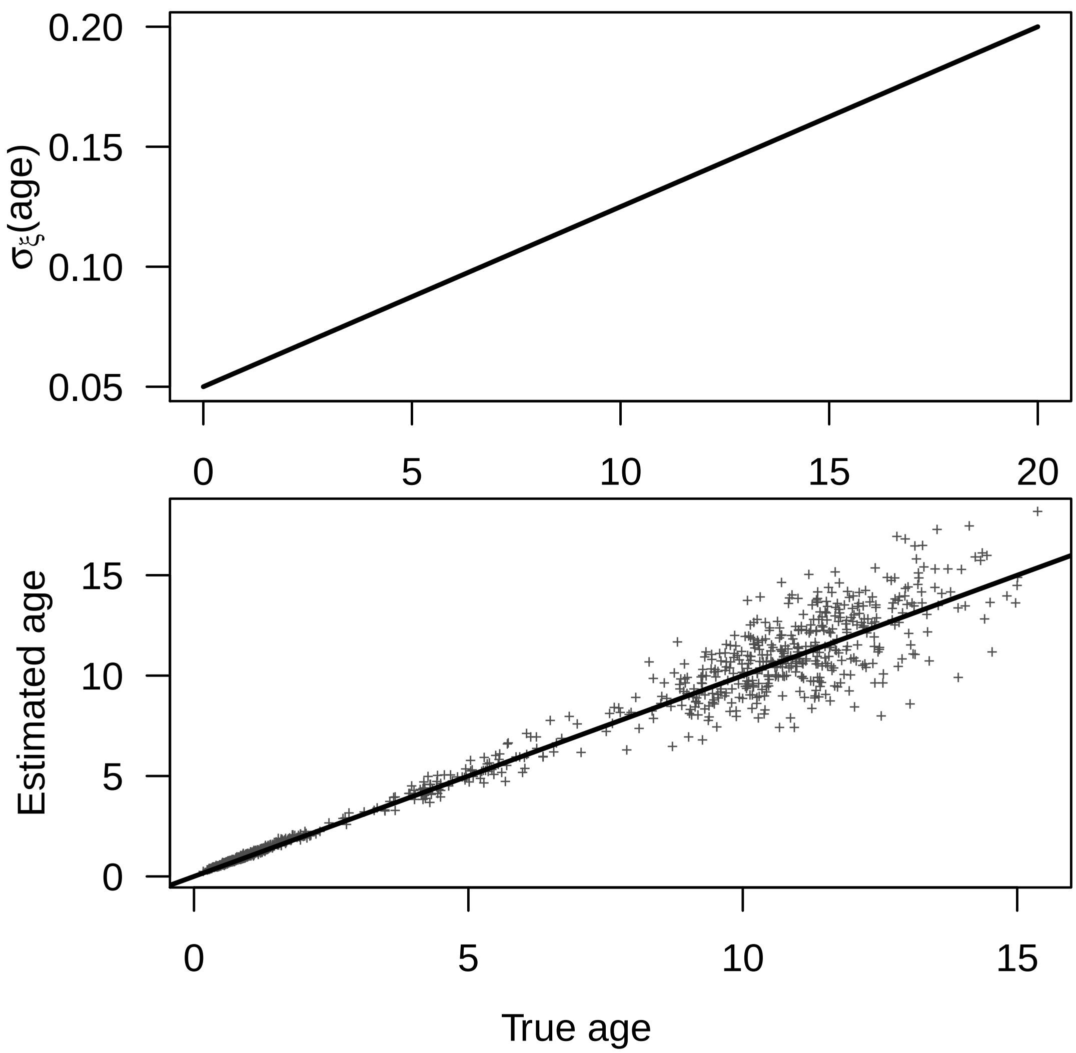

Figure 3.7: Top panel: Age ME coefficient of variation as a function of true age. Bottom panel: Estimated ages versus true ages and the 1-1 reference line for Example 2.

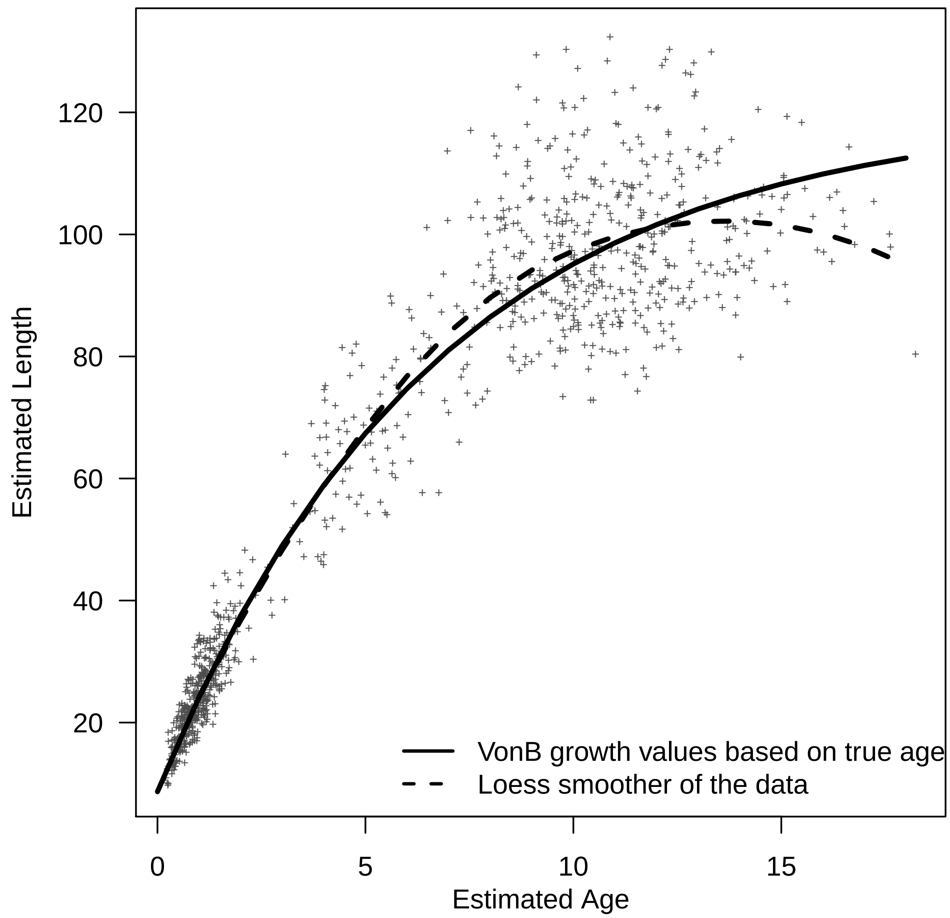

Figure 3.8: Length versus age for Example 2. Both variables are measured with error.

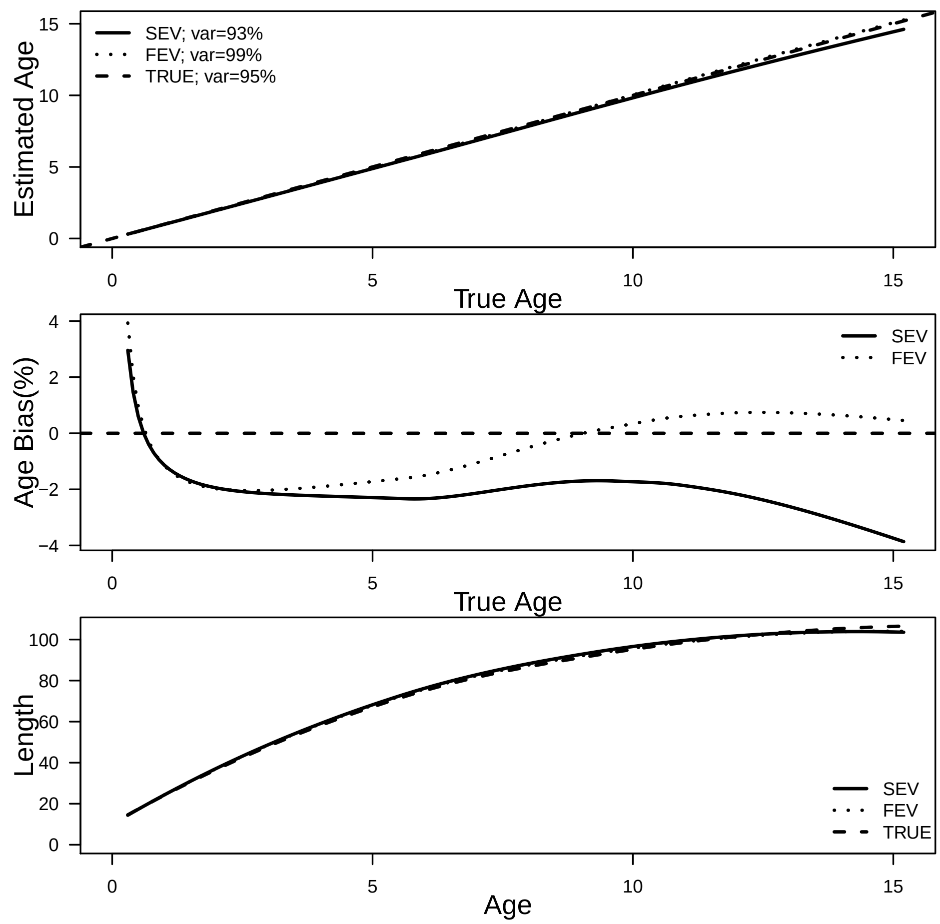

Figure 3.9: Some model diagnostics from Example 2. Top panel: Smoothed model predicted ages versus true simulation values. Middle panel: Bias in model predicted ages, in percent of true values. Bottom panel: Smoothed lengths with ME as functions of true and model predicted ages.

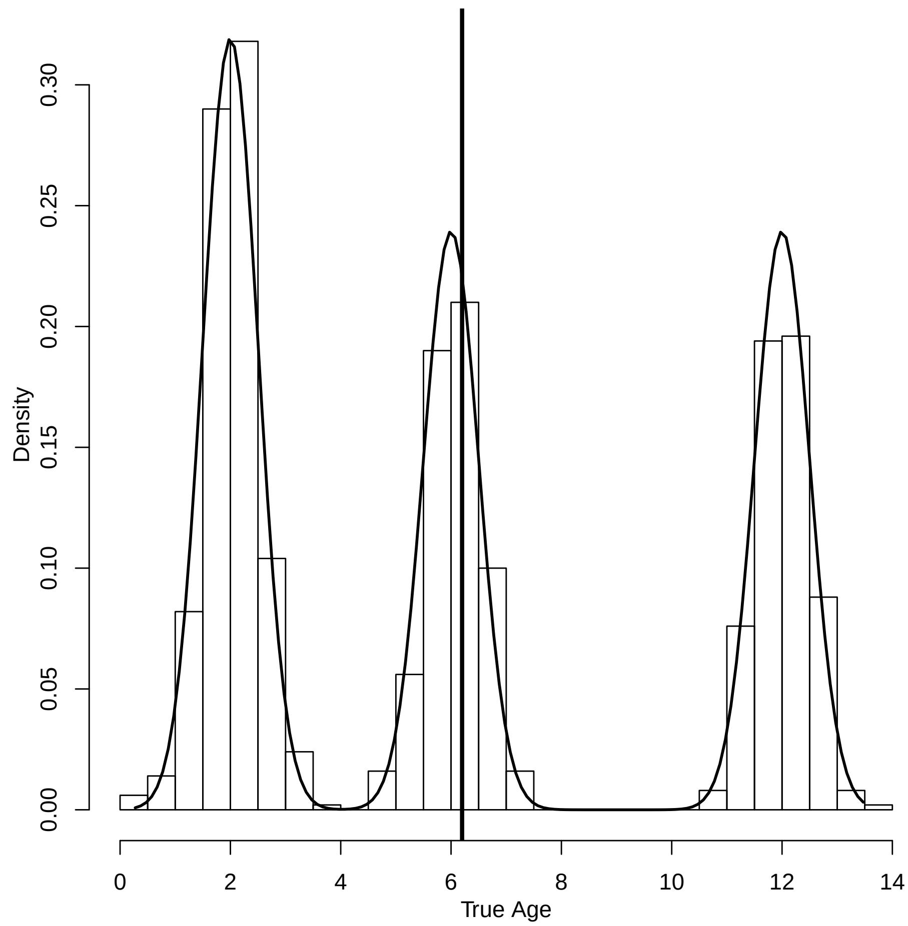

Figure 3.10: Distribution of $N = 1000$ true ages for Example 3.

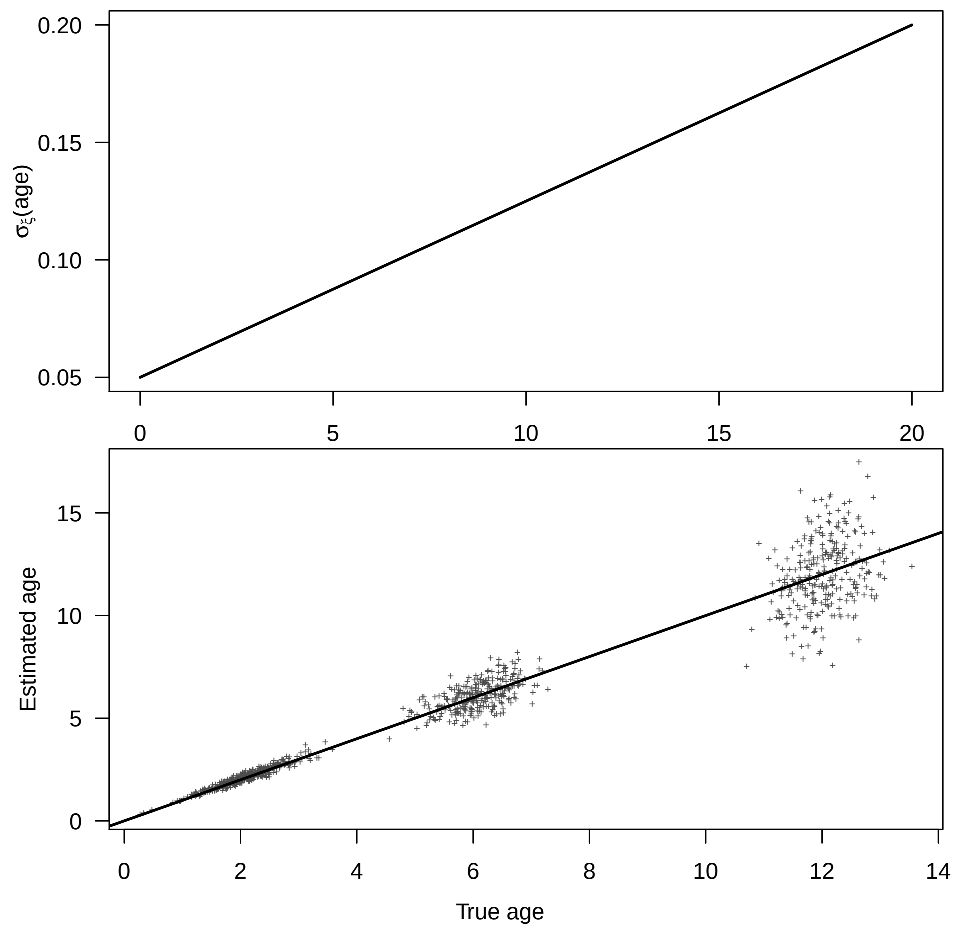

Figure 3.11: Top panel: Age ME coefficient of variation as a function of true age. Bottom panel: Estimated ages versus true ages and the 1-1 reference line for Example 3.

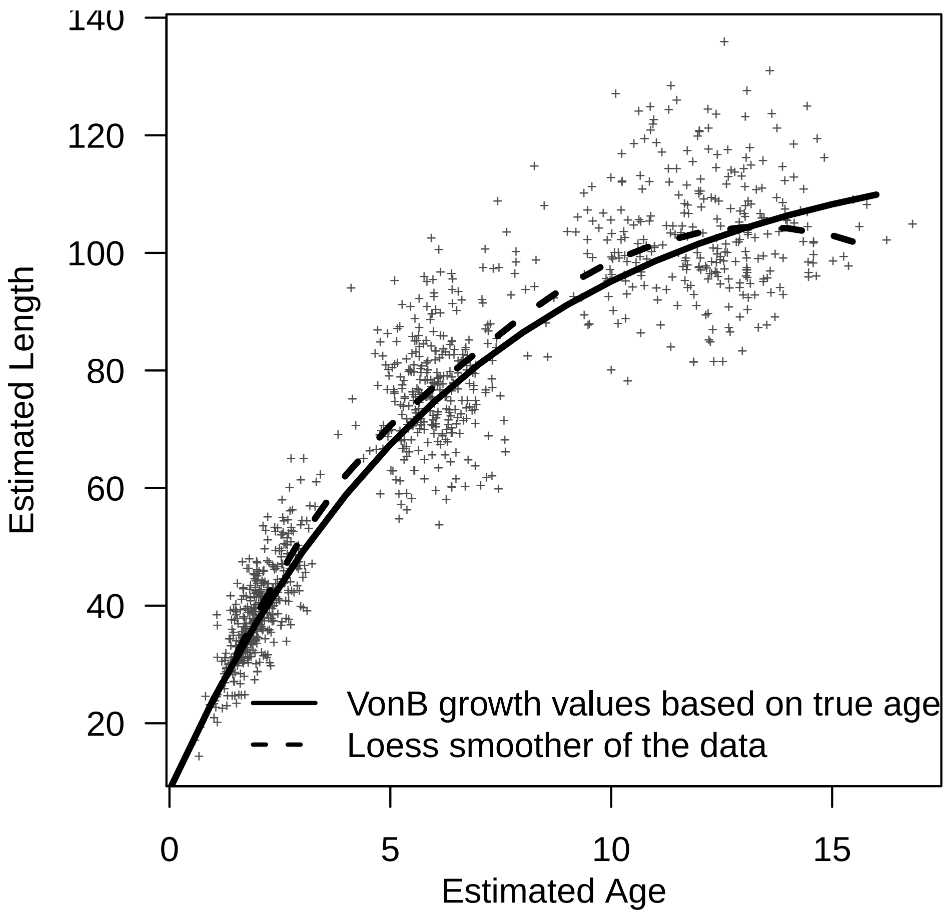

Figure 3.12: Length versus age for Example 3. Both variables are measured with error.

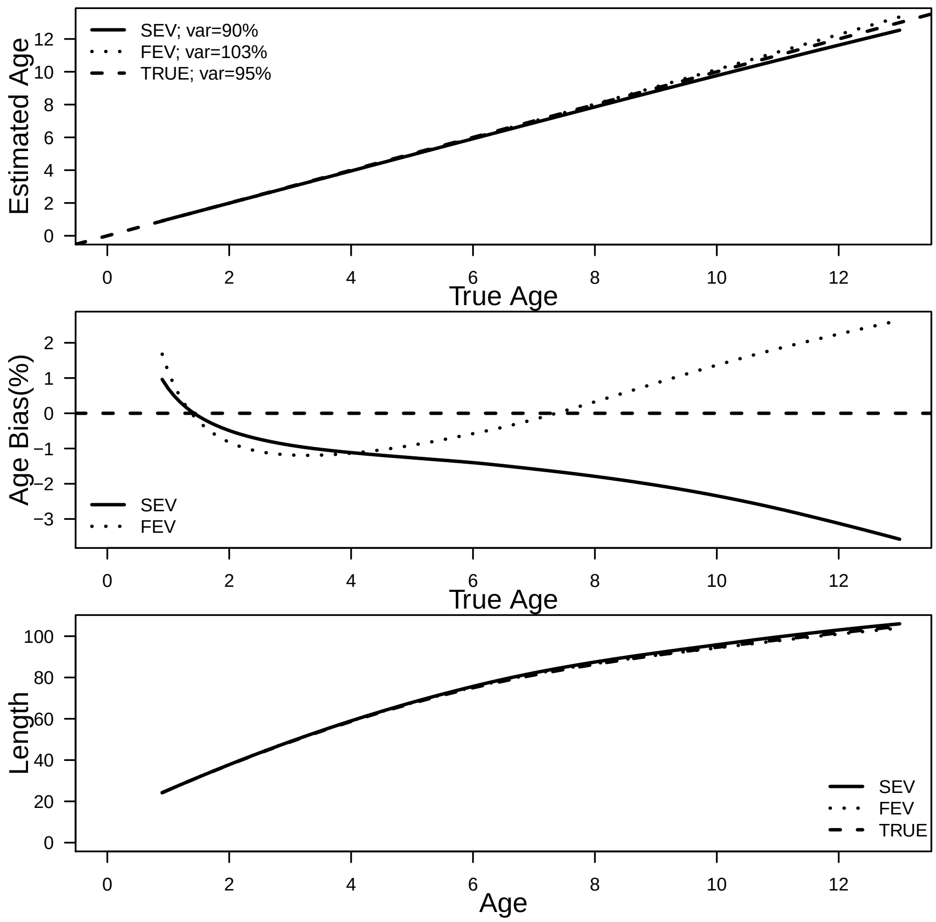

Figure 3.13: Some model diagnostics from Example 3. Top panel: Smoothed model predicted ages versus true simulation values. Middle panel: Bias in model predicted ages, in percent of true values. Bottom panel: Smoothed lengths with ME as functions of true and model predicted ages.

3.7 Tables

n_h	Freq
1	15
2	32
3	58
4	54
5	49
6	19
7	13
8	9
9	1

Table 3.1: Frequency of sample size per stratum, n_h.

		SEV				FEV			
Parameter	Value	est	CV(%)	L95	U95	est	CV(%)	L95	U95
l_∞	120.000	123.631	2	119.651	127.743	114.092	2	110.719	117.567
k	0.150	0.141	4	0.131	0.153	0.164	4	0.153	0.176
t_o	-0.500	-0.593	14	-0.757	-0.428	-0.495	15	-0.641	-0.348
α	7.00	6.926	5	6.318	7.593	-	-	-	-
β	1.00	1.013	5	0.920	1.116	-	-	-	-
σ_∞	0.100	0.089	9	0.074	0.107	0.102	8	0.087	0.118
σ_k	0.100	0.120	13	0.093	0.153	0.084	25	0.051	0.138
$\sigma_{o\xi}$	0.050	0.056	16	0.041	0.077	0.049	18	0.034	0.070
$\sigma_{1\xi}$	0.008	0.006	23	0.004	0.009	0.007	19	0.005	0.011

Table 3.2: Parameter estimates, coefficients of variation (CV), and 95% confidence intervals (L,U) for Example 1, using a data set of N=1000.

Parameter	Value	SEV				FEV			
		est	CV(%)	L95	U95	est	CV(%)	L95	U95
l_∞	120.000	117.333	1	115.267	119.436	115.084	1	113.084	117.120
k	0.150	0.157	2	0.152	0.162	0.161	2	0.156	0.166
t_o	-0.500	-0.474	3	-0.497	-0.451	-0.472	2	-0.495	-0.449
α	-	0.989	4	0.916	1.069	-	-	-	-
β	-	6.263	5	5.662	6.926	-	-	-	-
σ_∞	0.100	0.101	6	0.091	0.113	0.101	6	0.090	0.112
σ_k	0.100	0.098	8	0.083	0.115	0.093	9	0.078	0.110
$\sigma_{o\xi}$	0.050	0.057	8	0.049	0.067	0.055	8	0.047	0.064
$\sigma_{1\xi}$	0.008	0.006	11	0.005	0.007	0.006	10	0.005	0.008

Table 3.3: Parameter estimates, coefficients of variation (CV), and 95% confidence intervals (L,U) for Example 2, using a data set of N=1000.

Parameter	Value	SEV				FEV			
		est	CV(%)	L95	U95	est	CV(%)	L95	U95
l_∞	120.000	121.251	1	118.886	123.663	116.599	1	114.225	119.024
k	0.150	0.149	2	0.143	0.154	0.158	2	0.152	0.164
t_o	-0.500	-0.517	5	-0.565	-0.469	-0.491	5	-0.538	-0.444
α	-	1.973	4	1.816	2.143	-	-	-	-
β	-	3.132	5	2.847	3.445	-	-	-	-
σ_∞	0.100	0.095	6	0.085	0.107	0.097	6	0.087	0.109
σ_k	0.100	0.074	13	0.057	0.096	0.060	21	0.040	0.089
$\sigma_{o\xi}$	0.050	0.050	9	0.042	0.060	0.045	10	0.037	0.054
$\sigma_{1\xi}$	0.008	0.007	12	0.005	0.008	0.008	11	0.006	0.010

Table 3.4: Parameter estimates, coefficients of variation (CV), and 95% confidence intervals (L,U) for Example 3, using a data set of N=1000.

Parameter	Value	SEV		FEV	
		Mean est.	RMSE	Mean est.	RMSE
l_∞	120.000	120.68	21.338	111.145	107.572
k	0.150	0.148	0.063	0.173	0.283
t_o	-0.500	-0.526	0.924	-0.421	1.324
α	7.00	7.087	3.677	-	-
β	1.00	0.989	0.538	-	-
σ_∞	0.100	0.101	0.084	0.110	0.147
σ_k	0.100	0.083	0.323	0.037	0.877
$\sigma_{o\xi}$	0.050	0.058	0.133	0.042	0.156
$\sigma_{1\xi}$	0.008	0.006	0.026	0.009	0.022

Table 3.5: Mean parameter (Mean est.) estimates and root mean square error (RMSE) for Example 1, using 1000 data sets of N=1000.

Parameter	Value	SEV		FEV	
		Mean est.	RMSE	Mean est.	RMSE
l_∞	120.000	118.21	22.119	115.752	52.261
k	0.150	0.156	0.067	0.160	0.122
t_o	-0.500	-0.474	0.304	-0.470	0.382
α	-	0.972	-	-	-
β	-	6.142	-	-	-
σ_∞	0.100	0.099	0.062	0.099	0.072
σ_k	0.100	0.096	0.096	0.0903	0.158
$\sigma_{o\xi}$	0.050	0.053	0.055	0.0502	0.051
$\sigma_{1\xi}$	0.008	0.006	0.019	0.007	0.014

Table 3.6: Mean parameter (Mean est.) estimates and root mean square error (RMSE) for Example 2, using 1000 data sets of N=1000.

		SEV		FEV	
Parameter	Value	Mean est.	RMSE	Mean est.	RMSE
l_∞	120.000	119.395	14.574	115.360	57.353
k	0.150	0.152	0.043	0.160	0.131
t_o	-0.500	-0.480	0.360	-0.471	0.561
α	-	1.916	-	-	-
β	-	3.250	-	-	-
σ_∞	0.100	0.099	0.064	0.101	0.085
σ_k	0.100	0.096	0.127	0.084	0.307
$\sigma_{o\xi}$	0.050	0.055	0.094	0.049	0.078
$\sigma_{1\xi}$	0.008	0.006	0.022	0.007	0.016

Table 3.7: Mean parameter (Mean est.) estimates and root mean square error (RMSE) for Example 3, using 1000 data sets of N=1000.

Chapter 4

Growth Model Including Sampling Technique Effect

In this Chapter we extend the growth model to account for variations that may be caused by the way fish are caught and sampled for length and age measurements. In Chapters 2 and 3, we assumed that fish are randomly sampled from the population, even though in practice fish are caught using sampling trawls that target and capture fish by size and species during the fishing operation. Thus, all sampling trawls are size-selective and species-selective to varying degrees (Walsh 1996). In the data we examine, all fish in research survey trawl catches were first measured for length and then assigned to 3 cm length classes. A fixed number of fish were selected from the length classes for age determination. This is known as size-stratified age-sampling. The second sample provides the length-at-age data that is used to fit the VonB growth model but the fraction sampled for ages in each length class is important as well.

Size (length) selective samples are biased towards the size selectivity of the fishing gear used to catch the fish and this produced biased estimates of the VonB growth parameters (Candy et al., 2007). Goodyear (1995) simulated length-at-age data to show that length selectivity introduce bias into estimates of the VonB growth parameters and this result was proved analytically by Candy (2005). Thus to get an unbiased and precise estimates of the VonB growth parameters for length-at-age data, the VonB model has to be extended to take into account:

- Size selectivity of the fishing gear used to catch the fish.
- Size-stratified age sampling (Candy et al., 2007).

Candy et al. (2007) used a step function to describe the size selectivity function and account for the size-stratified age sampling by using sampling probabilities. We used a similar approach to Candy et al. (2007), however, we assume the selectivity function has a linear logistic form.

4.1 Model

4.1.1 Size-Selective Capture

Let C be a random variable that indicates whether a fish is caught or not in a sample. Thus

$$C = \begin{cases} 1 & \text{if a fish is caught,} \\ 0 & \text{otherwise.} \end{cases}$$

If Y^T is the true length of the fish, then the capture probability is

$$Pr(C = 1|Y^T = y) = s(y),$$

where $s(y)$ is the probability a fish is caught given length y. We assume this selectivity function is known from external sources. The selectivity function for commercial fishery data may be quite complicated and have substantial effects on growth data. However, the research survey data used in this study does not have these issues. We assume the selectivity function has a linear logistic form

$$logit[s(y)] = \beta_0 + \beta_1 y \tag{4.0}$$

such that the size at 50% selection is 10 cm and the size at 95% selection is 20 cm. That is,

$$\begin{aligned} \beta_0 + \beta_1 10 &= log(1) = 0 \\ \beta_0 + \beta_1 20 &= log(0.95/0.05) = 2.94. \end{aligned}$$

Hence, $\beta_0 = -2.94$, and $\beta_1 = 0.294$ and Figure 4.1 shows the selectivity function for fish length 1 cm to 30 cm. We let x_{50} be the size of the fish at 50% selection and x_{95} the size at 95% selection. In general, if the selectivity $s(x_{50}) = 0.5$ and $s(x_{95}) = 0.95$, then

$$\begin{aligned} \beta_0 + \beta_1 x_{50} &= 0 \\ \beta_0 + \beta_1 x_{95} &= 2.94. \end{aligned}$$

Thus, $\beta_1 = 2.94/(x_{95} - x_{50})$ and $\beta_0 = 2.94 x_{50}/(x_{95} - x_{50})$. The selectivity of most fish captured in the survey will be 1 when $x_{50} = 10$ and $x_{95} = 20$.

The age of the i^{th} fish caught including measurement error (ME) is

$$A_i = A_i^T + \xi_i \qquad \text{where} \qquad \xi_i \overset{iid}{\sim} N\left\{0, \sigma_\xi^2(A_i^T)\right\}.$$

The pdf of the age for the catch is

$$\Pr(A_i = a \mid \gamma, \beta) = \frac{a^{\gamma-1}\exp(-a/\beta)}{\beta^\gamma \Gamma(\gamma)} \tag{4.1}$$

where $\gamma = (A_i^T/\sigma_\xi(A_i^T))^2$ and $\beta = \sigma_\xi^2(A_i^T)/A_i^T$.

The length of the i^{th} fish caught including measurement error (ME) is

$$Y_i = Y_i^T + \varepsilon_i \qquad \text{where} \qquad \varepsilon_i \overset{iid}{\sim} N(0, \sigma_\varepsilon^2).$$

The pdf of the length for the catch is

$$f(Y_i = y_i \mid Y_i^T = y_i^T, \sigma_\varepsilon) = \frac{1}{\sigma_\varepsilon\sqrt{2\pi}} \exp\left[-\frac{\{y_i - y_i^T\}^2}{2\sigma_\varepsilon^2}\right] \tag{4.2}$$

where

$$Y_i^T = L_{\infty i}\left(1 - e^{-K_i A_i^T}\right).$$

The probability that the i^{th} fish caught has measured age a and length y, is

$$\Pr(A_i = a, Y_i = y) = \iint \Pr(A_i = a|A_i^T)\Pr(Y_i = y|Y_i^T)\Pr(A_i^T, Y_i^T|C = 1)\partial A_i^T \partial Y_i^T. \quad (4.3)$$

The probability that a fish with true age a and true length y caught is derived as follows

$$\begin{aligned}\Pr(A_i^T, Y_i^T|C_i = 1;\theta) &= \frac{\Pr(A_i^T, Y_i^T, C_i = 1;\theta)}{\Pr(C_i = 1;\theta)} \\ &= \frac{\Pr(C_i = 1|A_i^T, Y_i^T;\theta)\Pr(A_i^T, Y_i^T;\theta)}{\Pr(C_i = 1;\theta)} \\ &= \frac{s(Y_i^T)\Pr(Y_i^T|A_i^T;\theta)\Pr(A_i^T;\theta)}{\Pr(C_i = 1;\theta)}\end{aligned} \quad (4.4)$$

$$\Pr(A_i^T, Y_i^T|C_i = 1;\theta) = \begin{cases} \dfrac{s\left\{Y(A_i^T, L_{\infty i}, K_i)\right\}\Pr(A_i^T;\theta)}{\int s\left\{Y(a, L_{\infty i}, K_i)\right\}\Pr(A_i^T = a;\theta)\partial a} & \text{if } Y_i^T = Y(A_i^T, L_{\infty i}, K_i), \\ 0 & \text{otherwise.} \end{cases} \quad (4.5)$$

The denominator in Eqn. 4.5 does not have a closed form solution so we approximate as

$$\int s\left\{Y(a, L_{\infty i}, K_i)\right\}\Pr(A_i^T = a;\theta)\partial a \simeq \int_{s_{50,i}}^{\infty} \Pr(A_i^T = a;\theta)\partial a = \Pr(A_i^T > s_{50}). \quad (4.6)$$

For the survey data and the species we examine, the selectivity is near to 1 for most ages so including this is not important. However, this will be much more important for commercial data or for species with slower growth rates.

4.1.2 Size-Selective Sampling

The sampling scheme used for the data we investigate involves measuring the length of all fish caught in research tows and subsampling a portion of the catch to determine ages. It is usually too costly and time-consuming to determine the age of all fish caught. A fixed number of fish are selected from the length classes to measure for age. This is a form of two-phase stratified random sampling (TPSRS), where the stratification depends on the response variable length. We let h be the index of stratum and we assume there is a total of H strata.

Let N_h denote the total number of fish caught in length class h and let n_h be the number of fish that were randomly subsampled to determine age. The total number of fish caught is $\sum_h N_h = N$ and the total sampled for age determination is $\sum_h n_h = n$. The probability that a fish in length class h is sampled for age is $\pi_h = n_h/N_h$.

If all fish were measured for age, then the total SEV log-likelihood would be

$$l_{SEV}^T(\theta; D) = \sum_{h=1}^{H} \sum_{i=1}^{N_h} l_{hi}(\theta) = N\bar{l_p}(\theta)$$

where

$$l_{hi}(\theta) = log\{Pr(A_i = a_i, Y_i = y_i; \theta)\},$$

and $\bar{l}_p(\theta)$ is the average log-likelihood for the entire phase 1 sample. Estimation and statistical inference for θ would be based on this total log-likelihood. When TPSRS is employed and the subsampling probabilities are different for each stratum (i.e., length class); the log-likelihood for the sample responses is,

$$l_{SEV}^{S}(\theta; D) = \sum_{h=1}^{H} \sum_{i=1}^{n_h} l_{hi}(\theta) = N\bar{l}_S(\theta).$$

This may produce biased parameter estimation and inference. A problem is that $E_\pi\{\bar{l}_S(\theta)\} \neq \bar{l}_p(\theta)$ where the expectation is with respect to the probability a fish is sampled (π), similar to the common design-based approach to statistical inference with survey sampling. Let I_{hi} be an indicator variable that is

$$I_{hi} = \begin{cases} 1 & \text{if fish } i \text{ is selected for age sampling,} \\ 0 & \text{otherwise.} \end{cases}$$

Note that $E_\pi(l_{hi}) = \pi_h$. Then,

$$l_{SEV}^{S}(\theta; D) = \sum_{h=1}^{H} \sum_{i=1}^{N_h} I_{hi} l_{hi}(\theta)$$

and

$$E_\pi\{\bar{l}_S(\theta)\} = \sum_{h=1}^{H} \sum_{i=1}^{N_h} \pi_h l_{hi}(\theta) \neq \bar{l}_p(\theta).$$

This can lead to biased inferences depending on the magnitude of bias for $\bar{l}_S(\theta)$. Using the Horvtiz-Thompson (HT) procedure from survey sampling (Horvitz & Thompson, 1952), the bias of $\bar{l}_S(\theta)$ can be removed by weighting with the inverse selection probability, which gives the HT design-weighted log-likelihood as

$$l_{SEV}^{HT}(\theta; D) = \frac{n}{N} \sum_{h=1}^{H} \pi_h^{-1} \sum_{i=1}^{n_h} l_{hi}(\theta) = n\bar{l}_{HT}(\theta), \tag{4.7}$$

where $E_\pi\{\bar{l}_{HT}(\theta) = \bar{l}_p(\theta)$. Note that if fish were sampled in proportion to the total number in each length class, so that $\pi_h = \frac{n}{N}$, then $l^{HT}_{SEV}(\theta; D) = l^{S}_{SEV}(\theta; D)$ and no sampling adjustment is required. Such a sampling design is basically equivalent to simple random sampling.

4.2 Simulation studies.

We investigate the performance of the HT-adjusted SEV method in estimating the VonB growth parameters separating the ME in age and BT variation. We also compared the SEV method with the commonly used nonlinear least squares (NLS) method. Simulations were conducted in R. Model performance was measured using:

- coefficient of variation (CV) and confidence intervals (CI) for a single data set of 937 length-at-age measurements.
- mean square error (RMSE) for 1000 data sets of 937 length-at-age measurements.

4.2.1 Data generation

We randomly generated $N = 10000$ ages from a gamma distribution with $\alpha = 7$ and $\beta = 1$ so that $E(A) = 7 = Var(A)$. Lengths were generated from the VonB model, with $l_\infty = 120$, $\sigma_\infty = 0.120$, $k = 0.2$, and $\sigma_k = 0.1$ in Equation 3.3. We set the age when length is zero to be $t_o = -0.070$. We set $\sigma_\varepsilon = 0.05$, and for age ME we used $\sigma_{o\xi} = 0.05$

and $\sigma_{1\xi} = 0.0075$. We stratified the data by 1 cm length classes and sampled a maximum of 10 fish per length stratum. This produced 937 length and age measurements which is a sample size similar to Chapter 3 examples. The data are shown in Figure 4.2 (black symbols), along with the rest of the simulated unsampled measurements (red symbols). The loess smoother of the sampled growth data clearly indicates that this data provides biased estimates of the growth curve if the size-selective sampling design is not included. The problem is that the fish that reach the largest sizes are usually all sampled because there are so few of them (usually less than 10 per 1 cm length class), whereas slower growing fish are subsampled. Hence, large-sized, older fish are over represented in the two-phase sample. If stratified random sampling with proportional allocation was used then larger fish would not be over-represented. We also computed the loess smoother using second phase sampling weights for each observation, $w_i = nN_h/n_hN$ (weight for the i^{th} observation), and the resulting smoother was much closer to the true population growth function. Note that $E_\pi(\sum_{i=1}^{n} w_i) = n$. We repeated this procedure 1000 times to produce 1000 data sets consisting of 937 length-at-age measurements. We used the same weighting approach and nonlinear least squares (nls) procedure in R, assuming only measurement error in length to estimate l_∞ and k.

4.3 Result

The NLS and SEV methods converged in all 1000 data sets. The π-weighted and un-weighted NLS results for a single data set of size 937 (Table 4.1) show equally unreliable results because profile likelihood CI from both methods did not contain the true parameter values. The average RMSE of the NLS π-weighted and un-weighted estimates for 1000 data sets is shown in Table 4.3. The high number for the RMSE is not surprising because the estimation model is seriously misspecified in this case. Note that the commonly used NLS model does not properly account for variation in the data.

It is remarkable and quite unexpected how well the unweighted SEV model estimates l_∞ and k (Table 4.2; Figure 4.3). The unweighted SEV estimates are fairly close to their true values and their 95% CIs contain the true values. The π-weighted estimates are less reliable; the CIs for l_∞ and k do not include the true values. Although these results are only for one simulated data set, they may suggest that the unweighted SEV model is able to accommodate two-phase sampling effects in the distributions of L_∞ and K. This is illustrated in Figure 4.4 for a single data set. The two-phase length-stratified sampling for age induces an age dependent effect in the conditional mean of L_∞ and K that the π-weighted and unweighted SEV models are able to predict. This age dependent effect is not included in the first phase of sampling (Figure 4.5), as we expect, because our simulation model randomly assigned $L_{\infty i}$ and K_i to a fish independent of its size. Note that the range of ages in Figure 4.5 is much smaller than in Figure 4.4, because the per-stratum sample size for length is much larger than for the ages-at-length sample size (Figure 4.6). Usually there are less than five phase-two age samples per stratum while there is 30 or more phase-one length samples per stratum. The average age of these phase-one samples is much less variable than the average age of the phase-two samples.

The mean of 1000 estimated parameters and their RMSE for the 1000 data sets are shown in Table 4.4 for the SEV approach. The SEV approach was shown to be more reliable than the NLS (compare RMSE in Tables 4.3 and 4.4); indicating the effect of the BT and ME in age variation on the growth parameters. The SEV π-weighted and unweighted estimates are all close to their population true values, however, the unweighted estimates have lower RMSE than π-weighted estimates (see Table 4.4) even though the π_i's are related to the response variable (length) which is also prone to error. This is surprising because Rao et al. (1999) found that the HT weighted approach can lead to unrealistic results if the basic design weights (π_i) are unrelated to the response variable (length in this case). Other accurate weighting approaches have been proposed by Rao (1966), Hajek (1971), and others. However, these approaches seem difficult to implement in our situation, having measurement error in both length and age, and between individual variability. We do not pursue the HT π-weighted SEV. However, why the unweighted SEV performed better than HT π-weighted SEV seem like a useful area for future research.

4.4 Figures

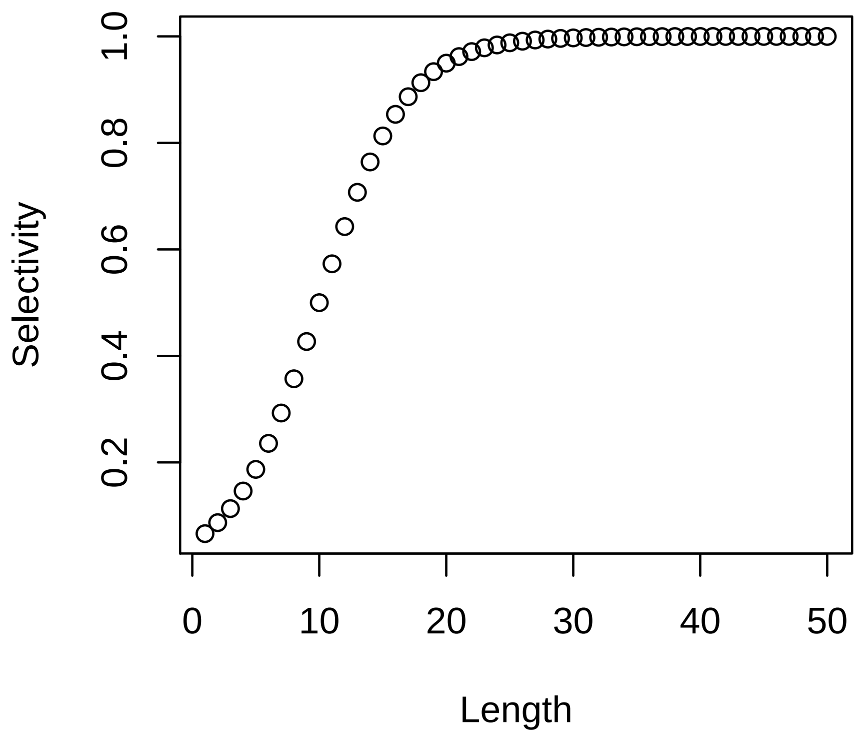

Figure 4.1: Selectivity function for fish lengths 1 cm to 30 cm with $\beta_0 = -2.94$ and $\beta_1 = 0.294$.

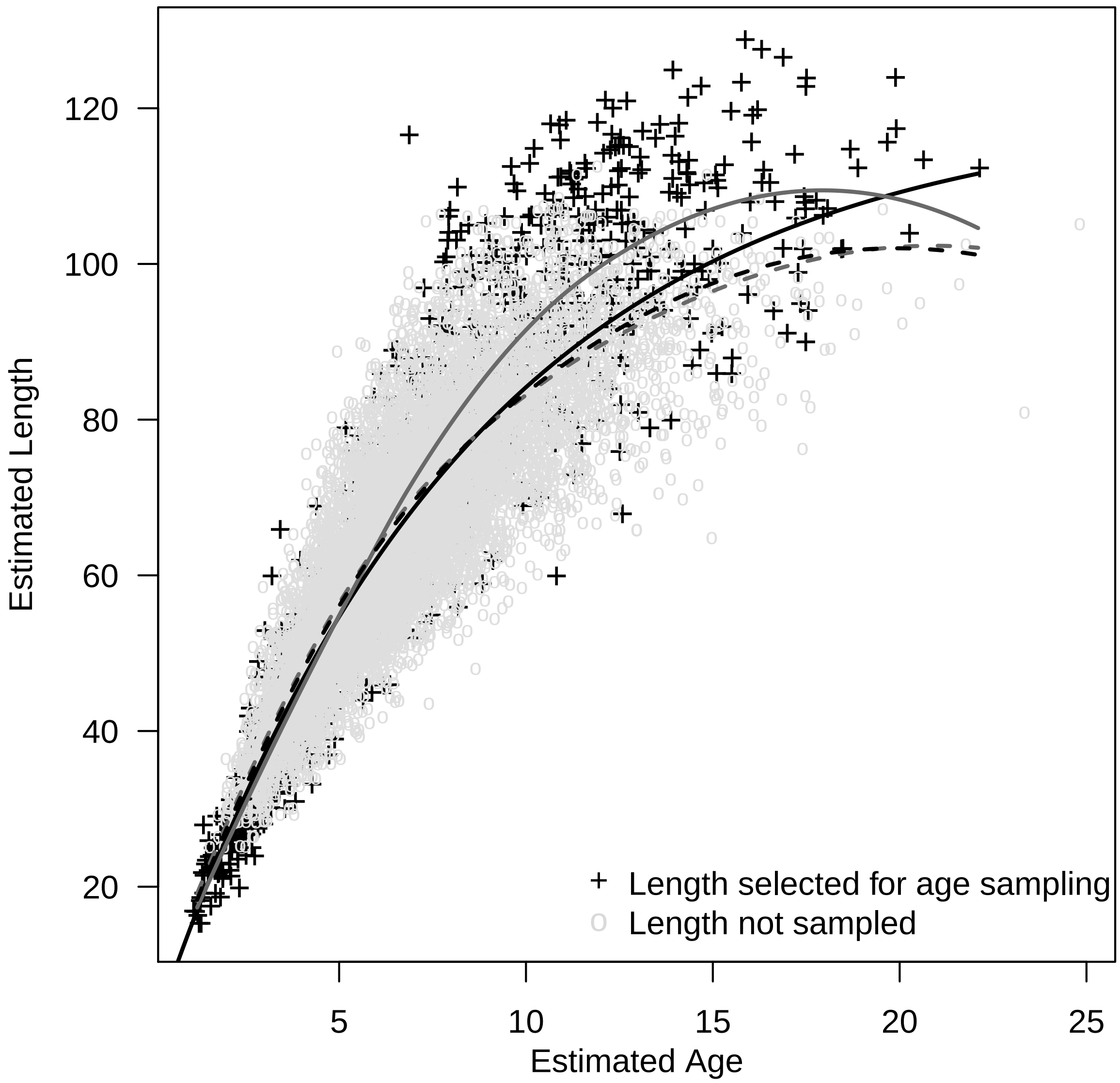

Figure 4.2: Illustration of length-stratified data. The solid black line is the true Von Bertalanffy population growth model. The solid grey line is a loess smoother of the sampled growth data. The black dashed line is a loess smoother of both the sampled and un-sampled growth data and the grey dashed line is the design-weighted loess smoother of the sample data.

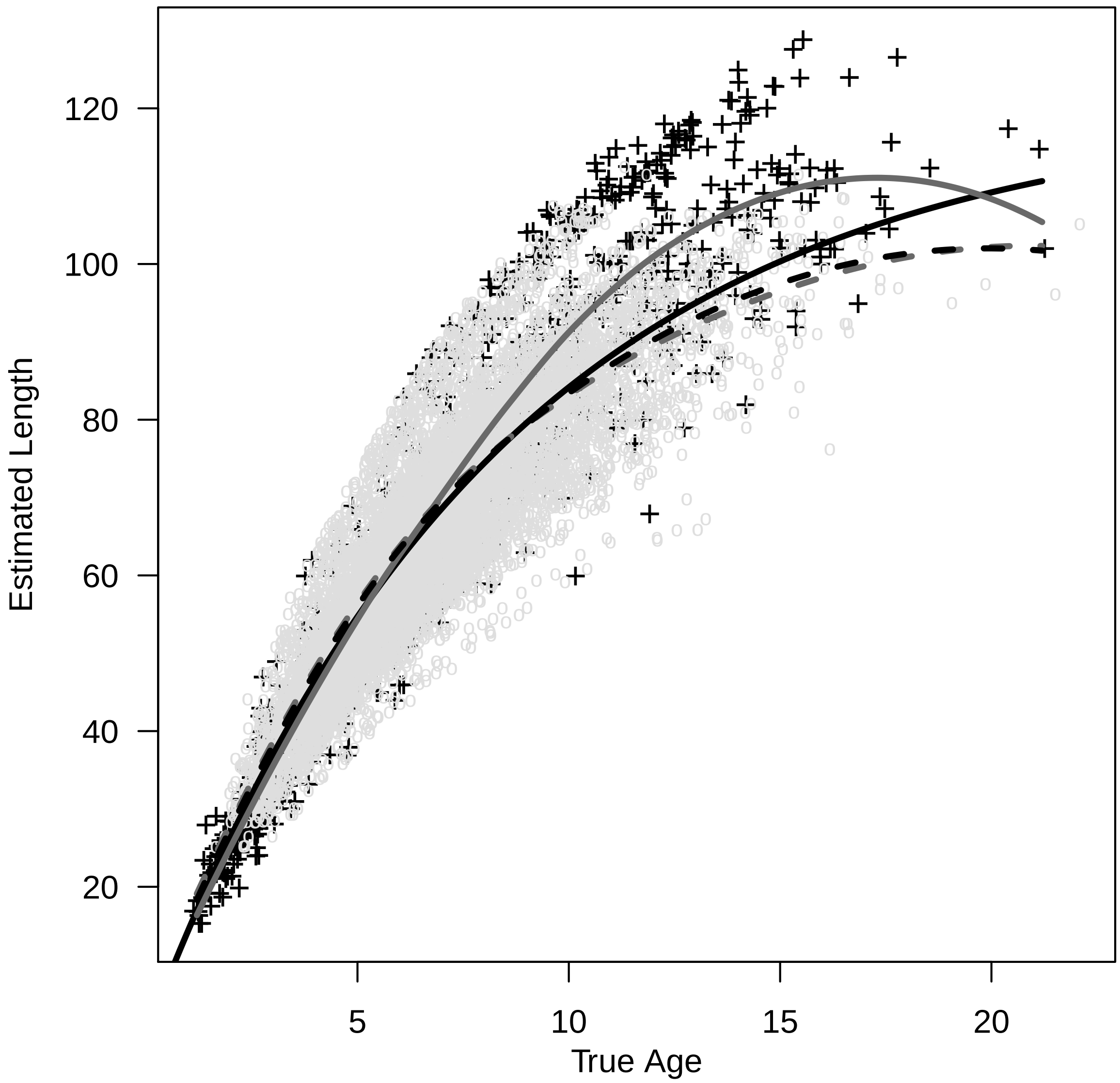

Figure 4.3: Von Bertalanffy growth model fits to the length-stratified simulated example data. The +s indicate lengths and ages for sampled fish. The solid black line is the true population growth model. The black dashed line is the π-weighted SEV estimate of the marginal growth function and the grey dashed line is the un-weighted estimated function. The o's indicate growth function estimates based on π-weighted and un-weighted nonlinear least squares based only on length observation errors (LOE).

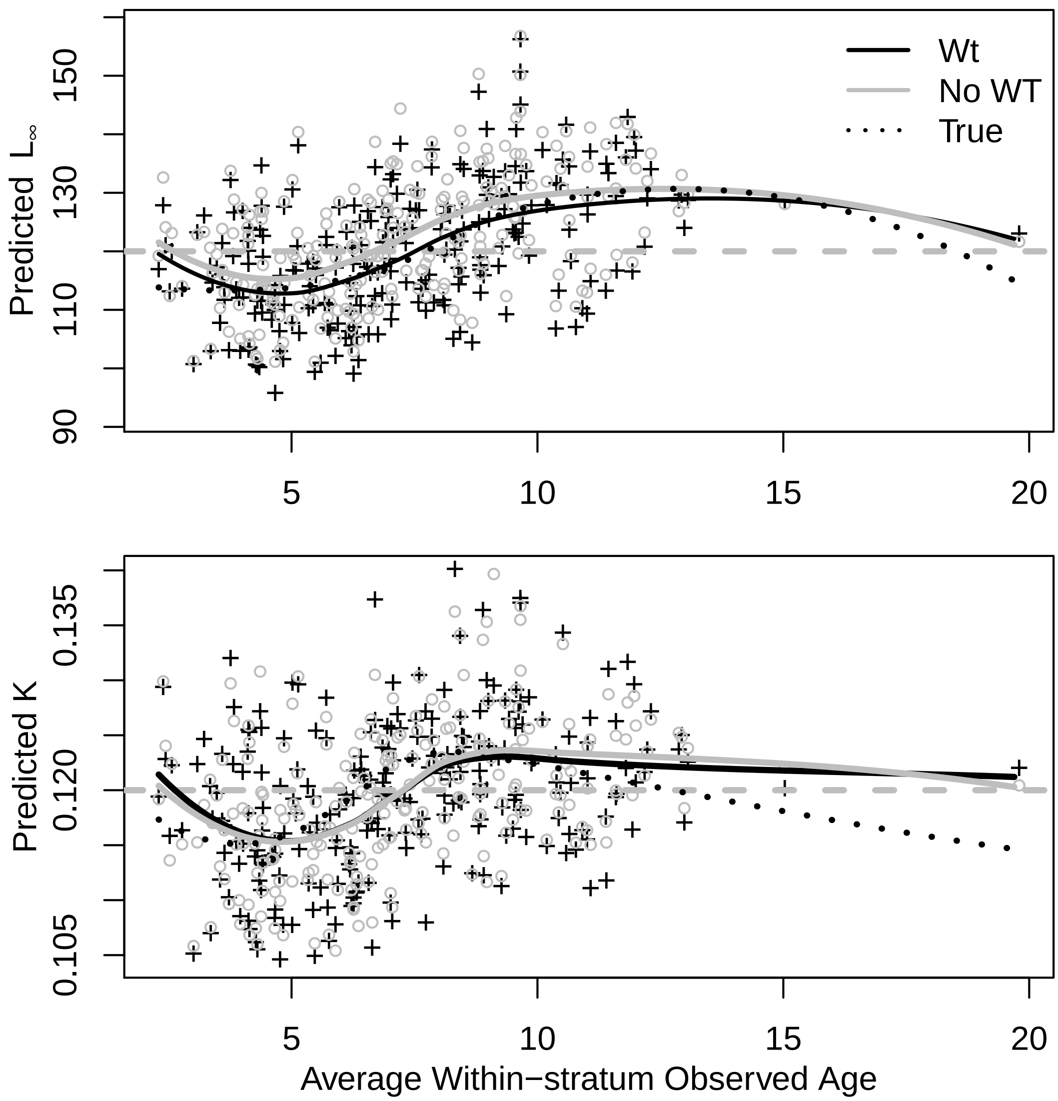

Figure 4.4: Unweighted (black +'s) and π-weighted (grey o's) SEV predictions of individual Von Bertalanffy growth parameters versus the average age of fish for each stratum in the phase 2 age-length sample.

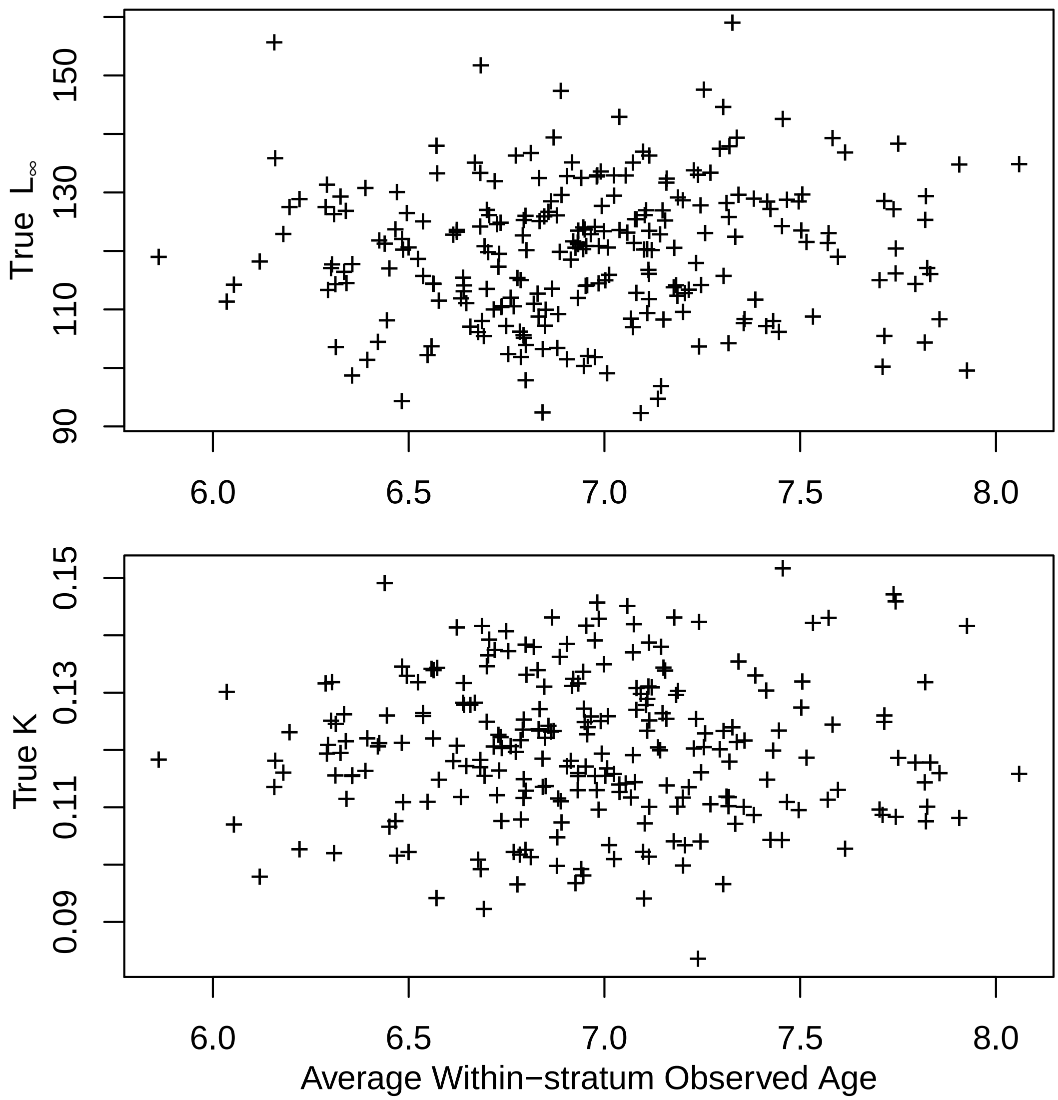

Figure 4.5: Simulated individual Von Bertalanffy growth parameters versus the average age of fish for each stratum in the phase 1 length sample.

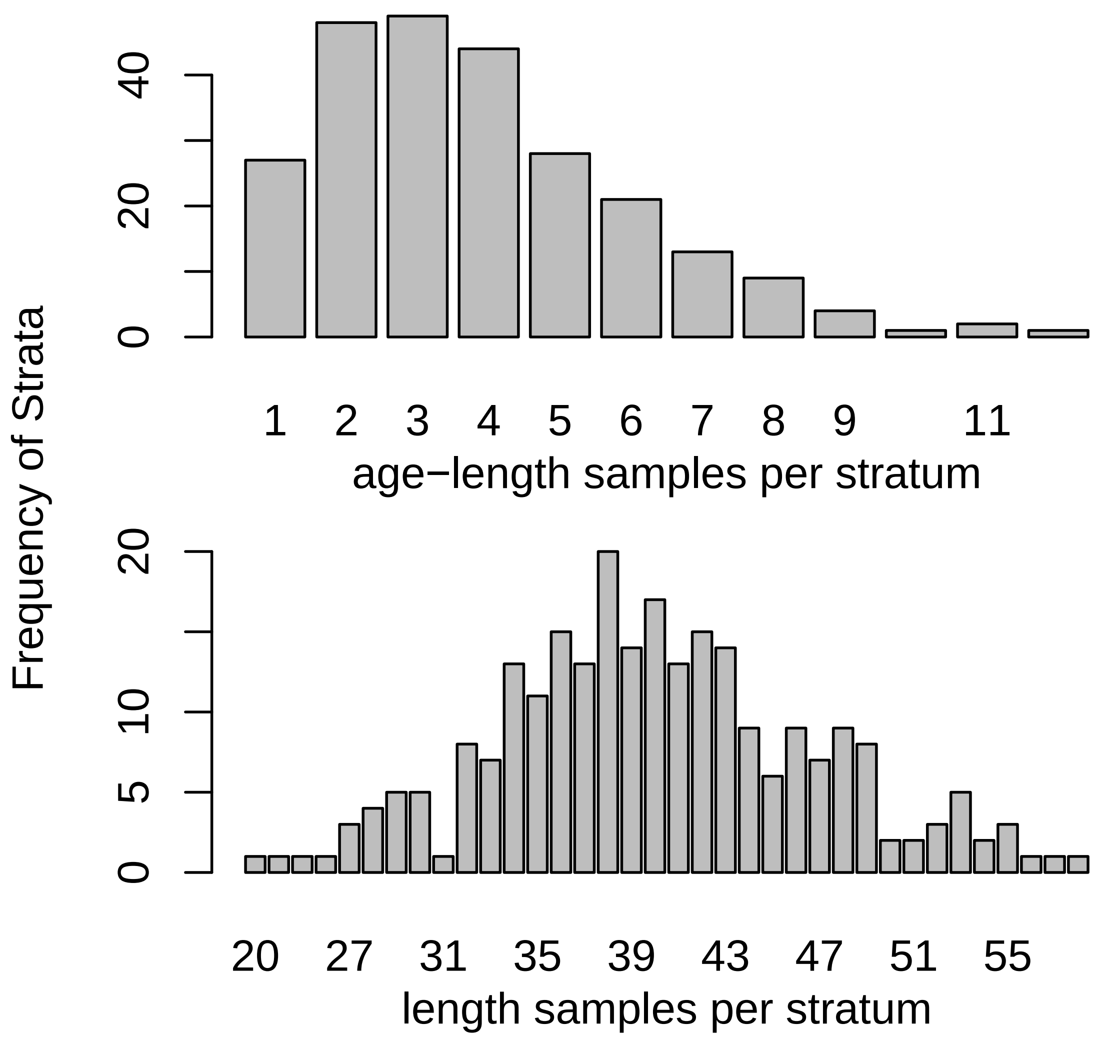

Figure 4.6: Frequency of phase 2 age-length samples per stratum (top panel) and phase 1 length samples per stratum (bottom panel).

4.5 Tables

		π-weighted				unweighted			
Parameter	Value	est	CV(%)	L95	U95	est	CV(%)	L95	U95
l_∞	120.000	107.957	3	104.045	112.733	134.078	3	129.444	139.164
k	0.120	0.145	8	0.131	0.155	0.108	6	0.102	0.115
t_o	-0.070	0.000	∞	-0.220	-	0.000	∞	-0.054	-

Table 4.1: Nonlinear least squares parameter estimates, coefficients of variation (CV), and 95% confidence intervals (L,U) for a length-stratified sample with a single data set of size 937.

		π-weighted				unweighted			
Parameters	Value	estimate	CV(%)	L95	U95	estimate	CV(%)	L95	U95
l_∞	120.000	115.122	2	111.363	119.009	117.498	2	113.839	121.274
k	0.120	0.131	3	0.123	0.140	0.124	3	0.118	0.131
t_o	-0.070	0.116	46	0.012	0.220	0.018	176	-0.043	0.079
α		3.169	5	2.888	3.476	3.133	5	2.869	3.422
β		2.315	5	2.087	2.568	2.342	5	2.125	2.580
σ_∞	0.100	0.104	9	0.087	0.124	0.099	9	0.084	0.118
σ_k	0.100	0.087	26	0.052	0.145	0.091	19	0.063	0.131
$\sigma_{o\xi}$	0.050	0.054	11	0.044	0.068	0.050	12	0.039	0.062
$\sigma_{1\xi}$	0.008	0.005	17	0.004	0.007	0.007	13	0.005	0.008

Table 4.2: SEV parameter estimates, estimates, coefficients of variation (CV), and 95% confidence intervals (L,U) for a length-stratified sample with a single data set of size 937.

		π-weighted		unweighted	
Parameters	Value	Mean est.	RMSE	Mean est.	RMSE
l_∞	120.000	109.860	95.539	135.753	146.436
k	0.120	0.142	0.210	0.108	0.117
t_o	-0.070	-0.038	0.688	0.000	0.639

Table 4.3: Nonlinear least squares parameter estimates and root mean square error (RMSE) for a length-stratified sample with a 1000 data sets of size 937.

		π-weighted		unweighted	
Parameters	Value	Mean est.	RMSE	Mean est.	RMSE
l_∞	120.000	119.724	19.702	121.595	16.193
k	0.120	0.123	0.039	0.118	0.025
t_o	-0.070	0.007	0.544	-0.060	0.218
α	-	3.154	-	3.182	-
β	-	2.376	-	2.337	-
σ_∞	0.100	0.076	0.162	0.094	0.079
σ_k	0.100	0.105	0.227	0.104	0.138
$\sigma_{o\xi}$	0.050	0.058	0.055	0.053	0.038
$\sigma_{1\xi}$	0.008	0.005	0.017	0.006	0.009

Table 4.4: SEV parameter estimates and root mean square error (RMSE) for a length-stratified sample with 1000 datasets of size 937.

Chapter 5

Modelling growth of 2J3KL Northern cod, 2009-2013

The Northern cod (*Gadus morhua*) was the most abundant and most valuable groundfish stock in the Northwest Atlantic Ocean. However, in recent years they are only found in a few confined areas next to the north and east coasts of Newfoundland and Labrador. Northern cod once represented almost one-half of the total Canadian cod catch. It is easier and more profitable to catch than most other groundfish stocks. Thus Northern cod is a very important fish stock to the Department of Fisheries and Oceans (DFO) Canada and it remains one of the most valuable groundfish in Newfoundland and Labrador. This makes understanding their growth patterns a highly important subject, especially in the context of sustainable fishery management.

In this chapter, we apply the extended VonB growth model and unweighted SEV approach, discussed in Chapter 4, to the length-at-age cod data collected in Northwest Atlantic Fisheries Organization (NAFO) divisions 2J, 3K, and 3L (see Figure 1.4) during 2009 to 2013.

5.1 Sampling scheme

DFO conducts two research surveys (bottom trawl) per year off the coast of Newfoundland and Labrador. The autumn survey covers divisions 2J3KL (2J, 3K and 3L) and runs from October to December. The trawl surveys follow a stratified random sampling scheme where each NAFO division is divided into a certain number of strata. This is depicted in Figure 5.1, the grey lines in the map indicate the strata borders, which are largely based on ocean depth. Stratum size ranges from small to large. An observation from the sample is the number of cod caught from each NAFO division in one standardized research trawl tow during the year.

The sampling scheme used for the DFO data we investigate involves measuring the length of all fish caught in research tows for each year (2009-2013), classifying them into length classes (3 cm length class), which is our population sample. A portion of the population sample from each length class is subsampled to determine ages. A total of 60 fish per

3 cm length class for each NAFO division (2J, 3K and 3L) were sampled for age determination (Table 5.2). This produced a sample total (from 2009 - 2013) of 8055 length-at-age measurements, a sample similar to the simulated data in Chapter 4.

Figure 5.2 shows the number of aged cod aggregated into 3 cm length classes for population and sample lengths. The population and sample length frequencies are totalled over all sampling areas (2J, 3K and 3L) and fishing seasons (from 2009 to 2013). The largest sizes of fish are all sampled because there are so few of them (Figure 5.2 and Table 5.3), whereas slower growing fish are subsampled. The total number of fish caught by year and length class and the number that was sampled for age determination by year and length class are found in Table 5.3. Figure 5.3 is the length and age measurements of the sample data we investigated accumulated over all the years (2009 -2013); Figure 5.4 shows the length and age plot by year and NAFO division.

The data for each year and NAFO division are highly stratified, the number of strata for the years (2009 - 2013) ranges from 50 to 73. The maximum number of observations sampled from any given stratum is between 1 and 200 (Tables 5.4, 5.5, and 5.6). In almost all years, over 70% of the strata sampled contained more than 10 observations (see Tables 5.2, 5.4, 5.5, and 5.6).

5.2 Full model: all parameters different for each year and division.

We use the extended VonB growth model (Eqn. 4.1.1) developed in Chapter 4 to estimate the VonB growth parameter, BT variation, ME in age variation and the true age distribution of 2J3KL northern cod by year and NAFO division. Note that each NAFO division is divided into strata. Fish caught in different strata of the same NAFO division have varied growth rates, but we assume that the only variation within stratum is due to ME in length and age (ME in length is assumed known). This is represented in our full model; the estimates with their corresponding confidence intervals (CIs) are shown in Appendix A.1. We did not report the CIs of $\sigma_{1\xi}$ and σ_{∞} estimates. This is because the standard errors of the estimates that reached the lower bound are not reliable.

The full model estimates by year and NAFO division (see Appendix A.1) shows the VonB growth parameters (l_{∞}, k and t_o) are different for each division and year. However, the estimates of the ageing error slope and BT variance parameters ($\sigma_{1\xi}$ and σ_{∞}) are 0 for all years and divisions (Appendix A.1). Which suggest that these parameters are very small values across year and division. The estimate of the ageing error intercept variance parameter ($\sigma_{o\xi}$) is surprising. We were expecting $\sigma_{o\xi}$ to be the same across year and division since the ageing error depends on the individual reading the age. It appears there is confounding between $\sigma_{1\xi}$ and σ_{∞}; we will not investigate this confounding in this research.

However, this seem like a useful area for future research.

We calculated the negative log likelihood ratio (LLR), the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) for all models. However, the LLR test for model selection was employed since the reliability of the AIC and BIC for mixed-effect model selection is questionable. Some papers have proposed improvements in model selection techniques for mixed-effect models. This is beyond the scope of this report.

5.2.1 Reduced model (RM) 1: remove ageing error slope variance parameter ($\sigma_{1\xi}$) from the full model.

We remove $\sigma_{1\xi}$ from the full model by assuming $\sigma_{1\xi}$ is a small known value that is the same for all years and divisions. The estimates for RM 1 with their corresponding CIs are shown in Appendix A.2. Note that, the estimates of σ_∞ for all years and divisions reach the lower bound in this model. The CIs for σ_∞ estimates are not reported. This model (RM 1) and the full model gave the same negative likelihood (nll) value (see Table 5.7). Removing 15 parameters left the nll the same. This indicates the model with fewer parameters (RM 1) is more efficient than the full model. RM 1 is now our primary model.

5.2.2 RM 2: remove between-individual variance parameter (σ_∞) from RM 1.

We remove σ_∞ from RM 1 by assuming σ_∞ is a small known value, that is the same for all years and divisions. The estimates with their corresponding CIs for this model are shown in Appendix A.3. This model (RM 2), the RM 1 and the full model all gave the same negative likelihood (nll) value (see Table 5.7). This shows that the assumption that σ_∞ is a small and known value, that is the same for all years and divisions is satisfied. Thus σ_∞ can be removed.

5.2.3 RM 3: combine between-individual variance parameter (σ_k) across year and division.

In RM 3 we combine σ_k from RM 2 across years and divisions. Thus we assume that σ_k is the same for all years and divisions. The estimates with their corresponding CIs are attached in Appendix A.4. The LLR test is performed in order to compare the RM 3 and RM 2. The null hypothesis for this test is, σ_k is the same across years and divisions (i.e., between-individual variation in the growth of fish is the same across years and divisions). The small p value of $0 < 0.001$ (see Table 5.7) strongly indicates that σ_k is different for each year and division. The RM 2 is preferred over the RM 3.

5.2.4 RM 4: combine between-individual variance parameter (σ_k) for each division across the years.

In RM 4 we combine σ_k from RM 2 for each divisions (2J, 3K and 3L) across the years. Thus we assume that σ_k is the same for the same divisions across all years. The estimates with their corresponding CIs are attached in Appendices A.5 and A.1. The LLR test is performed in order to compare the RM 4 and RM 2. The p value of $0.098 > 0.001$ (see Table 5.7) indicates that σ_k is the same for the same division the across years. The RM 4 is preferred over the RM by the LLR test and the AIC. RM 4 is now our primary.

5.2.5 RM 5: combine ageing error intercept variance parameter ($\sigma_{o\xi}$) across year and division.

For RM 5, we assume that $\sigma_{o\xi}$ is the same for all years and divisions. The estimates with their corresponding CIs are attached in Appendix A.6. The LLR test is performed in order to compare RM 5 and RM 4. The null hypothesis for this test is, $\sigma_{o\xi}$ is the same across years and divisions (i.e., ageing error intercept variance parameter is the same across year and division). The small p value of $0 < 0.001$ (see Table 5.7) indicates that $\sigma_{o\xi}$ is different for each year and division. This is surprising as we were expecting $\sigma_{o\xi}$ to be the same across year and division, since the ageing error depends on the individual reading the age. RM 4 is selected over RM 5.

5.2.6 RM 6: combine VonB growth parameter (k) across year and division.

In RM 6 we assume that k is the same for fish caught in the same division across the years. A p value of $0 < 0.001$ (see Table 5.7) strongly suggests that k is not equivalent for the same division across the years. Thus RM 4 is selected over RM 5. The estimates for RM 6 with corresponding CI's are shown in Appendices A.7.

5.2.7 RM 7: combine VonB growth parameter (k) for 3K and 3L divisions for 2009, 2010, 2012, and 2013.

In RM 7 we assume that k is the same for fish caught in divisions 3K and 3L for the years 2009, 2010, 2012 and 2013. This will reduce the parameters in our primary model (RM 4) by 4. A p value of $0.021 > 0.001$ (see Table 5.7) suggests strong evidence that k is equivalent for 3K and 3L divisions for the years 2009, 2010, 2012, and 2013. Thus the LLR test and the BIC preferred RM 7 over RM 4. The estimates for RM 7 with corresponding CI's are shown in Appendices A.8 and A.2. RM 7 is now our primary.

5.2.8 RM 8: combine VonB growth parameter (l_∞) across year and division.

Now, we wish to test if l_∞ is the same for the same division across the years. The estimates with their corresponding CIs are attached in Appendix A.9. The LLR test is performed in order to compare RM 8 and RM 7. The null hypothesis for this test is, l_∞ is the same across years and divisions (i.e., all fish achieve the same maximum sizes across years and divisions). The small p value of $0 < 0.001$ (see Table 5.7) indicates that l_∞ is different for each year and division. Thus RM 7 is selected over RM 8.

5.2.9 RM 9: combine VonB growth parameter (t_o) across year and division.

For RM 9 we assume that t_o is the same across years and divisions. The estimates for RM 9 with corresponding CI's are shown in Appendix A.10. A smalll p value of $0 < 0.001$ (see Table 5.7) indicates that t_o is different across divisions and years. Thus RM 7 is selected over RM 9.

5.2.10 RM 10: combine the true age distribution parameter (α) across year and division.

Now, we wish to test if α is the same across years and divisions. The estimates with their corresponding CIs for RM 10 are attached in Appendix A.11. The LLR test is performed in order to compare RM 10 and RM 7. The null hypothesis for this test is, α is the same across years and divisions. The small p value of $0 < 0.001$ (see Table 5.7) indicates that α is different for each year and division as we were expecting. Thus RM 7 is selected over RM 10.

5.2.11 RM 11: combine the true age distribution parameter (β) across year and division.

For RM 11 we test if β is the same across years and divisions. The estimates with their corresponding CIs for RM 11 are attached in Appendix A.12. The LLR test is performed in order to compare RM 11 and RM 7. The small p value of $0 < 0.001$ (see Table 5.7) indicates that β is different for each year and division as we were expecting. Thus RM 7 is our final model.

5.3 Tables

Area	Division	Strata
1	2H	All Strata
2	2J	201-206, 213-226, 234, 236-239
3	2J	207-212, 227-233, 235, 240
4	3K	617-618, 620, 622-625, 630-631, 633,645-649
5	3K	619, 621, 626-629, 634-644, 650-654
6	3L	328, 341-343, 348-350, 363-365, 370-372,384-385, 390
7	3L	344-347, 366, 368-369, 386-389, 391-392,729-751
8	3NO	All Strata

Table 5.1: Sampling structure.

Length cl.	2009 2J	2009 3K	2009 3L	2010 2J	2010 3K	2010 3L	2011 2J	2011 3K	2011 3L	2012 2J	2012 3K	2012 3L	2013 2J	2013 3K	2013 3L
4	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
7	1	0	1	0	0	0	0	0	7	0	0	2	1	2	8
10	0	0	0	0	0	0	0	0	0	0	4	2	3	9	12
13	0	7	3	0	0	0	0	0	0	0	9	8	5	6	7
16	1	5	8	2	15	5	0	1	2	0	13	16	24	38	30
19	13	21	26	6	58	25	3	5	2	4	23	23	39	55	42
22	28	31	33	21	74	22	6	8	0	13	39	23	42	61	45
25	18	32	45	15	61	23	7	10	6	8	33	18	39	55	32
28	23	40	43	14	46	29	18	31	18	9	22	14	37	53	37
31	36	51	41	14	49	25	24	51	22	15	25	19	38	49	43
34	30	39	28	21	46	19	28	44	32	24	37	21	39	53	31
37	19	34	31	16	48	24	30	38	36	40	44	31	28	48	38
40	25	42	33	27	55	24	19	36	32	59	45	42	25	42	42
43	26	40	36	18	49	20	16	36	37	50	48	40	28	51	38
46	17	38	29	16	42	22	13	33	29	56	54	37	35	47	45
49	6	32	27	17	42	23	12	34	30	40	48	36	38	59	39
52	8	37	24	12	40	13	5	31	32	27	35	36	35	58	41
55	0	26	14	2	38	12	3	28	16	22	36	25	28	58	48
58	5	23	14	1	25	15	0	24	16	11	33	27	20	48	37
61	1	15	11	0	22	8	2	20	15	10	36	28	10	40	40
64	1	19	12	2	20	9	1	20	14	4	34	15	11	43	35
67	0	16	10	1	17	5	0	21	13	4	28	21	3	39	26
70	0	12	9	1	19	8	0	14	9	0	27	12	5	37	21
73	0	15	4	0	18	3	0	22	9	0	21	6	3	34	21
76	0	7	6	0	10	5	0	19	4	0	16	12	1	30	23
79	0	6	1	0	7	4	0	9	7	1	16	5	1	27	12
82	0	3	1	0	4	4	0	9	4	0	12	10	1	15	8
85	0	5	0	0	4	1	0	8	3	0	11	7	1	14	9
88	0	1	0	0	3	1	0	4	1	0	7	7	0	13	8
91	0	1	0	0	0	1	0	3	1	1	7	0	0	8	4
94	0	0	0	0	0	0	0	1	0	1	4	2	0	2	3
97	0	0	0	0	1	1	0	0	0	0	2	2	0	0	1
100	0	0	0	0	1	0	0	0	0	0	1	1	0	1	1
103	0	0	1	0	1	0	0	0	0	0	1	1	0	0	1
106	0	0	1	0	0	0	0	0	0	0	2	0	1	1	0
109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
121	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0

Table 5.2: Number of cod sampled from each length class (length cl.) for age determination by year and NAFO division

	Population number by year					Number sampled for age by year					Sampled totals
Length class	2009	2010	2011	2012	2013	2009	2010	2011	2012	2013	by length class
1	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	41791	41791	0	0	0	1	1	2
7	88421	0	0	291295	1399081	2	0	7	2	11	22
10	0	0	0	0	0	0	0	0	6	24	30
13	574792	0	0	4264772	1431847	10	0	0	17	18	45
16	1426736	1212956	240789	3983675	8833439	14	22	3	29	93	161
19	4269153	7854663	1005702	5337264	33017901	60	89	10	51	136	346
22	5361483	9669875	1589719	8454724	40779286	92	117	14	76	148	447
25	6619069	5146605	1779987	7454747	22376579	95	99	23	59	126	402
28	9172719	4328690	6662793	3435299	14501333	106	89	67	45	127	434
31	12201674	5829681	11220796	5640158	23261626	128	88	98	59	130	503
34	9898640	8026767	14191966	6757423	19982062	97	86	107	83	123	496
37	15317175	9859780	11587318	7910950	11809670	84	88	106	115	114	507
40	25000663	15151088	6559203	15060449	8316907	100	106	88	147	109	550
43	24366334	12621797	8021887	20803297	9730327	102	88	89	144	117	540
46	18775353	9561244	6629892	18127504	12023396	84	81	77	149	127	518
49	12316735	7255113	7132660	13543834	18716595	66	82	78	125	136	487
52	7100395	6630608	7081699	9547777	21026803	69	65	70	98	134	436
55	5281100	5619067	3842264	8404600	16854037	40	52	47	83	135	357
58	4190400	4071473	3483516	6636392	12058497	42	41	41	72	105	301
61	3231849	2918032	3542898	5546391	10381293	27	30	37	75	90	259
64	2898582	2752178	2950837	3553496	7791726	32	31	35	53	89	240
67	2217524	2195112	2305465	3280806	5822370	26	23	34	53	68	204
70	1679058	1953126	1686282	2033002	4180141	21	28	23	40	63	175
73	1287756	1505051	1875827	1633234	3752018	19	21	31	27	58	156
76	772674	1203401	1192331	1317315	2998290	13	15	23	28	54	133
79	296934	562307	1122533	1251831	1632612	7	11	16	22	40	96
82	239239	378435	587356	851251	1130872	4	8	13	22	24	71
85	218533	187911	648852	712146	1240113	5	5	11	18	24	63
88	35405	148473	149459	549945	998185	1	4	5	14	21	45
91	35405	36059	157949	301247	583138	1	1	5	8	12	27
94	0	0	67064	249027	248620	0	0	1	7	5	13
97	0	66649	0	134496	34091	0	2	0	4	1	7
100	0	50353	0	78616	124504	0	1	0	2	2	5
103	42206	35967	0	73205	39663	1	1	0	2	1	5
106	48343	0	0	72669	86331	1	0	0	2	2	5
109	0	0	0	0	40810	0	0	0	0	1	1
112	0	0	0	0	0	0	0	0	0	0	0
115	0	0	0	0	0	0	0	0	0	0	0
118	0	0	0	0	0	0	0	0	0	0	0
121	0	100707	0	0	0	0	2	0	0	0	2

Table 5.3: Total number of fish caught and the number that was sampled for age determination by year and length class.

Nafo div	Stratum	2009	2010	2011	2012	2013
2J	201	0	1	0	4	12
	202	0	1	0	3	2
	203	3	4	0	11	5
	204	1	0	0	0	2
	205	5	5	1	7	13
	206	8	13	7	25	49
	207	27	18	13	38	21
	208	30	0	6	5	15
	209	2	13	1	16	41
	210	4	14	18	49	53
	211	8	4	5	4	47
	212	2	1	0	0	1
	213	28	22	30	19	21
	214	27	3	18	13	22
	215	12	7	4	69	20
	216	8	4	12	3	7
	217	0	1	0	0	0
	222	27	32	23	16	32
	223	0	3	0	0	0
	227	0	0	0	0	4
	228	23	44	32	78	109
	229	25	8	17	5	25
	234	7	2	0	1	2
	235	7	2	0	1	7
	237	1	4	0	26	25
	238	1	0	0	5	6
	240	2	0	0	1	0

Table 5.4: Number of cod sampled from each strata in NAFO division for each year.

Nafo div	Stratum	2009	2010	2011	2012	2013
3K	328	27	0	19	43	50
	341	17	0	18	8	4
	342	10	0	10	4	5
	343	9	0	4	4	0
	344	10	60	17	41	98
	345	61	79	50	109	143
	346	143	66	47	36	44
	347	14	29	20	76	13
	348	17	9	7	7	38
	349	11	0	5	7	17
	350	22	0	19	12	45
	363	3	0	13	0	17
	364	5	0	8	2	12
	365	0	7	0	1	0
	366	59	77	79	46	124
	368	16	24	12	6	1
	369	2	0	1	26	5
	370	2	0	2	1	1
	371	1	0	0	7	4
	372	1	0	2	27	30
	384	3	0	3	3	5
	385	0	0	0	2	3
	386	0	0	3	2	20
	387	0	0	24	9	20
	388	33	0	13	12	23
	389	20	0	9	36	62
	390	0	0	3	1	13
	391	4	0	2	3	12
	392	2	0	7	19	21

Table 5.5: Number of cod sampled from each strata in NAFO division for each year.

Nafo div	Stratum	2009	2010	2011	2012	2013
3L	617	11	15	11	23	39
	618	15	22	12	15	45
	619	0	11	1	16	59
	620	13	51	27	63	78
	621	20	22	0	90	81
	622	4	14	3	3	5
	623	4	20	26	20	25
	624	0	8	14	44	140
	625	5	11	23	10	75
	626	23	17	6	27	79
	627	23	9	4	7	38
	628	38	63	66	49	85
	629	4	14	5	33	17
	630	1	0	2	1	1
	631	1	0	0	8	0
	633	22	13	24	21	25
	634	1	16	70	60	17
	635	5	9	40	28	27
	636	51	81	47	61	36
	637	33	67	61	65	68
	638	183	200	68	63	68
	639	136	149	50	65	88
	640	5	5	0	1	0

Table 5.6: Number of cod sampled from each strata in NAFO division for each year.

5.4 Testing the models.

Model	Table Ref.	2nll	No. par	Chisq stat	P-value	AIC	Δ AIC	BIC	Δ BIC
Full	A.1	84236.009	135.000	-	-	84506.010	65.363	85450.206	383.567
RM 1	A.2	84236.009	120.000	-	-	84476.009	35.362	85315.295	248.655
RM 2	A.3	84236.009	105.000	-	-	84446.009	5.362	85180.384	113.744
RM 3	A.4	84272.972	91.000	36.963	0.001	84454.972	14.325	85091.430	24.791
RM 4	A.5	84254.647	93.000	18.638	0.098	84440.647	0.000	85091.093	24.454
RM 5	A.6	84517.659	79.000	263.012	0.000	84675.659	235.012	85228.189	161.549
RM 6	A.7	84266.169	79.000	207.018	0.000	84619.665	179.018	85172.194	105.555
RM 7	A.8	84461.665	89.000	11.523	0.021	84444.169	3.523	85066.640	0.000
RM 8	A.9	84437.913	75.000	171.743	0.000	84587.913	147.266	85112.466	45.827
RM 9	A.10	84588.113	75.000	321.944	0.000	84738.113	297.466	85262.667	196.027
RM 10	A.11	84686.664	75.000	420.495	0.000	84836.664	396.018	85361.218	294.578
RM 11	A.12	84818.193	75.000	552.024	0.000	84968.193	527.546	85492.747	426.107

Table 5.7: Negative log likelihood (nll), likelihood ratio test statistics (chisq stat), number of parameters (No. Par), P values, AIC and BIC for all 12 models.

5.5 Figures

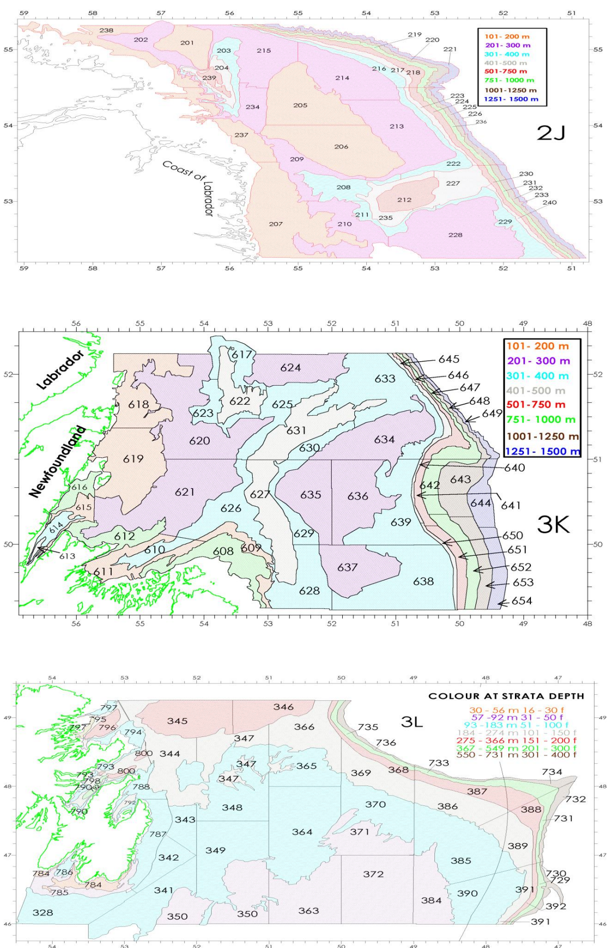

Figure 5.1: NAFO division 2J3KL, with numbers indicating strata.

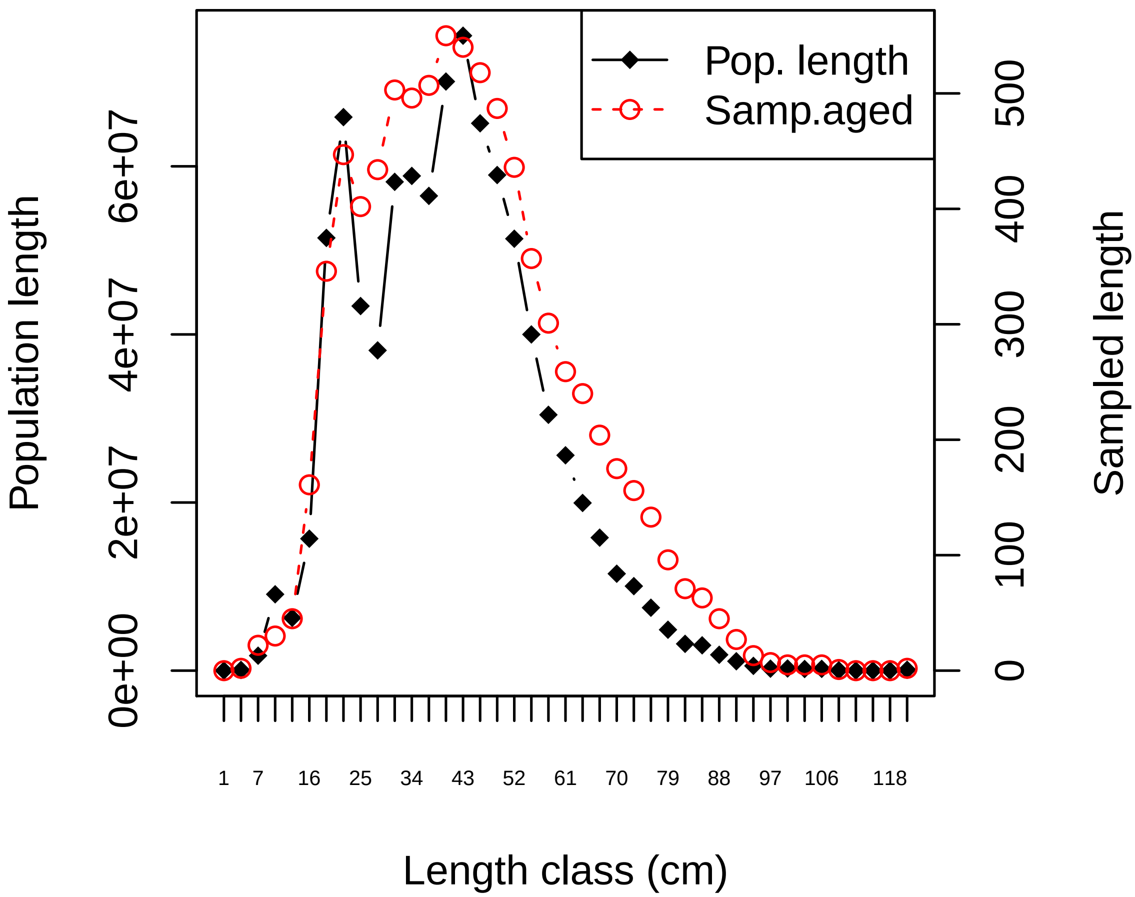

Figure 5.2: Number of tow catches by 3 cm length class (solid line, diamond) and the number sampled for age (dashed line, circles) from 2009 to 2013.

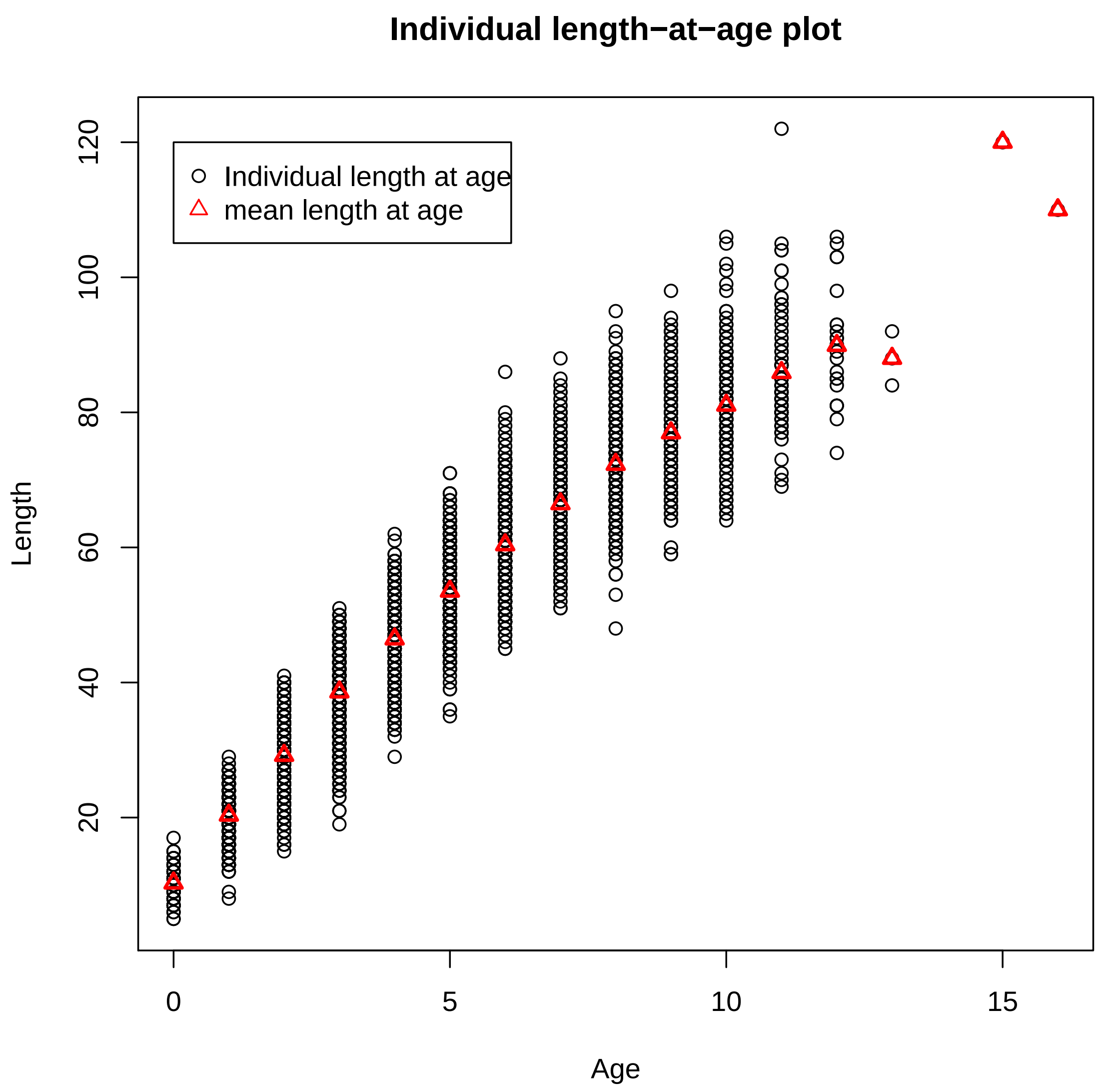

Figure 5.3: Individual length-at-age (Ind. length-at-age, circle) data and the mean length-at-age (mean length-at-age, diamonds) for Northwest Atlantic Fisheries Organization (NAFO) divisions 2J, 3K, and 3L (circles) from 2009-2013.

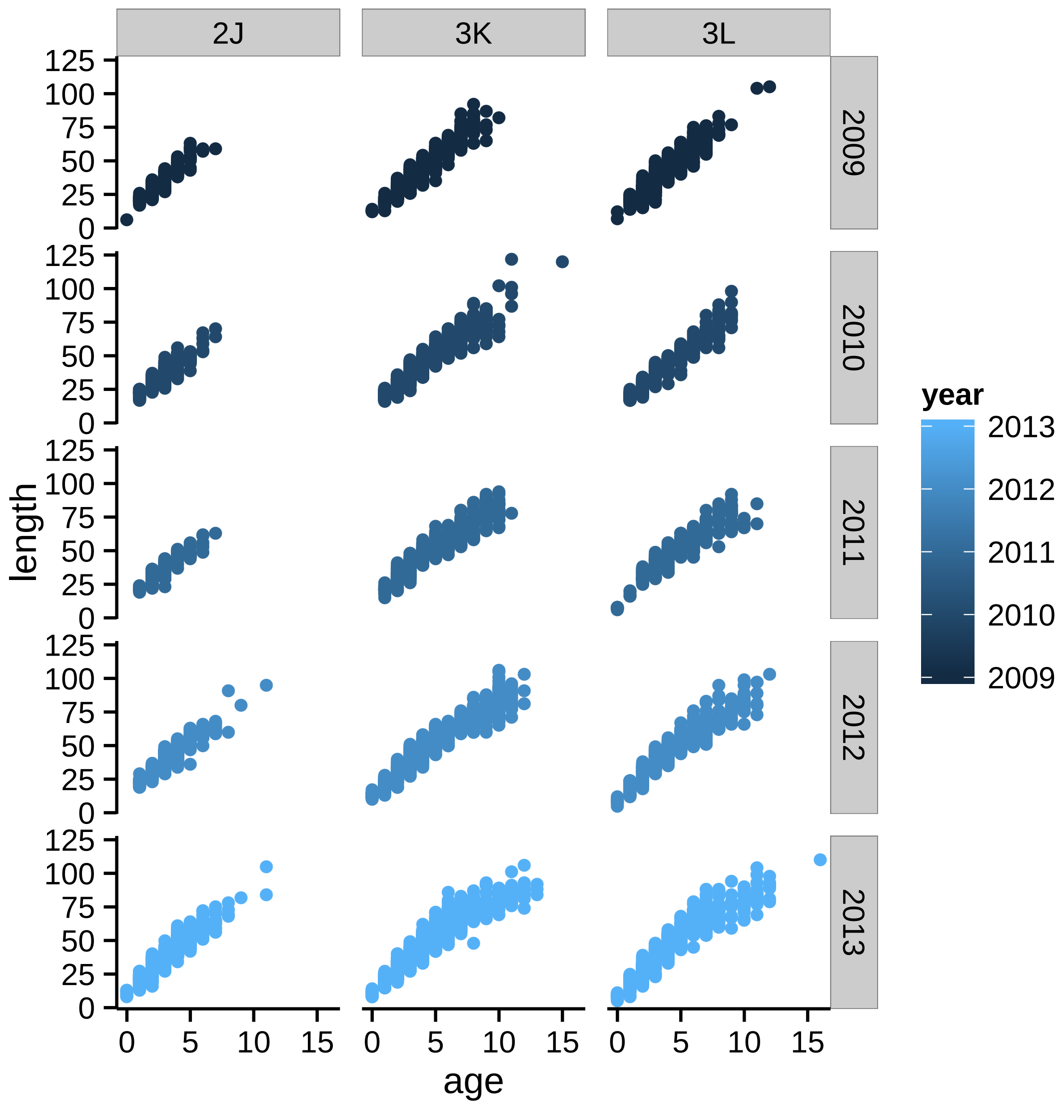

Figure 5.4: Individual length and age data by NAFO division and year for Northwest Atlantic Fisheries Organization (NAFO) divisions 2J, 3K, and 3L (circles).

Chapter 6

Summary and discussion

The focus of this practicum was to model the growth of northern cod. In developing this model, we first extended the commonly used Von Bertalanffy (VonB) growth model to include measurement error (ME) in length and account for between-individual (BT) variation. We investigate the commonly used maximum likelihood estimators (MLE) that assume only ME in length (MEL) is present and does not include BT variation in L_∞ and K; and the MLE that assumes both ME in length and BT variation (MELBT) in VonB parameters. We confirmed in Chapter 2 that the MLE of MEL method was unreliable, since the MLE of MEL method was misspecified in 2 data sets; data sets with only between-individual (BT) variation, and data sets with BT plus ME in length. The MLE of MELBT seems to be reliable in estimating the VonB growth parameters for all data sets. However, we realized that the parameters that described the BT and ME in length variations, σ_∞

and σ_ε were completely confounded.

We fixed this confounding in Chapter 3 using a spatial model and included ageing error to have a more realistic error structure. We expected the ageing error standard deviation (σ_ξ) to be confounded with σ_∞ and σ_k. As a result we assumed σ_∞ and σ_k are negligible for fish caught in the same year and stratum. That is, variation in growth rate of individuals within a stratum was due to measurement error in length and age (σ_ε and σ_ξ) only, and the value σ_ε is known. This made it possible to estimate σ_ξ.

We studied two ME models described by Carroll et al. (2006): the functional errors-in-variables (FEV) model and the structural errors-in-variables (SEV) model. We used a simulation study to confirm that the SEV model was reliable even when the true age distribution is different from a Gamma distribution. We concluded from the study that the SEV method was more reliable than the REML FEV method we investigated. The SEV was the estimating method we pursued in Chapters 4 and 5.

In Chapter 4, we extended our SEV growth model to account for variations that may be caused by the:

- Size selectivity of the fishing gear used to catch the fish.
- Size-stratified age sampling.

The SEV growth model developed in this study has several advantages. First, it contributes to modelling fish growth using individual length-at-age data sets. It also enabled us to estimate the true age distribution of the fish. Finally, the model accounts for BT variation in growth and the sampling designs used to collect the length-at-age measurements. All of these are not accounted for in the commonly used VonB model.

There are many factors that affect individual growth that are not included into our extended VonB growth model (see Filipe et al., 2010; Taylor et al., 2005; Shelton & Mangel, 2012). We extended the VonB growth model to account for BT variation based on the assumption that l_∞ and k are independent. It is important to note that various studies have been done to show that l_∞ and k are correlated (see Shelton & Mangel, 2012). Also be able to estimate σ_ξ we had to assume that the only variation in growth rate of individuals within a stratum, was due to measurement error in length and age (σ_ε and σ_ξ), and the value for σ_ε is known.

We applied this extended growth model to model the growth of the 2J3KL Northern cod by year and NAFO division. We used the LLR test for model selection. We did not used the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) for model selection, since we were unsure how reliable they are for mixed-effect model selection. Some papers have proposed improvements for mixed-effect models. This is beyond the scope of this report. Our results shows the existence of individual variation in growth and ME in age. According to the LLR test, the best model indicated: 1) different growth patterns across divisions and years. 2) Between individual variation in growth is the same for the same division across years. 3) The ME in age and true age distribution are different for each year and division.

In conclusion, we hope this practicum will serve as a resource for further modelling of 2J3KL Northern cod growth to predict unbiased fishery trends; this will help maintain a healthy and sustainable fish population and fishing industry.

Appendices

Appendix A

Estimates of parameters of 2J3KL Northern cod, 2009-2013 by year and division.

A.1 Full model

		2J			3K			3L		
Year	Parameter	Est	L95	U95	Est	L95	U95	Est	L95	U95
2009	l_∞	103.442	79.417	134.736	259.506	148.917	452.221	350.000	132.273	926.112
	k	0.138	0.095	0.201	0.040	0.021	0.077	0.031	0.011	0.093
	t_o	-0.338	-0.515	-0.160	-0.648	-0.869	-0.427	-0.345	-0.590	-0.100
	α	6.071	4.931	7.475	4.240	3.740	4.807	5.431	4.683	6.298
	β	0.435	0.349	0.542	0.888	0.775	1.017	0.627	0.535	0.734
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.023	0.007	0.081	0.018	0.000	4.558	0.047	0.030	0.073
	$\sigma_{o\xi}$	0.196	0.178	0.215	0.200	0.187	0.215	0.226	0.211	0.242
	$\sigma_{1\xi}$	0.000	-	-	0.000	-	-	0.000	-	-
2010	l_∞	188.084	59.106	598.516	245.243	177.736	338.390	340.227	145.712	794.407
	k	0.058	0.014	0.251	0.037	0.025	0.054	0.028	0.011	0.072
	t_o	-0.852	-1.494	-0.209	-1.204	-1.344	-1.064	-1.013	-1.222	-0.804
	α	5.441	4.342	6.819	2.779	2.517	3.068	3.232	2.777	3.761
	β	0.536	0.421	0.683	1.322	1.184	1.477	1.140	0.965	1.347
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.018	0.008	0.040	0.018	0.000	13.593	0.018	0.004	0.092
	$\sigma_{o\xi}$	0.221	0.198	0.246	0.189	0.179	0.199	0.170	0.157	0.184
	$\sigma_{1\xi}$	0.000	-	-	0.000	-	-	0.000	-	-
2011	l_∞	121.632	69.549	212.717	158.121	130.298	191.884	111.226	101.807	121.516
	k	0.098	0.044	0.219	0.070	0.053	0.093	0.126	0.110	0.145
	t_o	-0.801	-1.287	-0.315	-0.984	-1.226	-0.741	-0.284	-0.437	-0.132
	α	8.886	6.994	11.291	3.983	3.519	4.508	4.114	3.452	4.902
	β	0.328	0.255	0.420	1.127	0.986	1.289	1.030	0.853	1.242
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.046	0.009	0.238	0.068	0.030	0.157	0.018	0.006	0.058
	$\sigma_{o\xi}$	0.164	0.147	0.183	0.171	0.161	0.183	0.221	0.205	0.238
	$\sigma_{1\xi}$	0.000	-	-	0.000	-	-	0.000	-	-
2012	l_∞	227.028	121.813	423.121	126.297	117.110	136.204	121.305	112.579	130.708
	k	0.050	0.023	0.106	0.107	0.095	0.120	0.114	0.102	0.128
	t_o	-0.846	-1.129	-0.563	-0.530	-0.630	-0.430	-0.342	-0.434	-0.250
	α	8.799	7.474	10.358	2.637	2.370	2.933	2.819	2.487	3.195
	β	0.397	0.335	0.470	1.648	1.461	1.858	1.485	1.292	1.708
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.018	0.008	0.042	0.049	0.026	0.091	0.029	0.009	0.091
	$\sigma_{o\xi}$	0.167	0.155	0.179	0.247	0.234	0.260	0.212	0.200	0.226
	$\sigma_{1\xi}$	0.000	-	-	0.000	-	-	0.000	-	-
2013	l_∞	136.585	111.721	166.982	122.230	114.379	130.619	118.236	112.271	124.517
	k	0.099	0.074	0.133	0.112	0.100	0.124	0.122	0.112	0.134
	t_o	-0.403	-0.606	-0.200	-0.435	-0.512	-0.359	-0.173	-0.260	-0.085
	α	3.274	2.874	3.729	2.371	2.174	2.585	2.548	2.300	2.823
	β	0.910	0.788	1.051	1.757	1.592	1.940	1.598	1.423	1.795
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.038	0.019	0.076	0.018	0.003	0.096	0.028	0.010	0.075
	$\sigma_{o\xi}$	0.239	0.224	0.255	0.231	0.221	0.242	0.254	0.241	0.268
	$\sigma_{1\xi}$	0.000	-	-	0.000	-	-	0.000	-	-

Table A.1: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). All parameters different for each year and division (full model).

A.2 Reduced model (RM) 1

Year	Parameter	2J Est	2J L95	2J U95	3K Est	3K L95	3K U95	3L Est	3L L95	3L U95
2009	l_∞	103.442	79.417	134.736	259.489	148.923	452.144	350.000	132.274	926.110
	k	0.138	0.095	0.201	0.040	0.021	0.077	0.031	0.011	0.093
	t_o	-0.338	-0.515	-0.160	-0.648	-0.868	-0.427	-0.345	-0.590	-0.100
	α	6.071	4.931	7.475	4.240	3.740	4.807	5.431	4.683	6.298
	β	0.435	0.349	0.542	0.888	0.775	1.017	0.627	0.535	0.734
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.023	0.007	0.081	0.018	0.001	4.534	0.047	0.030	0.073
	$\sigma_{o\xi}$	0.196	0.178	0.215	0.200	0.187	0.215	0.226	0.211	0.242
2010	l_∞	188.007	59.170	597.369	245.243	177.735	338.392	340.162	145.720	794.056
	k	0.058	0.014	0.251	0.037	0.025	0.054	0.028	0.011	0.072
	t_o	-0.852	-1.493	-0.210	-1.204	-1.344	-1.064	-1.013	-1.222	-0.804
	α	5.441	4.342	6.819	2.779	2.517	3.068	3.232	2.777	3.761
	β	0.536	0.421	0.683	1.322	1.184	1.477	1.140	0.965	1.347
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.018	0.008	0.040	0.018	0.001	13.544	0.018	0.004	0.092
	$\sigma_{o\xi}$	0.221	0.198	0.246	0.189	0.179	0.199	0.170	0.157	0.184
2011	l_∞	121.625	69.550	212.694	158.119	130.298	191.881	111.226	101.807	121.515
	k	0.098	0.044	0.219	0.070	0.053	0.093	0.126	0.110	0.145
	t_o	-0.801	-1.287	-0.315	-0.984	-1.226	-0.741	-0.284	-0.437	-0.132
	α	8.886	6.993	11.291	3.983	3.519	4.508	4.114	3.452	4.902
	β	0.328	0.255	0.420	1.127	0.986	1.289	1.030	0.853	1.243
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.046	0.009	0.238	0.068	0.030	0.157	0.018	0.006	0.058
	$\sigma_{o\xi}$	0.164	0.147	0.183	0.171	0.161	0.183	0.221	0.205	0.238
2012	l_∞	227.010	121.820	423.029	126.297	117.110	136.204	121.305	112.578	130.708
	k	0.050	0.023	0.106	0.107	0.095	0.120	0.114	0.102	0.128
	t_o	-0.846	-1.129	-0.563	-0.530	-0.630	-0.430	-0.342	-0.434	-0.250
	α	8.799	7.474	10.358	2.637	2.370	2.933	2.819	2.487	3.195
	β	0.397	0.335	0.470	1.648	1.461	1.858	1.485	1.292	1.708
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.018	0.008	0.042	0.049	0.026	0.091	0.029	0.009	0.091
	$\sigma_{o\xi}$	0.167	0.155	0.179	0.247	0.234	0.260	0.212	0.200	0.226
2013	l_∞	136.582	111.724	166.970	122.230	114.379	130.619	118.235	112.271	124.516
	k	0.099	0.074	0.133	0.112	0.100	0.124	0.122	0.112	0.134
	t_o	-0.403	-0.606	-0.200	-0.435	-0.512	-0.359	-0.173	-0.260	-0.085
	α	3.274	2.874	3.729	2.371	2.174	2.585	2.548	2.300	2.823
	β	0.910	0.788	1.051	1.757	1.592	1.940	1.598	1.423	1.795
	σ_∞	0.000	-	-	0.000	-	-	0.000	-	-
	σ_k	0.038	0.019	0.076	0.018	0.003	0.096	0.028	0.010	0.075
	$\sigma_{o\xi}$	0.239	0.224	0.255	0.231	0.221	0.242	0.254	0.241	0.268

Table A.2: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The ageing error slope variance parameters ($\sigma_{1\xi}$) removed from the full model (Table A.1).

A.3 RM 2

		2J			3K			3L		
Year	Parameter	Est	L95	U95	Est	L95	U95	Est	L95	U95
2009	l_∞	103.441	79.416	134.733	259.492	148.920	452.162	350.000	132.270	926.138
	k	0.138	0.095	0.201	0.040	0.021	0.077	0.031	0.011	0.093
	t_o 3	-0.338	-0.515	-0.160	-0.648	-0.868	-0.427	-0.345	-0.590	-0.100
	α	6.071	4.931	7.475	4.240	3.740	4.807	5.431	4.683	6.298
	β	0.435	0.349	0.542	0.888	0.775	1.017	0.627	0.535	0.734
	σ_k	0.023	0.007	0.081	0.018	0.001	4.535	0.047	0.030	0.073
	$\sigma_{o\xi}$	0.196	0.178	0.215	0.200	0.187	0.187	0.226	0.211	0.242
2010	l_∞	188.008	59.159	597.489	245.246	177.736	338.398	340.170	145.713	794.133
	k	0.058	0.014	0.251	0.037	0.025	0.054	0.028	0.011	0.072
	t_o	-0.852	-1.493	-0.210	-1.204	-1.344	-1.064	-1.013	-1.222	-0.804
	α	5.441	4.342	6.819	2.779	2.517	3.068	3.232	2.777	3.761
	β	0.536	0.421	0.683	1.322	1.184	1.477	1.140	0.965	1.347
	σ_k	0.018	0.008	0.040	0.018	0.001	13.550	0.018	0.004	0.092
	$\sigma_{o\xi}$	0.221	0.198	0.246	0.189	0.179	0.199	0.170	0.157	0.184
2011	l_∞	121.626	69.553	212.686	158.121	130.298	191.885	111.226	101.807	121.516
	k	0.098	0.044	0.219	0.070	0.053	0.093	0.126	0.110	0.145
	t_o	-0.801	-1.287	-0.315	-0.984	-1.226	-0.741	-0.284	-0.437	-0.132
	α	8.886	6.994	11.291	3.982	3.519	4.508	4.114	3.452	4.902
	β	0.328	0.255	0.420	1.127	0.986	1.289	1.030	0.853	1.242
	σ_k	0.046	0.009	0.238	0.068	0.030	0.157	0.018	0.006	0.058
	$\sigma_{o\xi}$	0.164	0.147	0.183	0.171	0.161	0.183	0.221	0.205	0.238
2012	l_∞	227.013	121.820	423.040	126.297	117.110	136.205	121.305	112.579	130.708
	k	0.050	0.023	0.106	0.107	0.095	0.120	0.114	0.102	0.128
	t_o	-0.846	-1.129	-0.563	-0.530	-0.630	-0.430	-0.342	-0.434	-0.250
	α	8.799	7.474	10.358	2.637	2.370	2.933	2.819	2.487	3.195
	β	0.397	0.335	0.470	1.648	1.461	1.858	1.485	1.292	1.708
	σ_k	0.018	0.008	0.042	0.049	0.026	0.091	0.029	0.009	0.091
	$\sigma_{o\xi}$	0.167	0.155	0.179	0.247	0.234	0.260	0.212	0.200	0.226
2013	l_∞	136.584	111.723	166.977	122.230	114.379	130.619	118.235	112.271	124.516
	k	0.099	0.074	0.133	0.112	0.100	0.124	0.122	0.112	0.134
	t_o	-0.403	-0.606	-0.200	-0.435	-0.512	-0.359	-0.173	-0.260	-0.085
	α	3.274	2.874	3.729	2.371	2.174	2.585	2.548	2.300	2.823
	β	0.910	0.788	1.051	1.757	1.592	1.940	1.598	1.423	1.795
	σ_k	0.038	0.019	0.076	0.018	0.003	0.096	0.028	0.010	0.075
	$\sigma_{o\xi}$	0.239	0.224	0.255	0.231	0.221	0.242	0.254	0.241	0.268

Table A.3: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The between-individual variance parameter (σ_∞) is removed from RM 1 (Table A.2).

A.4 RM 3

Year	Parameter	2J Est	2J L95	2J U95	3K Est	3K L95	3K U95	3L Est	3L L95	3L U95
2009	l_∞	102.646	79.263	132.926	265.359	152.408	462.019	350.000	127.331	962.059
	k	0.139	0.096	0.202	0.040	0.021	0.076	0.032	0.010	0.099
	t_o	-0.343	-0.520	-0.166	-0.648	-0.853	-0.443	-0.286	-0.522	-0.050
	α	6.063	4.925	7.464	4.232	3.736	4.794	5.459	4.709	6.330
	β	0.436	0.350	0.543	0.890	0.777	1.018	0.623	0.532	0.729
	$\sigma_{o\xi}$	0.195	0.178	0.214	0.199	0.187	0.212	0.225	0.210	0.241
2010	l_∞	201.133	60.666	666.844	247.110	178.350	342.381	338.982	146.217	785.875
	k	0.053	0.013	0.225	0.037	0.025	0.055	0.028	0.011	0.071
	t_o	-0.921	-1.362	-0.479	-1.159	-1.296	-1.021	-1.013	-1.222	-0.804
	α	5.415	4.328	6.775	2.783	2.521	3.072	3.231	2.777	3.760
	β	0.539	0.424	0.685	1.320	1.182	1.474	1.140	0.965	1.347
	$\sigma_{o\xi}$	0.220	0.198	0.244	0.188	0.179	0.199	0.170	0.157	0.184
2011	l_∞	126.017	67.848	234.058	148.246	126.585	173.613	111.322	101.984	121.516
	k	0.096	0.040	0.230	0.077	0.061	0.097	0.126	0.110	0.145
	t_o	-0.744	-1.250	-0.239	-0.864	-1.066	-0.663	-0.283	-0.432	-0.134
	α	9.041	7.113	11.490	3.984	3.520	4.509	4.110	3.450	4.897
	β	0.322	0.251	0.413	1.126	0.985	1.288	1.030	0.854	1.243
	$\sigma_{o\xi}$	0.165	0.148	0.184	0.170	0.160	0.181	0.221	0.205	0.237
2012	l_∞	220.937	121.723	401.016	125.722	116.732	135.405	121.408	112.657	130.838
	k	0.051	0.025	0.106	0.107	0.095	0.120	0.113	0.101	0.127
	t_o	-0.825	-1.106	-0.544	-0.550	-0.649	-0.452	-0.344	-0.430	-0.259
	α	8.802	7.478	10.361	2.628	2.362	2.923	2.821	2.489	3.197
	β	0.397	0.335	0.470	1.653	1.466	1.865	1.485	1.291	1.707
	$\sigma_{o\xi}$	0.167	0.155	0.179	0.246	0.234	0.260	0.212	0.199	0.226
2013	l_∞	137.667	114.238	165.900	122.365	114.505	130.763	118.165	112.270	124.370
	k	0.098	0.076	0.126	0.112	0.101	0.125	0.122	0.112	0.133
	t_o	-0.421	-0.532	-0.309	-0.442	-0.518	-0.366	-0.178	-0.253	-0.103
	α	3.272	2.873	3.727	2.371	2.174	2.586	2.546	2.298	2.820
	β	0.910	0.788	1.051	1.756	1.591	1.939	1.600	1.425	1.796
	$\sigma_{o\xi}$	0.239	0.224	0.255	0.231	0.221	0.242	0.254	0.241	0.268

			Est.	L95	U95
All year	all division	σ_k	0.0310	0.025	0.039

Table A.4: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The between-individual variance parameter (σ_k) combined across year and division from RM 2 (Table A.3).

A.5 RM 4

Year	Parameter	2J Est	2J L95	2J U95	3K Est	3K L95	3K U95	3L Est	3L L95	3L U95
2009	l_∞	103.909	79.657	135.544	264.934	151.996	461.790	350.000	134.600	910.105
	k	0.138	0.095	0.201	0.039	0.020	0.076	0.031	0.011	0.091
	t_o	-0.333	-0.509	-0.157	-0.654	-0.862	-0.446	-0.332	-0.565	-0.100
	α	6.076	4.935	7.481	4.234	3.737	4.797	5.429	4.683	6.295
	β	0.435	0.349	0.542	0.889	0.777	1.018	0.627	0.535	0.734
	$\sigma_{o\xi}$	0.196	0.179	0.215	0.200	0.188	0.212	0.225	0.210	0.241
2010	l_∞	205.741	59.312	713.672	246.985	178.348	342.038	332.922	146.627	755.915
	k	0.051	0.012	0.228	0.037	0.025	0.055	0.028	0.011	0.071
	t_o	-0.962	-1.421	-0.502	-1.182	-1.321	-1.044	-1.010	-1.218	-0.802
	α	5.396	4.313	6.751	2.781	2.519	3.070	3.229	2.775	3.757
	β	0.541	0.426	0.688	1.321	1.183	1.476	1.141	0.966	1.348
	$\sigma_{o\xi}$	0.219	0.197	0.244	0.189	0.179	0.199	0.169	0.156	0.183
2011	l_∞	124.126	69.055	223.114	152.735	128.545	181.477	110.762	101.619	120.727
	k	0.096	0.042	0.222	0.074	0.058	0.095	0.126	0.110	0.145
	t_o	-0.788	-1.283	-0.294	-0.897	-1.110	-0.685	-0.293	-0.442	-0.144
	α	8.939	7.036	11.356	3.983	3.519	4.508	4.106	3.446	4.893
	β	0.326	0.254	0.418	1.127	0.985	1.288	1.031	0.855	1.245
	$\sigma_{o\xi}$	0.164	0.147	0.183	0.171	0.161	0.182	0.220	0.205	0.237
2012	l_∞	216.050	121.624	383.786	126.297	117.119	136.195	121.263	112.559	130.641
	k	0.053	0.026	0.106	0.107	0.095	0.120	0.114	0.102	0.128
	t_o	-0.821	-1.100	-0.541	-0.530	-0.629	-0.430	-0.344	-0.430	-0.259
	α	8.783	7.461	10.338	2.637	2.370	2.933	2.818	2.487	3.193
	β	0.398	0.335	0.471	1.648	1.461	1.858	1.486	1.293	1.709
	$\sigma_{o\xi}$	0.167	0.155	0.179	0.247	0.234	0.260	0.212	0.200	0.226
2013	l_∞	136.389	113.787	163.481	122.390	114.520	130.802	118.040	112.198	124.187
	k	0.099	0.078	0.127	0.112	0.100	0.124	0.121	0.111	0.132
	t_o	-0.401	-0.510	-0.293	-0.435	-0.511	-0.359	-0.192	-0.271	-0.114
	α	3.273	2.874	3.727	2.370	2.174	2.585	2.538	2.291	2.812
	β	0.910	0.788	1.051	1.757	1.592	1.940	1.605	1.429	1.802
	$\sigma_{o\xi}$	0.239	0.224	0.255	0.231	0.221	0.242	0.254	0.241	0.267

			Est.	L95	U95
All year	2J	σ_k	0.018	0.010	0.034
	3K		0.040	0.030	0.054
	3L		0.041	0.028	0.059

Table A.5: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The between-individual variance parameter (σ_k) combined across year for the same division from RM 2 (Table A.3).

A.6 RM 5

Year	Parameter	2J Est	2J L95	2J U95	3K Est	3K L95	3K U95	3L Est	3L L95	3L U95
2009	l_∞	102.302	78.142	133.933	262.519	148.277	464.780	350.000	134.441	911.182
	k	0.142	0.097	0.209	0.040	0.020	0.078	0.031	0.011	0.091
	t_o	-0.314	-0.495	-0.133	-0.660	-0.875	-0.445	-0.333	-0.564	-0.103
	α	6.112	4.936	7.569	4.235	3.730	4.808	5.419	4.681	6.273
	β	0.429	0.342	0.538	0.884	0.771	1.014	0.630	0.539	0.736
2010	l_∞	209.531	59.326	740.033	256.621	175.998	374.177	337.783	126.570	901.454
	k	0.050	0.011	0.226	0.036	0.023	0.055	0.028	0.009	0.085
	t_o	-0.973	-1.429	-0.517	-1.206	-1.357	-1.054	-0.998	-1.236	-0.759
	α	5.393	4.315	6.741	2.795	2.527	3.091	3.241	2.771	3.792
	β	0.542	0.427	0.688	1.302	1.163	1.457	1.119	0.940	1.331
2011	l_∞	136.534	61.034	305.428	156.499	126.397	193.769	110.790	101.742	120.643
	k	0.086	0.029	0.258	0.073	0.053	0.099	0.126	0.110	0.144
	t_o	-0.827	-1.413	-0.241	-0.912	-1.161	-0.663	-0.298	-0.445	-0.151
	α	9.148	7.046	11.876	4.012	3.526	4.564	4.097	3.443	4.876
	β	0.313	0.238	0.411	1.101	0.957	1.267	1.036	0.860	1.248
2012	l_∞	195.418	115.358	331.043	126.830	118.117	136.187	121.163	112.428	130.577
	k	0.061	0.031	0.118	0.105	0.094	0.117	0.115	0.102	0.129
	t_o	-0.742	-1.043	-0.441	-0.538	-0.632	-0.445	-0.345	-0.431	-0.259
	α	8.892	7.443	10.623	2.640	2.378	2.930	2.818	2.486	3.194
	β	0.386	0.321	0.465	1.666	1.482	1.873	1.485	1.291	1.707
2013	l_∞	136.068	114.581	161.585	123.019	115.289	131.269	118.027	112.572	123.746
	k	0.099	0.078	0.125	0.110	0.099	0.122	0.119	0.110	0.129
	t_o	-0.403	-0.506	-0.299	-0.437	-0.511	-0.363	-0.196	-0.269	-0.123
	α	3.266	2.875	3.711	2.376	2.180	2.589	2.534	2.292	2.800
	β	0.920	0.799	1.059	1.763	1.599	1.944	1.634	1.460	1.829

			Est.	L95	U95
All year	2J	σ_k	0.018	0.010	0.033
	3K		0.042	0.031	0.057
	3L		0.042	0.029	0.059

			Est.	L95	U95
All year	all division	$\sigma_{o\xi}$	0.216	0.212	0.219

Table A.6: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The ageing error intercept variance parameter ($\sigma_{o\xi}$) combined across year and division from RM 4 (Table A.5).

A.7 RM 6

Year	Parameter	2J Est	2J L95	2J U95	3K Est	3K L95	3K U95	3L Est	3L L95	3L U95
2009	l_∞	147.564	138.994	156.662	143.651	136.677	150.981	152.468	144.054	161.374
	t_o 2	-0.453	-0.572	-0.333	-0.347	-0.449	-0.245	-0.043	-0.171	0.086
	α	6.362	5.204	7.780	3.309	2.911	3.762	5.236	4.518	6.068
	β	0.413	0.335	0.511	0.921	0.804	1.055	0.651	0.556	0.763
	$\sigma_{o\xi}$	0.196	0.178	0.215	0.202	0.190	0.214	0.225	0.210	0.240
2010	l_∞	135.644	126.725	145.190	132.752	128.105	137.568	136.115	129.918	142.606
	t_o	-0.790	-0.980	-0.599	-0.793	-0.871	-0.716	-0.626	-0.717	-0.534
	α	5.282	4.232	6.594	4.104	3.622	4.650	3.029	2.605	3.521
	β	0.554	0.437	0.702	1.438	1.288	1.606	1.227	1.039	1.450
	$\sigma_{o\xi}$	0.220	0.198	0.245	0.192	0.182	0.202	0.174	0.161	0.189
2011	l_∞	134.516	125.448	144.240	137.930	132.590	143.486	143.747	136.694	151.164
	t_o	-0.813	-1.034	-0.593	-0.764	-0.874	-0.654	-0.419	-0.529	-0.310
	α	9.031	7.125	11.447	2.580	2.339	2.845	4.388	3.697	5.209
	β	0.322	0.252	0.412	1.145	1.003	1.307	0.957	0.797	1.148
	$\sigma_{o\xi}$	0.164	0.147	0.184	0.171	0.161	0.182	0.221	0.205	0.238
2012	l_∞	147.444	140.026	155.256	144.479	138.659	150.543	146.460	140.102	153.106
	t_o	-0.605	-0.752	-0.457	-0.637	-0.714	-0.559	-0.441	-0.509	-0.373
	α	8.613	7.310	10.148	3.925	3.475	4.433	2.954	2.612	3.340
	β	0.406	0.342	0.481	1.595	1.417	1.796	1.407	1.228	1.612
	$\sigma_{o\xi}$	0.168	0.156	0.181	0.246	0.234	0.260	0.213	0.200	0.226
2013	l_∞	150.984	143.725	158.610	144.278	138.568	150.223	145.825	139.090	152.886
	t_o	-0.442	-0.518	-0.367	-0.558	-0.612	-0.503	-0.341	-0.409	-0.273
	α	3.309	2.911	3.762	2.710	2.439	3.010	2.620	2.362	2.906
	β	0.899	0.780	1.035	1.712	1.552	1.889	1.545	1.375	1.736
	$\sigma_{o\xi}$	0.239	0.224	0.255	0.231	0.221	0.241	0.255	0.242	0.268

Year	Division	Parameter	Est.	L95	U95
All year	2J	σ_k	0.018	0.010	0.034
	3K		0.034	0.025	0.047
	3L		0.044	0.030	0.065
All year	all division	k	0.087	0.082	0.091

Table A.7: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The VonB growth parameter (k) combined across year and division from RM 4 (Table A.5).

A.8 RM 7

Year	Parameter	2J Est	2J L95	2J U95	3K Est	3K L95	3K U95	3L Est	3L L95	3L U95
2009	l_∞	103.909	79.657	135.546	174.405	139.689	217.749	350.000	198.306	617.734
	k	0.031	0.017	0.059	0.107	0.095	0.120			
	t_o	-0.333	-0.509	-0.157	-0.475	-0.639	-0.311	-0.331	-0.513	-0.150
	α	6.076	4.935	7.481	4.155	3.668	4.706	5.428	4.688	6.284
	β	0.435	0.349	0.542	0.909	0.794	1.040	0.627	0.536	0.733
	$\sigma_{o\xi}$	0.196	0.179	0.215	0.201	0.189	0.213	0.225	0.210	0.241
2010	l_∞	156.881	86.291	285.217	246.946	178.335	341.954	305.042	174.785	532.371
	k	0.118	0.110	0.127	0.138	0.095	0.201			
	t_o	-0.853	-1.194	-0.513	-1.182	-1.320	-1.044	-0.990	-1.159	-0.821
	α	5.336	4.270	6.667	2.781	2.519	3.070	3.218	2.769	3.741
	β	0.548	0.431	0.695	1.321	1.183	1.476	1.146	0.971	1.352
	$\sigma_{o\xi}$	0.220	0.198	0.244	0.189	0.179	0.199	0.169	0.157	0.183
2011	l_∞	153.639	85.228	276.962	152.729	128.539	181.471	110.762	101.619	120.727
	k	0.037	0.025	0.055	0.126	0.110	0.145	0.072	0.033	0.155
	t_o	-0.929	-1.325	-0.532	-0.897	-1.110	-0.685	-0.293	-0.442	-0.144
	α	9.037	7.127	11.459	3.983	3.519	4.508	4.106	3.446	4.893
	β	0.322	0.251	0.412	1.127	0.985	1.288	1.031	0.855	1.245
	$\sigma_{o\xi}$	0.165	0.147	0.184	0.171	0.161	0.182	0.220	0.205	0.237
2012	l_∞	216.015	121.640	383.613	126.297	117.119	136.195	118.681	113.176	124.453
	k	0.053	0.026	0.106	0.112	0.100	0.124			
	t_o	-0.821	-1.100	-0.541	-0.530	-0.629	-0.430	-0.327	-0.402	-0.252
	α	8.783	7.461	10.338	2.637	2.370	2.933	2.801	2.474	3.172
	β	0.398	0.335	0.471	1.648	1.461	1.858	1.496	1.302	1.719
	$\sigma_{o\xi}$	0.167	0.155	0.179	0.247	0.234	0.260	0.213	0.200	0.227
2013	l_∞	186.884	146.661	238.138	122.405	114.531	130.821	119.592	114.305	125.123
	k	0.066	0.049	0.089	0.074	0.058	0.095			
	t_o	-0.508	-0.608	-0.407	-0.435	-0.511	-0.359	-0.205	-0.276	-0.133
	α	3.368	2.959	3.833	2.370	2.174	2.585	2.544	2.297	2.819
	β	0.880	0.763	1.015	1.757	1.592	1.940	1.601	1.425	1.798
	$\sigma_{o\xi}$	0.239	0.224	0.255	0.231	0.221	0.242	0.254	0.241	0.267

Year	Division	Parameter	Est.	L95	U95
All year	2J	σ_k	0.018	0.010	0.034
	3K		0.040	0.030	0.053
	3L		0.039	0.027	0.059

Table A.8: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The VonB growth parameter (k) combined across 2009, 2010, 2012 and 2013 division 3K and 3L from RM 4 (Table A.5).

A.9 RM 8

Year	Parameter	2J Est	2J L95	2J U95	3K Est	3K L95	3K U95	3L Est	3L L95	3L U95
2009	k	0.088	0.083	0.094	0.093	0.093	0.087			
	t_o	-0.428	-0.558	-0.298	-0.255	-0.255	-0.346	-0.215	-0.328	-0.103
	α	6.360	5.207	7.767	4.192	4.192	3.714	4.635	4.043	5.313
	β	0.413	0.335	0.510	0.897	0.897	0.788	0.755	0.652	0.873
	$\sigma_{o\xi}$	0.196	0.179	0.215	0.202	0.202	0.190	0.226	0.211	0.242
2010	k	0.095	0.090	0.100	0.094	0.094	0.088			
	t_o	-0.830	-1.001	-0.658	-0.912	-0.912	-0.998	-0.575	-0.674	-0.477
	α	5.211	4.224	6.428	2.566	2.566	2.330	3.124	2.692	3.627
	β	0.563	0.451	0.703	1.456	1.456	1.306	1.176	0.999	1.386
	$\sigma_{o\xi}$	0.220	0.198	0.244	0.191	0.191	0.181	0.175	0.162	0.190
2011	k	0.077	0.073	0.082	0.092	0.092	0.086	0.083	0.077	0.090
	t_o	-0.837	-1.022	-0.651	-0.803	-0.803	-0.934	-0.382	-0.509	-0.255
	α	9.139	7.312	11.424	3.907	3.907	3.464	4.466	3.765	5.297
	β	0.318	0.252	0.400	1.152	1.152	1.011	0.936	0.781	1.122
	$\sigma_{o\xi}$	0.165	0.148	0.184	0.171	0.171	0.161	0.222	0.206	0.239
2012	k	0.094	0.087	0.101	0.093	0.093	0.087			
	t_o	-0.586	-0.754	-0.418	-0.595	-0.595	-0.681	-0.401	-0.468	-0.335
	α	8.459	7.193	9.947	2.707	2.707	2.439	2.949	2.613	3.330
	β	0.414	0.349	0.490	1.594	1.594	1.419	1.404	1.229	1.603
	$\sigma_{o\xi}$	0.168	0.156	0.181	0.247	0.247	0.234	0.213	0.200	0.227
2013	k	0.094	0.089	0.100	0.085	0.085	0.080			
	t_o	-0.459	-0.531	-0.387	-0.520	-0.520	-0.582	-0.292	-0.355	-0.229
	α	3.180	2.810	3.599	2.425	2.425	2.225	2.634	2.384	2.910
	β	0.947	0.828	1.085	1.705	1.705	1.547	1.532	1.372	1.711
	$\sigma_{o\xi}$	0.237	0.223	0.253	0.231	0.231	0.221	0.255	0.242	0.268

			Est.	L95	U95
All year	2J	σ_k	0.018	0.010	0.034
	3K		0.042	0.031	0.056
	3L		0.042	0.030	0.058

			Est.	L95	U95
All year	all division	l_∞	138.804	133.945	143.839

Table A.9: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The VonB growth parameter (l_∞) combined across year and division from RM 7 (Table A.8).

A.10 RM 9

		2J			3K			3L		
Year	Parameter	Est	L95	U95	Est	L95	U95	Est	L95	U95
2009	l_∞	128.083	95.646	171.522	172.165	148.028	200.238	189.408	157.421	227.896
	k	0.058	0.047	0.072	0.110	0.101	0.120			
	α	5.877	4.799	7.197	4.134	3.665	4.662	4.570	3.996	5.226
	β	0.450	0.363	0.558	0.914	0.803	1.041	0.758	0.656	0.876
	$\sigma_{o\xi}$	0.195	0.178	0.213	0.201	0.189	0.214	0.230	0.215	0.247
2010	l_∞	103.144	88.787	119.823	137.421	129.545	145.777	196.694	163.266	236.965
	k	0.092	0.085	0.101	0.101	0.071	0.142			
	α	5.530	4.474	6.835	2.945	2.664	3.255	3.716	3.194	4.324
	β	0.526	0.420	0.660	1.229	1.100	1.374	0.969	0.822	1.142
	$\sigma_{o\xi}$	0.224	0.201	0.248	0.197	0.187	0.208	0.180	0.166	0.195
2011	l_∞	102.483	88.141	119.158	124.557	116.424	133.259	117.633	106.422	130.025
	k	0.091	0.085	0.099	0.110	0.096	0.126	0.134	0.110	0.162
	α	9.295	7.437	11.615	4.021	3.550	4.554	3.627	3.088	4.260
	β	0.312	0.248	0.394	1.114	0.973	1.275	1.175	0.987	1.398
	$\sigma_{o\xi}$	0.165	0.148	0.184	0.173	0.162	0.184	0.219	0.203	0.235
2012	l_∞	151.127	124.154	183.962	124.422	117.102	132.199	137.635	128.678	147.216
	k	0.087	0.068	0.110	0.104	0.096	0.113			
	α	9.230	7.848	10.856	2.655	2.390	2.949	2.781	2.470	3.130
	β	0.377	0.319	0.446	1.634	1.452	1.840	1.508	1.322	1.721
	$\sigma_{o\xi}$	0.169	0.157	0.181	0.247	0.234	0.260	0.212	0.199	0.226
2013	l_∞	187.989	161.137	219.316	126.661	119.370	134.396	133.536	125.005	142.649
	k	0.066	0.056	0.079	0.106	0.096	0.116			
	α	3.403	3.015	3.841	2.349	2.155	2.560	2.307	2.096	2.538
	β	0.870	0.762	0.993	1.776	1.610	1.959	1.794	1.608	2.001
	$\sigma_{o\xi}$	0.240	0.225	0.255	0.231	0.221	0.242	0.257	0.244	0.271

			Est.	L95	U95
All year	2J	σ_k	0.018	0.010	0.034
	3K		0.033	0.023	0.047
	3L		0.059	0.045	0.076
			Est.	L95	U95
All year	all division	t_o	-0.495	-0.528	-0.462

Table A.10: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The VonB growth parameter (t_o) combined across year and division from RM 7 (Table A.8).

A.11 RM 10

Year	Parameter	2J Est	2J L95	2J U95	3K Est	3K L95	3K U95	3L Est	3L L95	3L U95
2009	l_∞	89.508	74.986	106.842	163.692	134.391	199.381	298.914	180.477	495.074
	k	0.035	0.020	0.062	0.102	0.090	0.116			
	t_o	-0.416	-0.546	-0.287	-0.525	-0.682	-0.368	-0.511	-0.688	-0.334
	β	0.815	0.753	0.883	1.157	1.091	1.227	1.048	0.984	1.117
	$\sigma_{o\xi}$	0.198	0.180	0.217	0.201	0.189	0.214	0.226	0.211	0.242
2010	l_∞	111.495	82.745	150.233	285.597	192.398	423.941	277.214	168.902	454.983
	k	0.116	0.107	0.125	0.161	0.124	0.210			
	t_o	-0.829	-1.110	-0.548	-1.194	-1.337	-1.050	-0.960	-1.129	-0.790
	β	0.902	0.825	0.986	1.114	1.056	1.174	1.123	1.047	1.204
	$\sigma_{o\xi}$	0.223	0.200	0.248	0.189	0.179	0.199	0.169	0.157	0.183
2011	l_∞	106.372	79.582	142.180	145.043	124.836	168.522	108.008	100.016	116.640
	k	0.032	0.020	0.050	0.128	0.113	0.145	0.109	0.070	0.168
	t_o	-1.019	-1.355	-0.683	-0.906	-1.112	-0.699	-0.376	-0.491	-0.260
	β	0.898	0.821	0.983	1.373	1.294	1.456	1.298	1.212	1.389
	$\sigma_{o\xi}$	0.167	0.149	0.187	0.171	0.161	0.182	0.220	0.204	0.237
2012	l_∞	140.332	108.403	181.665	131.647	121.184	143.013	121.350	115.324	127.692
	k	0.085	0.059	0.122	0.106	0.094	0.119			
	t_o	-0.917	-1.138	-0.696	-0.490	-0.592	-0.388	-0.307	-0.382	-0.233
	β	1.078	1.009	1.152	1.312	1.241	1.387	1.270	1.195	1.350
	$\sigma_{o\xi}$	0.170	0.158	0.184	0.247	0.234	0.261	0.213	0.200	0.227
2013	l_∞	177.706	143.096	220.687	128.998	119.612	139.119	123.546	117.528	129.872
	k	0.070	0.054	0.092	0.079	0.063	0.099			
	t_o	-0.500	-0.596	-0.404	-0.385	-0.464	-0.306	-0.147	-0.217	-0.078
	β	0.904	0.850	0.962	1.252	1.191	1.317	1.228	1.163	1.297
	$\sigma_{o\xi}$	0.239	0.224	0.255	0.232	0.222	0.242	0.255	0.242	0.268

			Est.	L95	U95
All year	2J	σ_k	0.018	0.010	0.034
	3K		0.043	0.033	0.057
	3L		0.040	0.027	0.058

			Est.	L95	U95
All year	all division	α	3.283	3.175	3.395

Table A.11: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The true age distribution parameter (α) combined across year and division from RM 7 (Table A.8).

A.12 RM 11

		2J			3K			3L		
Year	Parameter	Est	L95	U95	Est	L95	U95	Est	L95	U95
2009	l_∞	83.622	72.064	97.034	156.389	130.735	187.077	278.791	175.371	443.200
	k	0.037	0.022	0.063	0.100	0.087	0.115			
	t_o	-0.432	-0.544	-0.319	-0.498	-0.653	-0.343	-0.531	-0.709	-0.354
	α	2.564	2.371	2.773	3.342	3.169	3.525	3.149	2.971	3.337
	$\sigma_{o\xi}$	0.199	0.181	0.218	0.201	0.189	0.213	0.225	0.210	0.241
2010	l_∞	103.273	80.044	133.244	271.993	188.599	392.262	262.552	166.552	413.887
	k	0.115	0.106	0.125	0.174	0.138	0.219			
	t_o	-0.846	-1.119	-0.572	-1.189	-1.331	-1.048	-0.952	-1.119	-0.785
	α	2.775	2.550	3.020	3.077	2.931	3.231	3.163	2.965	3.373
	$\sigma_{o\xi}$	0.223	0.201	0.248	0.189	0.179	0.200	0.169	0.157	0.183
2011	l_∞	99.283	77.466	127.243	151.893	128.667	179.312	109.046	100.756	118.018
	k	0.033	0.022	0.051	0.127	0.111	0.145	0.118	0.080	0.174
	t_o	-1.012	-1.338	-0.686	-0.912	-1.124	-0.699	-0.338	-0.464	-0.213
	α	2.856	2.621	3.112	3.857	3.660	4.065	3.674	3.454	3.907
	$\sigma_{o\xi}$	0.167	0.149	0.187	0.171	0.160	0.182	0.220	0.204	0.237
2012	l_∞	137.329	107.309	175.747	134.992	123.578	147.461	122.547	116.309	129.120
	k	0.086	0.061	0.122	0.105	0.093	0.119			
	t_o	-0.925	-1.146	-0.705	-0.470	-0.573	-0.366	-0.295	-0.371	-0.219
	α	3.327	3.131	3.535	3.510	3.339	3.688	3.446	3.262	3.640
	$\sigma_{o\xi}$	0.170	0.157	0.183	0.250	0.237	0.264	0.214	0.201	0.228
2013	l_∞	165.870	136.149	202.078	130.905	121.039	141.577	124.809	118.537	131.413
	k	0.075	0.058	0.096	0.074	0.058	0.095			
	t_o	-0.540	-0.636	-0.445	-0.377	-0.456	-0.298	-0.144	-0.213	-0.075
	α	2.659	2.508	2.819	3.311	3.165	3.464	3.300	3.141	3.466
	$\sigma_{o\xi}$	0.238	0.223	0.254	0.234	0.224	0.245	0.257	0.244	0.271

			Est.	L95	U95
All year	2J	σ_k	0.018	0.010	0.034
	3K		0.043	0.033	0.057
	3L		0.040	0.027	0.058
			Est.	L95	U95
All year	all division	β	1.169	1.127	1.213

Table A.12: Structural errors-in-variables (SEV) parameter estimation results (Est.) of northern cod and 95% confidence interval (L,U) by year (2009-2013) and NAFO division (2J3KL). The true age distribution parameter (β) combined across year and division from RM 7 (Table A.8).

A.13 RM 4 estimates plot.

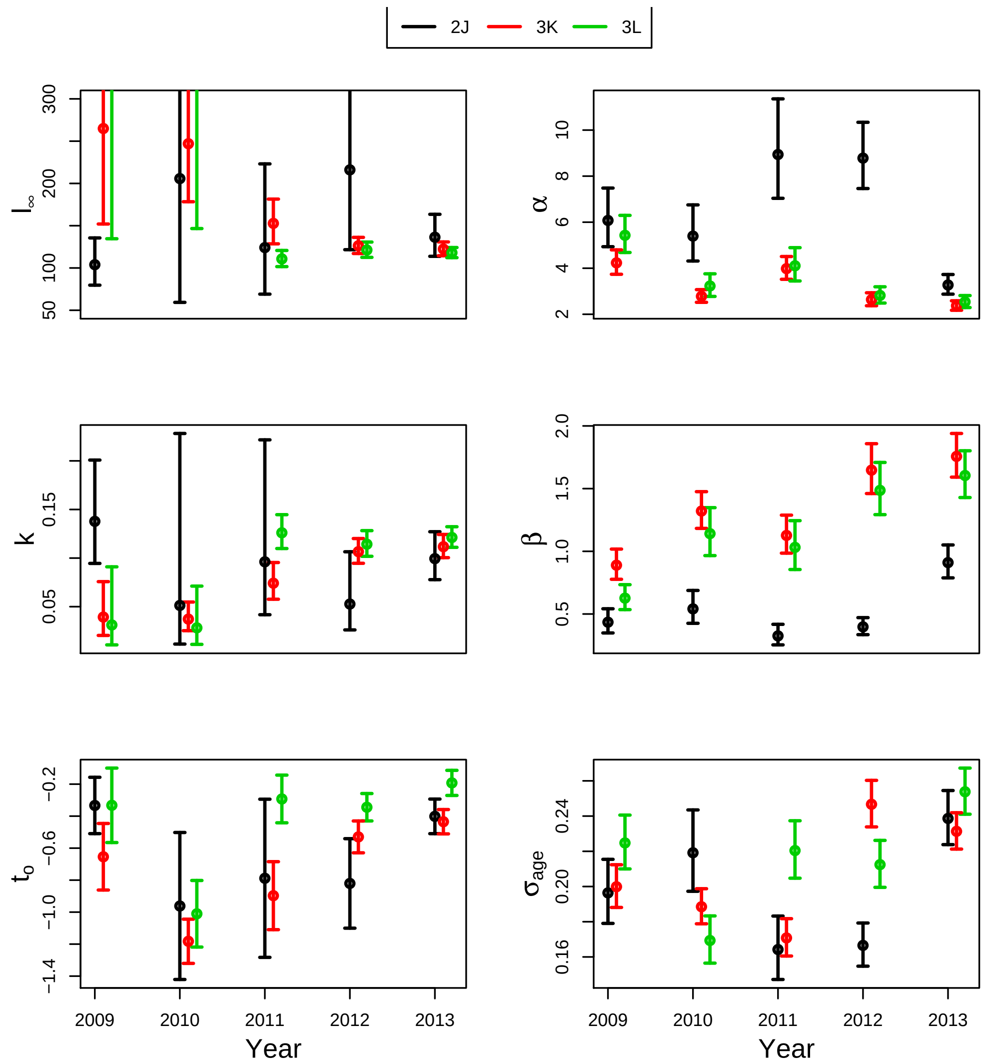

Figure A.1: Graph by year and NAFO division of reduced model 4 results in Table A.6.

A.14 RM 7 estimates plot.

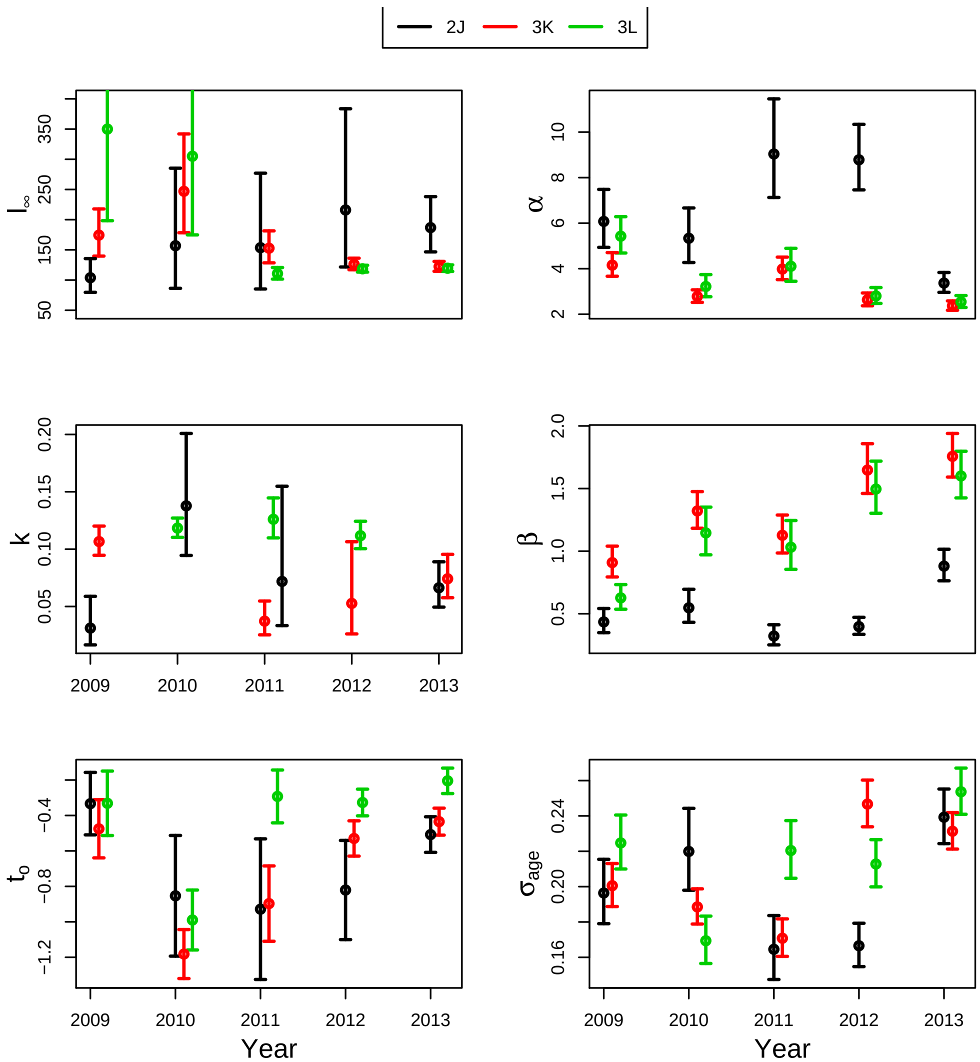

Figure A.2: Graph by year and NAFO division of reduced model 7 results in Table A.7.

Bibliography

[1] Alos J., Palmer M., Balle S., Grau A. M., and Morales-Nin B. (2010). Individual growth pattern and variability in Serranus scriba: a Bayesian analysis, *ICES Journal of Marine Science*, **67**, 502-512.

[2] Barndorff-Nielsen O.E, and Cox D.R. (1994). Inference and Asymptotics, *London: Chapman and Hall.*

[3] Berger J. O., Liseo, B. & Wolpert R. L. (1999). Integrated likelihood methods for eliminating nuisance parameters, *Statistical Science*, **14**(1), 1-28.

[4] Beverton R. J. H. and Holt S. J. (1957). On the dynamics of exploited fish populations, *Fishery investigations, London,* **Series II XIX**,(533 pp).

[5] Candy S. G., Constable A.J., Lamb T. and Williams R. (2007). A Von Bertalanffy growth model for toothfish at Heard Island fitted to length-at-age data and compared to observed growth from mark-recapture studies,*CCAMLR Science*, **14**, 4366.

[6] Candy S.G. (2005). Fitting a Von Bertalanffy growth model to length-at-age data accounting for length-dependent fishing selectivity and length-stratified sub-sampling of length frequency samples, Document WG-FSA-SAM-05/13. *CCAMLR, Hobart, Australia.*

[7] Carroll R. J., Ruppert D., Stefanski L. A, Crainiceanu C. M. (2006). Measurement error in nonlinear models, A modern perspective, *Chapman & Hall/CRC,Taylor & Francis Group,Boca Raton,London, New York.*

[8] 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 , *Can. J. Fish. Aquat. Sci.*, **64**.

[9] Cox D.R. (2006). Principles of Statistical Inference, *Cambridge: Cambridge University Press.*

[10] Filipe P. A., Braumann C. A., Roquete C. J. (2010). Multiphasic Individual Growth Models in Random Environments, *Springer Science+Business Media, LLC.*

[11] Fournier D. A., Skaug H. J., Ancheta J., Ianelli J., Magnusson A., Maunder M. N. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models, *Optimization Methods and Software*, **27**,(2), 233-249.

[12] Goodyear P. C. (1995). Mean size at age: An evaluation of sampling strategies with simulated Red Grouper data, *Trans. Am. Fish. Soc.*, **124**, 746755.

[13] Gulland, J. A. (1983). Fish stock assessment: a manual of basic methods, *Chichester, U.K., Wiley Interscience, FAO/Wiley series on food and agriculture*, **1:223 p**.

[14] Hajek J. (1971). Comments on a paper by Basu D. *In Foundations of statistical inference* (Godambe, V.P. and sprott, D.A. eds.), *Holt, Rienhart and Winston, Toronto.*

[15] Heifetz J., Delsa Anderl, Nancy E. Maloney and Thomas L. Rutecki (1998). Age validation and analysis of ageing error from marked and recaptured sablefish, Anoplopoma fimbria,*Fish. Bull.*, **97**, 256-263.

[16] Horvitz D.G., and Thompson D.J. (1952). A generalization of sampling without replacement from a finite universe, *JASA*, **47**, 663685.

[17] Hu N. (2013). Statistical analysis of length-biased and right-censored data, Available at https://books.google.ca/books?id=x296oAEACAAJ.

[18] Kimura D. K. (1990). Approaches to age-structured separable sequential population analysis, *Can. J. Fish. Aquat. Sci.*, **47**, 23642374

[19] Kitakado T. (2001). Application of the conditional score method for estimating a fish growth curve, *Japanese Journal of Biometrics*, **21**,(2),29-39.

[20] Kristensen K., Nielsen A., Berg C. W., Skaug H., and Bell B. (2015). TMB: Automatic Differentiation and Laplace Approximation, *arXiv preprint arXiv:1509.00660.*

[21] Morales-Nin B. (1992). Determination of growth in bony fishes from otolith microstructure, *FAO Fisheries Technical Paper* 322, Rome, FAO.

[22] Pauly D. (1984). Fish population dynamics in tropical waters: a manual for use with programmable calculators, *ICLARM Stud.Rev.*, 8, 325 p.

[23] Pauly D. and Morgan G. R. (1987). Length-based methods in fisheries research, *ICLARM Conf.Proc.*, 13, 468 p.

[24] Quist M.C., M.A. Pegg, and D.R. DeVries (2012). Age and Growth In: Fisheries Techniques, Third Edition (Zale, A.V., D.L. Parish and T.M. Sutton Eds.), *American Fisheries Society, Bethesda, Maryland*. 677-731

[25] Rao J. N. K., Arijit Chaudhuri, J. L. Eltinge, Robert E. Fay, J. K. Ghosh, Malay Ghosh, P. Lahiri and Danny Pfeffermann (1999). Some current trends in sample survey theory and methods (with discussion). *Sankhya*, **61**, 157.

[26] Rao J.N.K. (1966). Alternative estimators in PPS sampling for multiple characteristics. *Sankhya*, **A28**, 47-60.

[27] R Core Team (2014). R: A Language and Environment for Statistical Computing., *R Foundation for Statistical Computing, Vienna, Austria.*

[28] Ricker W.E., (1975). Computation and interpretation of biological statistics of fish populations, *Bull.Fish.Res.Board Can.*, **191**, 382p.

[29] Sainsbury J. (1980). Effect of individual variability on the von Bertalanffy growth equation. *Can. J. Fish. Aquat. Sci.*, **37**, 241-247.

[30] Searle S. R., Casella G., & McCulloch C. E. (1998). Variance components (Vol. 391), *John Wiley & Sons.*

[31] Severini T.A. (2000) Likelihood Methods in Statistics, *Oxford: Oxford University Press.*

[32] Shackell N. L., Stobo W. T., Frank K. T., and Brickman D. (1997). Growth of cod (Gadus morhua) estimated from mark-recapture programs on the Scotian Shelf and adjacent areas, *ICES Journal of Marine Science*, **54**, 383398.

[33] Shelton A. O. and Mangel M. (2012). Estimating von Bertalanffy parameters with individual and environmental variations in growth, *Journal of biological dynamics,* **6**,(2).

[34] Skaug H. J., & Fournier, D. A. (2006) Automatic approximation of the marginal likelihood in non-gaussian hierarchical models, *Computational Statistics and Data Analysis*, **51**,(2), 699-709.

[35] Smith E.B., Williams F.M., and Fisher C.R. (1997). Effects of intrapopulation variability on von Bertalanffy growth parameter estimates from equal markrecapture intervals, *Can. J. Fish. Aquat. Sci.*, **54**, 20252032.

[36] Sparre P. and Venema S. C. (1998). Introduction to tropical fish stock assessment, *FAO fisheries technical paper 306/1 Rev.2.*

[37] Taylor N. G., C. J. Walters, and S. J. D. Martell (2005). A new likelihood for simultaneously estimating Von Bertalanffy growth parameters, gear selectivity, and natural and fishing mortality, *Can. J. Fish. Aquat. Sci.*, **62**, 215223.

[38] Ursin E., (1968). A mathematical model of some aspects of fish growth, respiration and mortality, *J.Fish.Res.Board Can.*, **24**, 2355-2453.

[39] Walsh S. J. (1996). Efficiency of bottom sampling trawls in deriving survey abundance indices, *NAFO Sci. Coun. Studies*, **28**, 924.

[40] Wang You-Gan, Mervyn R, and Ian F. Somers (1995). A maximum likelihood approach for estimating growth from tag-recapture data. *Can. J. Fish. Aquat. Sci.*, **52**, 252-259.