



A spatiotemporal Richards–Schnute growth model and its estimation when data are collected through length-stratified sampling

Nan Zheng¹ · Noel Cadigan¹ · M. Joanne Morgan²

Received: 30 August 2019 / Revised: 15 May 2020 / Published online: 13 June 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

We propose a spatiotemporal generalized von Bertalanffy (vonB) growth model that also includes between-individual (BI) variation and male/female correlation. The generalized vonB model includes the effect of maturation on growth. The model and the methodology are applied to a long time-series of survey observations of age and length for American plaice on the Grand Bank off the northeast coast of Canada. The bias in age-length data due to size selectivity of the survey gear is accounted for. The survey design includes length-stratified age sampling which is a type of response selective sampling design for growth model estimation. We propose and implement a conditional empirical proportion likelihood approach for these data. Neglecting this sampling scheme can lead to seriously biased estimation results. We found that a 6-parameter growth model is necessary for capturing the biphasic growth patterns of the American plaice on the Grand Bank, and the survey gear selectivity and BI variation are important for a good model fit. We proposed an empirically optimal BI variation model for this data. Our estimation results indicate that there are substantial differences in size-at-age for male and female American plaice, and this changes over time and between regions.

Keywords Between-individual variation · Growth model · Length-stratified sampling · Spatiotemporal model · Template model builder · Von Bertalanffy

Handling Editor: Bryan F. J. Manly.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10651-020-00450-8>) contains supplementary material, which is available to authorized users.

✉ Nan Zheng
Nan.Zheng@mi.mun.ca

Extended author information available on the last page of the article

1 Introduction

Estimation of body growth is a fundamental part of fish stock assessment and the determination of sustainable levels of fishing. One of the most widely used growth models in fisheries science is the von Bertalanffy (vonB) model (Von Bertalanffy 1960) because it conforms to the observed growth of many fish species (e.g. Quist et al. 2012). The basic vonB model is

$$y(a) = l_{\infty} \left(1 - e^{-k(a-a_0)} \right), \quad (1)$$

where $y(a)$ denotes length at age a , l_{∞} is the maximum possible size (as $a \rightarrow \infty$), $k \geq 0$ is a growth rate parameter, and $a_0 (< 0)$ is the theoretical age at which the fish would have had zero length. The length at $a = 0$ is $l_0 = l_{\infty}(1 - e^{ka_0})$ and an alternative parameterization of the vonB model we use is

$$y(a) = l_{\infty} \left(1 - [1 - l_0/l_{\infty}] e^{-ka} \right). \quad (2)$$

Fish size-at-age may vary spatially and temporally and for different individuals for many reasons, including effects due to changes in water temperatures, maturation age, reproductive investment, activity levels, and prey quality and availability (e.g. Webber and Thorson 2016). Spatial separation of fish often leads to genetic and environmental differences, which in turn may result in variation in size-at-age and growth model parameters between fish in different areas. On the other hand, growth rates of fish may have spatial correlations, depending on distance, closeness to major ocean currents and lands, etc. Inclusion of such spatiotemporal correlations can lead to more efficient statistical inferences about size-at-age. Understanding spatiotemporal variation in size-at-age is also important for fish stock assessment models (e.g. Quinn and Deriso 1999) and especially spatial assessment models.

Growth models are usually estimated from samples in which fish length and age are measured. The samples usually contain data from multiple cohorts and this is another possible source of variability. Different cohorts may grow at different rates due to genetic and environmental variation, while individuals within the same cohort may have more similar growth rates. However, there will still be between-individual (BI) variation in growth among the individuals even in the same cohort. Failure to account for BI variation can lead to bias in estimating the population mean growth parameters and length-at-age, as noted by Sainsbury (1980) and Wang et al. (1995). In Sect. 4 we will discuss that an incorrect BI variation model combined with survey gear selectivity can lead to biased estimation of growth model parameters and consequently size-at-age. Variation in growth is also important for population and community dynamics (e.g. Shelton et al. 2013).

A focus of this paper is to analyze size-at-age for American plaice (*Hippoglossoides platessoides*) using a large dataset collected by Fisheries and Oceans Canada (DFO) during spring and fall surveys from 1978 to 2015 in Northwest Atlantic Fisheries Organization (NAFO) Divisions 3L, 3N and 3O off the east coast of Canada. These three spatial divisions are all neighbouring to each other, but also large enough to have

differences in environmental factors such as water temperature, depth, and ocean currents. We use these NAFO Divisions for the spatial aspect of our model, and investigate how growth rates change over time, or more specifically, over cohorts. We account for division and cohort variations using a mixed-effects modelling approach, where the vonB parameters for individual cohorts are considered to be random variables that are spatially and temporally correlated.

The data we analyze were obtained through two-phase length-stratified age sampling (LSAS), where in the first phase a large number of fish were randomly caught each year, measured for length, and divided into length strata of 2 cm. In the second phase, a fixed number of fish in each length stratum were randomly selected for age measurement which involves counting growth rings in fish ear bones and is much more expensive than length measurement. Such a sampling strategy can be effective in lowering estimation error, reducing age sampling expenses, improving population representativeness, and enhancing measurement manageability. However, Zheng and Cadigan (2019) demonstrated that growth model parameter estimation can be seriously biased based on LSAS data without accounting for the sampling strategy. The simulation results in Goodyear (2019) also indicate this. Zheng and Cadigan (2019) proposed two new inference approaches, namely, the full-data likelihood approach and the empirical proportion (EP) likelihood approach. They demonstrated that these two methods performed the best among several existing approaches in their simulation study. However Perreault et al. (2019) showed that the EP approach was more practically useful and performed best compared to a wide range of other methods used or applicable for dealing with LSAS in fisheries. This paper is focused on spatial time-series of growth data, and we apply the EP likelihood approach because it performs well, and is easier to implement compared to the full-data likelihood approach of Zheng and Cadigan (2019).

In practice fish are caught using sampling trawls that are size-selective to varying degrees (e.g. Walsh 1997). Size (i.e. length) selective samples are biased towards the size selectivity of the fishing gear used to catch the fish when there is BI variation in size at age, and this can result in biased estimates of the vonB growth parameters if the selectivity is not accounted for (e.g. Goodyear 1995; Candy 2005; Candy et al. 2007; Schueller et al. 2014). To get unbiased and more precise estimates of the vonB growth parameters, the model fitting procedure has to be extended to take into account the size selectivity of the fishing gear used to catch the fish. Candy et al. (2007) used a step function to describe the size selectivity function. We use a similar approach to Candy et al. (2007); however, we assume the selectivity function has a linear logistic form.

One drawback of the vonB model (1) is that it does not explicitly incorporate changes in growth after a fish matures, when surplus energy, especially for female fish, has to be allocated to both reproduction and growth, whereas before maturation more surplus energy is allocated to growth. This cost of reproduction should be considered when modeling fish growth (e.g., Brody 1945; Roff 1983; Roff et al. 2000; Lester et al. 2004, 2014), as it may result in different growth patterns in juvenile and adult stages. This biphasic growth can be modeled by two separate growth equations, each for a specific growth phase (e.g., Day and Taylor 1997; Soriano et al. 1992), or by a continuous function with a smooth transition between the two growth phases

(e.g., Laslett et al. 2002; Quince et al. 2008). Minte-Vera et al. (2016) presented a review of the existing biphasic growth models, and proposed two new models: the von Bertalanffy logistic- L_∞ (VB log- L_∞) and the Cost of Reproduction (CoR) models. In this study we applied the VB log- L_∞ model

$$y(a) = \tilde{l}_\infty \left(1 - \left[1 - l_0/\tilde{l}_\infty \right] e^{-ka} \right), \quad (3)$$

where $\tilde{l}_\infty = l_\infty \{1 + p(a)\Delta\} / \{1 + \Delta\}$, and Δ and $p(a)$ are defined in Sect. 2. We chose this approach among all the existing biphasic growth models because, (i) VB log- L_∞ is easier to implement, (ii) its performance is among the best under various circumstances (Minte-Vera et al. 2016), and (iii) it can estimate both growth and maturity parameters with only length-at-age data (no maturity-at-age data), which can be the case in many studies.

The outline of this article is as follows. In Sect. 2.1 we present a spatiotemporal growth model incorporating temporal cohort effects, spatial correlation between different divisions and between males and females, between-individual variation and reproduction cost. In Sect. 2.2, for data collected from LSAS, we extend the EP likelihood method as a conditional EP likelihood approach and for quick processing of a large spatiotemporal dataset. The issues with size-selective capture and calculation of model residuals for length-stratified sampling are addressed in Sects. 2.2.1 and 2.2.3 respectively. In Sect. 3 we apply the proposed models and estimation approach to the 3LNO American plaice data. Some further discussion of various issues in modeling and estimation is provided in Sect. 4.

2 Methods

In this section we propose a spatiotemporal model incorporating temporal linkage of sequential cohort effects, divisional and male/female correlations, between-individual variation and biphasic growth due to reproduction. We investigate the estimation of the model parameters with LSAS data. This study is aimed at 3LNO American plaice. Age determination for American plaice in Canadian waters has been validated and considered to be accurate (Morin et al. 2013). Length can usually be even more accurately determined. Therefore, in this paper we assume that the length and age measurement errors (MEs) are negligible and do not distinguish measured and true values. However, we provide some discussion about length and age MEs in Sect. 4. In the following all the lengths are in centimeters (cm).

Let the random variables $L_{\infty,c,d,g}$ and $K_{c,d,g}$ denote respectively the vonB asymptotic length and growth rate of the c 'th cohort for males or females ($g = M$ or F) in division d , where the upper case notation is used to denote random variables. Because male and female American plaice follow different growth patterns, we model them separately with different vonB parameters. In preliminary analysis the vonB model (1) could not fit the American plaice data reasonably well at younger and older ages at the same time. Similar fitting issues occurred with the other growth models such as Richards–Schnute growth models (Richards 1959; Schnute 1981). We obtained better

results when we explicitly accounted for the cost of reproduction in growth by applying the VB log- L_∞ model (3) proposed in Minte-Vera et al. (2016). To further increase the flexibility of the growth model in fitting large scale multiple-cohort length and age data, we combined the VB log- L_∞ model with the simplified Richards–Schnute (RS) model (e.g. Cadigan and Campana 2017), and proposed a RS log- L_∞ growth model for the population mean length at age a ,

$$\mu_{c,d,g}(a) = \tilde{L}_{\infty,c,d,g,a} \left\{ 1 - \left(1 - l_0 / \tilde{L}_{\infty,c,d,g,0} \right) e^{-K_{c,d,g} a^{\lambda_{d,g}}} \right\}, \quad (4)$$

where the subscripts c , d , g and a indicate respectively cohort, division, sex and age, and

$$\tilde{L}_{\infty,c,d,g,a} = \frac{L_{\infty,c,d,g}}{1 + \Delta_{\infty,d,g}} \left\{ 1 + p(a) \Delta_{\infty,d,g} \right\}.$$

The factor

$$p(a) = \frac{1}{1 + \exp(-\log(19)(a - a_{50,d,g})/(a_{95,d,g} - a_{50,d,g}))} \quad (5)$$

can be assumed to be a consequence of maturity (Minte-Vera et al. 2016). If $p(a)$ is interpreted as the proportion of fish mature at age a , $a_{50,d,g}$ and $a_{95,d,g}$ are respectively the ages at 50% and 95% maturity. Hence $a_{50,d,g}$ is constrained to be smaller than $a_{95,d,g}$. For younger age (juvenile fish), $p(a)$ is close to 0, and $\tilde{L}_{\infty,c,d,g,a}$ becomes $L_{\infty,c,d,g}/(1 + \Delta_{\infty,d,g})$. For older age (adult fish), $p(a)$ goes to 1, and $\tilde{L}_{\infty,c,d,g,a} \approx L_{\infty,c,d,g} = L_{\infty,c,d,g}/(1 + \Delta_{\infty,d,g}) + \Delta_{\infty,d,g} L_{\infty,c,d,g}/(1 + \Delta_{\infty,d,g})$. Therefore, $\Delta_{\infty,d,g}$ is the difference between the L_∞ 's of mature and juvenile fish as a proportion of the L_∞ of juvenile fish.

In (4) the powers $\lambda_{d,g}$ on age a represent an extension from the vonB model to the simplified RS model (Cadigan and Campana 2017). For the American plaice data used in this study (Sect. 3.1), the RS log- L_∞ model achieved significantly better fittings than the VB log- L_∞ model in terms of model selection criteria AIC (Akaike information criterion) and BIC (Bayesian information criterion), as well as model validation criteria such as residuals. In (4) and (5) $L_{\infty,c,d,g}$ and $K_{c,d,g}$ are random effects to incorporate spatiotemporal correlations; $\Delta_{\infty,d,g}$, $a_{50,d,g}$, $a_{95,d,g}$ and $\lambda_{d,g}$ are model parameters to be estimated using the maximum likelihood approach together with the spatiotemporal model parameters, the between-individual variation parameters and the selectivity parameters introduced later.

2.1 Spatiotemporal growth model

The cohort effects $L_{\infty,c,d,g}$ and $K_{c,d,g}$ are assumed to follow stationary log-AR(1) models (autoregressive model of order 1, e.g. Mills and Mills 1991),

$$\begin{aligned}
\log(L_{\infty,c,d,g}) &= \mu_{\infty,d,g} + \Gamma_{\infty,c,d,g}, \\
\Gamma_{\infty,c,d,g} &= \varphi_{\infty,d,g} \Gamma_{\infty,c-1,d,g} + \varepsilon_{\infty,c,d,g}, \quad \Gamma_{\infty,c,d,g} \in \mathcal{R}, \\
\log(K_{c,d,g}) &= \mu_{k,d,g} + \Gamma_{k,c,d,g}, \\
\Gamma_{k,c,d,g} &= \varphi_{k,d,g} \Gamma_{k,c-1,d,g} + \varepsilon_{k,c,d,g}, \quad \Gamma_{k,c,d,g} \in \mathcal{R},
\end{aligned} \tag{6}$$

where m_c is the total number of cohorts in the data, $\varepsilon_{\infty,c,d,g}$ and $\varepsilon_{k,c,d,g}$ are respectively $N(0, \delta_{\infty,d,g}^2)$ and $N(0, \delta_{k,d,g}^2)$, and $\mu_{\infty,d,g}$, $\varphi_{\infty,d,g}$, $\delta_{\infty,d,g}$, $\mu_{k,d,g}$, $\varphi_{k,d,g}$ and $\delta_{k,d,g}$ are model parameters. Some properties for the stationary AR(1) models are provided in Appendix A. The autocorrelations are

$$\begin{aligned}
\text{corr}(\log(L_{\infty,c-h,d,g}), \log(L_{\infty,c,d,g})) &= \varphi_{\infty,d,g}^h, \\
\text{corr}(\log(K_{c-h,d,g}), \log(K_{c,d,g})) &= \varphi_{k,d,g}^h;
\end{aligned}$$

that is, $\varphi_{\infty,d,g}$ and $\varphi_{k,d,g}$ are the lag 1 autocorrelations, and the autocorrelations decay with lag h as a power function.

Correlations in size-at-age between males and females and between spatial divisions are introduced through correlations among the error terms in (6)

$$\begin{aligned}
\text{corr}(\varepsilon_{\infty,c,d,g}, \varepsilon_{\infty,c',d',g'}) &= \rho_{\infty,dg,d'g'}, \\
\text{corr}(\varepsilon_{k,c,d,g}, \varepsilon_{k,c',d',g'}) &= \rho_{k,dg,d'g'}.
\end{aligned} \tag{7}$$

In the case study of Sect. 3 with 3 divisions (3L, 3N and 3O), there are 15 correlation parameters for each of L_{∞} (Table 4) and K (Table 5).

BI variation of an arbitrary individual i with age a and sex g in cohort c and division d can be represented by the standard deviation (SD) $\sigma_{c,d,g,a}$ of its length $y_i(a)$. Some authors (e.g. Piner et al. 2016) modeled $\sigma_{c,d,g,a} \propto \mu_{c,d,g}(a)$; however, in preliminary analyses of the American plaice data we found that this approach produced too large SD values at older ages. Some other authors (e.g. Prajneshu and Venugopalan 1999; Russo et al. 2009; Filipe et al. 2012) proved with stochastic growth models that the SD of BI variation monotonically increases, or non-monotonically converges, to a constant as age increases. However, direct implementation of these theoretical SD models did not lead to satisfying model fitting. After testing a series of SD models with the American plaice data, we found that the BI SD should increase first convexly and then concavely with the mean length, and finally converge to a constant. We propose the following SD model

$$\begin{aligned}
\sigma_{c,d,g,a} &= \text{CV}_{d,g} \times \mu_{c,d,g}(a) + h_{d,g}(\mu_{c,d,g}(a)) \times v_{d,g}, \\
h_{d,g}(\mu) &= \left\{ \frac{2}{1 + \exp(-\log(19) [\mu - l_0]/[l_{95,d,g} - l_0])} - 1 \right\}^{\tau_{d,g}}
\end{aligned} \tag{8}$$

with $\text{CV}_{d,g}$, $v_{d,g}$, $\tau_{d,g}$ and $l_{95,d,g}$ being positive parameters for each division and sex and $l_{95,d,g} > l_0$. The estimates of $\text{CV}_{d,g}$ are virtually zero for the American plaice data. We kept this CV term because this is a quite popular model for the variability in growth as well as observation error and hence we need to test its negligibility with the

data fitting results. When $CV_{d,g}$ is negligible, $v_{d,g}$ is the asymptotic value of $\sigma_{c,d,g,a}$ as mean length $\mu_{c,d,g}(a)$ goes to infinity, and $h_{d,g}(\cdot)$ gives the desired pattern of convergence to $v_{d,g}$.

Note that $\sigma_{c,d,g,a}$ actually may also include the SD of length ME (ME for measurement error). The length ME of the data in this study is negligible, hence here $\sigma_{c,d,g,a}$ mainly represents BI variation. Some authors model BI variation via random growth model parameters that vary from individual to individual. For instance, in Cadigan and Campana (2017) the BI variation is explicitly expressed as the BI variation in asymptotic length $L_{\infty,i}$, and the BI variation in growth rate K_i . Here the subscript i enumerates the individuals. In this work we simply model BI variation as a combined variation which may originate from many sources such as the BI variations in individual $L_{\infty,i}$, K_i , age zero length $l_{0,i}$, and etc.

The positive $y_i(a)$ can be conveniently modeled using a log-normal distribution with mean $\mu_{c,d,g}(a)$ and SD $\sigma_{c,d,g,a}$,

$$\begin{aligned}\log(y_i(a)) &= \log(\mu_{c,d,g}(a)) - \log(\sigma_{c,d,g,a}^2 / \mu_{c,d,g}^2(a) + 1) / 2 + \gamma_i, \\ \gamma_i &\sim N(0, \log(\sigma_{c,d,g,a}^2 / \mu_{c,d,g}^2(a) + 1)).\end{aligned}\quad (9)$$

It is straightforward to check that for log-normal model (9), the mean of $y_i(a)$ is $\mu_{c,d,g}(a)$ modeled by (4), and the SD of $y_i(a)$ is $\sigma_{c,d,g,a}$ modeled by (8).

2.2 Empirical proportion likelihood approach for length-stratified data

In this section we implement the empirical proportion (EP) likelihood approach proposed in Zheng and Cadigan (2019) to deal with LSAS. Here we focus on the case where the age measurement error is negligible, which is true for the American plaice data of this study. For addressing LSAS when the age measurement error has to be accounted for, we refer to Zheng and Cadigan (2019) when a valid model for age distribution is available, and Perreault et al. (2019) when such a model is unavailable. The simulation studies in Zheng and Cadigan (2019) and Perreault et al. (2019) indicated that the EP likelihood approach gives good estimation of model parameters and performs better than many other approaches.

Let S_j denote the j th length stratum (or interval) of usually 2 or 3 cm in practice. Here $j = 1, \dots, H_{\text{obs}}$ enumerate the strata with data observed, and $j = H_{\text{obs}} + 1, \dots, H_{\text{total}}$ enumerate the strata without data, namely the empty strata. H_{total} is the total number of strata with nonnegligible occupation probabilities. The EP density for the LSAS length-at-age data is

$$f_{\text{EP}}(Y = y \mid A = a; \text{LSAS}, \boldsymbol{\theta}) = \frac{\frac{n_h}{N_h} f(y \mid a; \boldsymbol{\theta})}{\sum_{j=1}^{H_{\text{obs}}} \frac{n_j}{N_j} Q_j(a; \boldsymbol{\theta}) + \sum_{j=H_{\text{obs}}+1}^{H_{\text{total}}} Q_j(a; \boldsymbol{\theta})}, \quad (10)$$

where $y \in S_h$, $\boldsymbol{\theta}$ is a vector of all the model parameters, LSAS indicates that the fish is in the second phase sample, $f(y \mid a; \boldsymbol{\theta})$ is the conditional distribution function of

length $Y = y$ given age $A = a$ for a random sample from the population (i.e. not length-stratified), N_h and n_h are respectively the first phase (i.e. length) and second phase (i.e. age) sample sizes in stratum S_h , and

$$Q_j(a; \boldsymbol{\theta}) = \Pr\{Y \in S_j \mid A = a; \boldsymbol{\theta}\}, \quad (11)$$

namely the conditional occupation probability for stratum S_j given age $A = a$.

2.2.1 Size-selective capture

Let C be a Bernoulli random variable indicating whether a fish is caught (1) or not (0) in a sample. The selectivity function $s(y)$ is the probability that a fish is caught given its length y ; that is

$$s(y) = \Pr(C = 1 \mid Y = y). \quad (12)$$

We assume the selectivity function follows a linear logistic model,

$$s(y) = \frac{1}{1 + \exp(-\log(19)[y - l_{50}]/[l_{95} - l_{50}])}, \quad (13)$$

where l_{50} and l_{95} are respectively the lengths at 50% and 95% selections, and constitute the model parameters to be estimated by the maximum likelihood approach.

The conditional distribution of the length of a captured fish (i.e. $C = 1$) is

$$f(y \mid C = 1, a; \boldsymbol{\theta}) = \frac{s(y) f(y \mid a; \boldsymbol{\theta})}{\Pr\{C = 1 \mid A = a; \boldsymbol{\theta}\}}, \quad (14)$$

which is derived in Appendix B.

2.2.2 Marginal likelihood for discrete length measurements

Let $\boldsymbol{\theta}$ be all the model parameters, and $\boldsymbol{\theta}_{d,g}$ be a subset of $\boldsymbol{\theta}$ including all the model parameters for spatial location d (e.g. NAFO Division for 3LNO American plaice) and sex g . Denote $\mathbf{R}_{c,d,g} = (L_{\infty,c,d,g}, K_{c,d,g})$, and define $S_{c,d,g}$ as the collection of all the second phase individuals of sex g in cohort c and division d .

The length data are usually discrete integer values in centimeters. We may account for this by discretizing the probability density function (pdf) used in the previous sections into the probability mass function (pmf) for discrete random variables as follows. For the i th individual in $S_{c,d,g}$, without considering the LSAS, the probability that its length y_i is an integer value l given $C_i = 1$, age a_i , random effects $\mathbf{R}_{c,d,g}$, and model parameters $\boldsymbol{\theta}_{d,g}$ is

$$\Pr\{y_i = l \mid C_i = 1, a_i, \mathbf{R}_{c,d,g}; \boldsymbol{\theta}_{d,g}\} = \frac{s(l) \text{LN}(l, \mu_{c,d,g}(a_i), \sigma_{c,d,g,a_i})}{\sum_{l'} s(l') \text{LN}(l', \mu_{c,d,g}(a_i), \sigma_{c,d,g,a_i})}, \quad (15)$$

where $\text{LN}(l, \mu, \sigma)$ denotes the log-normal density at l with mean μ and SD σ , $\mu_{c,d,g}(a_i)$ and σ_{c,d,g,a_i} are defined in (4) and (8) respectively, and the summation in the denominator is over all the possible integer length values. The denominator $\Pr\{C = 1 \mid A = a; \theta\}$ in (14) is canceled in the numerator and denominator of Eq. (15). The corresponding occupation probability Q_h for the h th stratum S_h conditional on age is

$$Q_h(C_i = 1, a_i, \mathbf{R}_{c,d,g}; \theta_{d,g}) = \sum_{l \in S_h} \Pr\{y_i = l \mid C_i = 1, a_i, \mathbf{R}_{c,d,g}; \theta_{d,g}\}. \quad (16)$$

For our American plaice case study the stratum size is 2 cm, so there are only two possible values of l in each S_h . The conditional EP probability for LSAS is then given by

$$\begin{aligned} & \Pr\{Y_i = l \mid C_i = 1, a_i, \mathbf{R}_{c,d,g}; \text{LSAS}, \theta_{d,g}\} \\ &= \frac{\frac{n_h}{N_h} s(l) \text{LN}(l, \mu_{c,d,g}(a_i), \sigma_{c,d,g,a_i})}{\sum_{j=1}^{H_{\text{obs}}} \frac{n_j}{N_j} Q_j(C_i = 1, a_i, \mathbf{R}_{c,d,g}; \theta_{d,g}) + \sum_{j=H_{\text{obs}}+1}^{H_{\text{total}}} Q_j(C_i = 1, a_i, \mathbf{R}_{c,d,g}; \theta_{d,g})} \\ &= \frac{\frac{n_h}{N_h} s(l) \text{LN}(l, \mu_{c,d,g}(a_i), \sigma_{c,d,g,a_i})}{\sum_{j=1}^{H_{\text{obs}}} \frac{n_j}{N_j} \sum_{l' \in S_j} s(l') \text{LN}(l', \mu_{c,d,g}(a_i), \sigma_{c,d,g,a_i}) + \sum_{j=H_{\text{obs}}+1}^{H_{\text{total}}} \sum_{l' \in S_j} s(l') \text{LN}(l', \mu_{c,d,g}(a_i), \sigma_{c,d,g,a_i})}. \end{aligned} \quad (17)$$

The cumbersome denominator summation in (15) cancels out in the EP formula (17), and this is the equation we use in our case study. The $\theta_{d,g}$ parameters and $\mathbf{R}_{c,d,g}$ random effects are all specific to the division and sex of the i th individual, and the random effects are also different for different cohorts. For the American plaice data of Sect. 3.1 with a stratum size of 2 cm, S_j are sets $\{2q, 2q + 1\}$ with $q = 0, 1, 2, \dots, H_{\text{total}}$. As defined in (10), $j = 1, \dots, H_{\text{obs}}$ enumerate the strata observed in the data; $j = H_{\text{obs}} + 1, \dots, H_{\text{total}}$ enumerate the possible but unobserved strata, which include the unobserved strata from length 0 to the length 15 cm greater than the longest observed length.

The EP likelihood function involving only the second phase length-at-age observations is

$$L_{\text{EP}}(\theta) = \int \left(\prod_g \prod_d \prod_c \prod_{i \in S_{c,d,g}} \Pr\{y_i \mid C_i = 1, a_i, \mathbf{R}_{c,d,g}; \text{LSAS}, \theta_{d,g}\} \right) g(\mathbf{L}, \mathbf{K} \mid \theta) d\mathbf{L} d\mathbf{K}, \quad (18)$$

where \mathbf{L} and \mathbf{K} are respectively the vectors of all the $L_{\infty,c,d,g}$'s and $K_{c,d,g}$'s, and $g(\mathbf{L}, \mathbf{K} \mid \theta)$ is the joint distribution function of \mathbf{L} and \mathbf{K} as defined by model (6)–(7).

We applied the Template Model Builder (TMB, Kristensen et al. 2015) package in R to calculate the marginal log-likelihood

$$l_{EP}(\boldsymbol{\theta}) = \log(L_{EP}(\boldsymbol{\theta})) \quad (19)$$

and its gradients. TMB uses Laplace approximation to evaluate the integration over random effects. The maximum EP likelihood estimates $\hat{\boldsymbol{\theta}}$ of the model parameters $\boldsymbol{\theta}$ are obtained by maximizing $l_{EP}(\boldsymbol{\theta})$ using the R (R Core Team 2018) function `nlminb()`. TMB can also make prediction for random effects.

The standard errors for parameter estimators and random effect predictors were obtained by using TMB R function “`sdreport()`”. The technical details for estimating these standard errors are given in Kristensen (2018). Note that the randomness in the empirical proportions, n_h/N_h , in (10) may introduce some extra variabilities in the finite sample maximum EP likelihood estimators.

2.2.3 Residuals

Equations (4)–(9) define the stochastic model for the distribution of length-at-age in the population. This would also be the distribution of length-at-age in a random sample from the population. However, the data we consider are not simple random samples, but involve size selectivity of the capture gear and 2 phase length-stratified sampling for ages. As a result, the distribution of the length-at-age data from the trawl survey sampling design does not follow model (4)–(9) even if this is exactly the true stochastic model for the population sampled. For this reason, the traditional random sample (RS) based error definition

$$\begin{aligned} e_i &= \log(y_i) - E(\log(Y_i)|a_i, \mathbf{R}_{c_i, d_i, g_i}; \text{RS}, \boldsymbol{\theta}_{d_i, g_i}) \\ &= \log(y_i) - \log(\mu_{c_i, d_i, g_i}(a_i)) + \log(\sigma_{c_i, d_i, g_i}^2 / \mu_{c_i, d_i, g_i}^2(a_i) + 1)/2 \end{aligned} \quad (20)$$

cannot reflect the model performance for LSAS data. Here c_i , d_i and g_i denote respectively the cohort, division and sex of the i th individual, and the log response is used due to log-normal model (9). The appropriate definition of the residual error is

$$\begin{aligned} e_i &= \log(y_i) - \chi_i, \\ \chi_i &= E(\log(Y_i)|C_i = 1, a_i, \mathbf{R}_{c_i, d_i, g_i}; \text{LSAS}, \boldsymbol{\theta}_{d_i, g_i}), \end{aligned} \quad (21)$$

where the expectation $E(\cdot)$ is with respect to the distribution of the second phase data after size selectivity and the length-stratified sampling. This distribution is given by (17). The residuals, which are the estimated errors, can be calculated as

$$\begin{aligned} r_i &= \log(y_i) - \hat{\chi}_i, \\ \hat{\chi}_i &= \sum_l \log(l) \times \Pr \left\{ Y_i = l \mid C_i = 1, a_i, \hat{\mathbf{R}}_{c_i, d_i, g_i}; \text{LSAS}, \hat{\boldsymbol{\theta}}_{d_i, g_i} \right\}, \end{aligned} \quad (22)$$

where hat denotes that the estimated parameters and the predicted random effects are used. The standard errors (SEs) of the residuals can be calculated with

Table 1 Numbers of male and female American plaice measured (N_{Div}), those with both length and age measured (n_{Div}), and the number of sets during the DFO spring and fall surveys from 1978 to 2015

Gender	Quantity	3L	3N	3O	All divisions
Male	N_{Div}	493,139	216,474	204,627	914,240
	n_{Div}	22,942	16,432	17,200	56,574
	Number of sets	4011	2505	2494	9010
Female	N_{Div}	478,955	231,682	229,129	939,766
	n_{Div}	32,410	23,631	24,291	80,332
	Number of sets	4376	2699	2692	9767

$$\text{SE}(r_i) = \sqrt{\sum_l \log^2(l) \times \Pr\{Y_i = l \mid C_i = 1, a_i, \hat{\mathbf{R}}_{c_i, d_i, g_i}; \text{LSAS}, \hat{\boldsymbol{\theta}}_{d_i, g_i}\} - \hat{\chi}_i^2}. \quad (23)$$

The standardized residual for the i th observation can then be evaluated as $r_i/\text{SE}(r_i)$. In (22) and (23) we used (17) to estimate respectively the means and standard errors for the LSAS residuals.

3 Results

In this section we apply the estimation approach proposed in Sect. 2.2 to fitting the spatiotemporal growth model proposed in Sect. 2.1 using data collected by DFO in NAFO Divisions 3L, 3N and 3O during their spring and fall research surveys from 1978 to 2015. In 2006 and 2015 the spring surveys were not completed. Fall surveys were not conducted in 3N and 3O before year 1990, and in 3L before 1981.

3.1 Data

During a spring or fall survey, within each division, a series of length-stratified samples, which are called “sets”, were collected throughout the region to spread the sampling across the division. Each set involved measuring the length of all fish caught in a research trawl tow, classifying them into 2 cm length strata, and subsampling a few or no otoliths from each length stratum. The sampling goal for all sets in each division and survey season was to obtain about 25 age measurements per 2cm length stratum by sex if length ≥ 10 cm, and about 15 age measurements per stratum without sex distinguishment if length < 10 cm. For the entire time-series almost 2 million fish have been measured for length and 147,000 fish have been aged (Table 1).

Table 2 illustrates the count data from one set. The length-stratified samples were collected in different spatial positions. They also had some subtle and complicated correlation because various sets needed to cooperate to ensure that the second phase samples in each length stratum were spread out across each division. However, in model (4)–(9), all the individuals in the same division are assumed to be identically

Table 2 Counts of female American plaice measured for length (row N_h) and those with both length and age measured (row n_h) in 2 cm length strata in a set within Division 3L collected during the DFO 2014 spring survey

Length bin	12.5	14.5	16.5	18.5	20.5	22.5	24.5	26.5	28.5	30.5
N_h	1	3	7	8	13	6	3	1	3	3
n_h	0	0	0	0	1	1	0	0	0	1
Length bin	32.5	34.5	36.5	38.5	40.5	42.5	44.5	46.5	48.5	
N_h	2	4	4	2	2	8	3	2	1	
n_h	0	1	1	1	0	2	2	1	0	

Midpoints of length bins are indicated above the counts

distributed; that is, under this model, within a division, the spatial separation of the sets can be ignored, and the individuals are identical or interchangeable. In this case Perreault et al. (2019) showed that one can treat all the sets in the same division and year as a single length-stratified sample. To be specific, when applying the EP likelihood approach, for any year, the empirical proportion for the h th stratum of division d and sex g is

$$\frac{\sum_{i=1}^{m_s} n_{h,g,i}}{\sum_{i=1}^{m_s} N_{h,g,i}},$$

where $i = 1, \dots, m_s$ enumerates all the sets in division d in that year, $N_{h,g,i}$ is the first phase sample size of stratum h and sex g in set i , and $n_{h,g,i}$ is the second phase sample size of stratum h and sex g in set i .

Survey gears used to collect the data in this study did not catch larval fish. This makes it difficult to give a realistic estimation of a_0 in (1) or l_0 in (4). For example, by fitting the models (1) and (4) to the data, the a_0 estimate is positive in Division 3L and l_0 estimates are negative for the three divisions. The simulation studies in Zheng and Cadigan (2019) and Perreault et al. (2019) also indicated that a_0 and l_0 cannot be estimated well with the existing approaches. A remedy for this issue is to borrow data from some other studies. Here we used the data in Region (1995) that American plaice larva is 0.62–0.75 cm long 5 days after hatching when the yolk absorption is complete. This can be regarded as age 0 length, namely $l_0 = (0.62 + 0.75)/2 = 0.685$ cm. In the following estimation we fix l_0 at 0.685.

The spring surveys were conducted from March to July, which are the spawning months of American plaice. The fall surveys were conducted from August to December, which on average are 4–5 months later than the spawning months. Hence for fall surveys we assume the age = measured age + adjustment. In this study we use the maximum likelihood approach to estimate this age adjustment for fall surveys.

A Yankee trawl was used for the surveys prior to 1983, which was changed to an Engel trawl in 1983. Since the fall of 1995, the survey gear has been the Campelen trawl. DFO conducted experiments to determine the conversion between Yankee and Engel (Gavaris and Brodie 1984) and between Engel and Campelen (Warren et al. 1997).

Gavaris and Brodie (1984) concluded that for American plaice < 28 cm the ratio of the Engel catches to Yankee catches was 0.5 with a standard error (SE) of 0.05, and for plaice ≥ 28 cm, the ratio was 1.3 with SE 0.09. Figure 8 in Warren et al. (1997) included the data for the ratio of Campelen catches to Engel catches of American plaice for lengths between 25 and 40 cm. We fit these data using nonlinear least squares, and estimated that the coefficient of variation (CV) around the fitted curve was 0.0968. This was based on assuming the ratio of catches was Gamma distributed with means at the fitted curve. Let s_Y , s_E and s_C denote the selectivities defined in (13) for the Yankee, Engel and Campelen trawls respectively, and $\text{dgamma}\{\text{data}|\text{parameters}\}$ denote the probability density function of a gamma distribution. To enhance the estimation of selectivities, we regarded these sampled ratios provided by Gavaris and Brodie (1984) and Warren et al. (1997) as data independent of the American plaice data, and added the following terms to the log-likelihood (19),

$$\begin{aligned} & \sum_{l=10}^{27} \log(\text{dgamma}\{0.5|\text{mean} = s_E(l)/s_Y(l), \text{variance} = 0.05^2\}) \\ & + \sum_{l=28}^{60} \log(\text{dgamma}\{1.3|\text{mean} = s_E(l)/s_Y(l), \text{variance} = 0.09^2\}) \\ & + \sum_{l=25}^{40} \log(\text{dgamma}\{r_{CE}(l)|\text{mean} = s_C(l)/s_E(l), \text{CV} = 0.0968\}), \end{aligned}$$

where $r_{CE}(l)$ is the observed selectivity ratio of Campelen to Engel at length l (Warren et al. 1997).

3.2 Estimation results

We estimated θ by maximizing the log EP likelihood (19). Parameter estimates and the corresponding standard errors (SEs) are given in Tables 3, 4, 5. The estimates of the selectivities of the survey gears using the method in Sect. 2.2.1 are provided in Table 6 with the corresponding SEs, and plotted in Fig. 1.

The autocorrelation parameters $\varphi_{\infty,d,g}$'s and $\varphi_{k,d,g}$'s are all larger than 0.89, and especially the autocorrelations of 3L males are extremely high with $\varphi_{\infty,3L,\text{male}} = 0.98$ and $\varphi_{k,3L,\text{male}} = 0.99$, indicating strong cohort or temporal correlation among the asymptotic lengths $L_{\infty,c,d,g}$'s and among the growth rates $K_{c,d,g}$'s of close cohorts and the same sex within the same division.

The estimated means of $L_{\infty,c,d,g}$ ($= \exp(E[\log(L_{\infty,c,d,g})] + \text{Var}[\log(L_{\infty,c,d,g})]/2)$) $= \exp(\mu_{\infty,d,g} + \delta_{\infty,d,g}^2/\{2[1 - \varphi_{\infty,d,g}^2]\})$ and of $K_{c,d,g}$ ($= \exp(E[\log(K_{c,d,g})] + \text{Var}[\log(K_{c,d,g})]/2)$) $= \exp(\mu_{k,d,g} + \delta_{k,d,g}^2/\{2[1 - \varphi_{k,d,g}^2]\})$ are provided in Table 7. These are values of L_{∞} and K averaged over all cohorts. The L_{∞} 's and K 's have significant cohort variations, as indicated by $\delta_{\infty,d,g}$'s and $\delta_{k,d,g}$'s that are significantly greater than zero. The larger $\delta_{\infty,d,g}$'s and $\delta_{k,d,g}$'s in 3L females suggest that 3L females have higher cohort variation in L_{∞} and K than 3N and 3O females. The predicted

Table 3 The EP likelihood estimation of the parameters in Model (4)–(9) for the 3LNO male (M) and female (F) American plaice data collected in DFO spring and fall surveys from 1978 to 2015

Sex	Div.		μ_∞	φ_∞	δ_∞	μ_k	φ_k	δ_k	ν	λ	Δ_∞	a_{50}	a_{95}	l_{95}	τ
M	3L	Est.	4.01	0.98	0.055	-0.92	0.99	0.14	3.44	1.06	115.78	5.85	27.09	20.36	1.18
		SE	0.22	0.022	0.0071	0.79	0.012	0.015	0.025	0.032	291.72	1.05	3.15	0.84	0.072
	3N	Est.	4.13	0.93	0.045	-1.87	0.89	0.15	3.11	1.22	0.70	12.86	22.68	21.39	0.59
		SE	0.11	0.028	0.0055	0.17	0.033	0.015	0.032	0.020	0.17	1.11	2.58	3.61	0.16
	3O	Est.	3.91	0.93	0.062	-1.80	0.92	0.15	3.45	1.13	0.45	9.01	15.55	9.40	3.70
		SE	0.11	0.028	0.0096	0.24	0.026	0.017	0.019	0.023	0.11	0.25	1.06	0.71	0.75
F	3L	Est.	4.21	0.93	0.097	-2.56	0.94	0.15	3.47	1.20	0.15	7.53	10.38	19.35	1.13
		SE	0.15	0.018	0.011	0.26	0.015	0.017	0.016	0.013	0.019	0.085	0.29	0.56	0.052
	3N	Est.	4.22	0.94	0.033	-1.78	0.89	0.12	3.41	1.04	1.35	8.06	23.52	18.55	1.07
		SE	0.064	0.029	0.0047	0.15	0.029	0.013	0.020	0.021	0.35	0.69	0.88	1.58	0.20
	3O	Est.	4.20	0.94	0.068	-3.27	0.91	0.12	3.77	1.40	-0.63	1.63	5.59	13.65	1.96
		SE	0.13	0.030	0.0082	0.18	0.028	0.013	0.019	0.032	0.073	0.53	0.24	0.81	0.27

“Div.”, “Est.” and “SE” are respectively the abbreviations for “Division”, “estimate” and “standard error”. The CV’s are virtually zero and hence not shown

Table 4 The estimated correlation coefficients $\rho_{\infty, dg, d'g'}$

		Female			Male		
		3L	3N	3O	3L	3N	3O
Female	3L	1.00	0.67	0.68	0.68	0.34	0.75
	3N	0.082	1.00	0.90	0.26	0.78	0.80
	3O	0.079	0.034	1.00	0.20	0.62	0.85
Male	3L	0.083	0.14	0.14	1.00	0.43	0.61
	3N	0.11	0.055	0.087	0.14	1.00	0.76
	3O	0.068	0.083	0.054	0.11	0.075	1.00

The estimates and the corresponding standard errors are given in the upper and lower triangles respectively

Table 5 The estimated correlation coefficients $\rho_{k, dg, d'g'}$

		Female			Male		
		3L	3N	3O	3L	3N	3O
Female	3L	1.00	0.58	0.65	0.79	0.38	0.72
	3N	0.10	1.00	0.91	0.39	0.92	0.86
	3O	0.072	0.029	1.00	0.25	0.73	0.89
Male	3L	0.042	0.11	0.081	1.00	0.40	0.53
	3N	0.14	0.028	0.072	0.13	1.00	0.80
	3O	0.083	0.053	0.031	0.085	0.065	1.00

The estimates and the corresponding standard errors are given in the upper and lower triangles respectively

Table 6 The estimated selectivities of the survey gears Yankee, Engel and Campelen

	Yankee		Engel		Campelen	
	l_{50}	l_{95}	l_{50}	l_{95}	l_{50}	l_{95}
Est.	34.55	67.73	31.28	48.38	11.01	12.65
SE	0.51	1.96	0.26	0.57	0.049	0.075

“Est.” and “SE” are respectively the abbreviations for “estimate” and “standard error”

cohort-average growth curves (Fig. 2) suggest that on average (i.e. for all cohorts) female American plaice in 3L are smaller than the same age females in 3N and 3O, which is also true for male American plaice less than about 16 years old. On average male American plaice in 3L, 3N and 3O are smaller than females of the same age and division after ages 5.6, 6.1 and 1.0 respectively. However the maximum differences for 3L and 3N are only 2.9 cm and 2.6 cm respectively up to age 24.5, while the maximum difference for 3O is 13.0 cm.

The predicted $L_{\infty, c, d, g}$'s (Fig. 3) and $K_{c, d, g}$'s (Fig. 4) show greater local variability if the corresponding autocorrelation parameter φ 's are smaller (e.g. L_{∞} and K of 3NO males, and K of 3N females), and show greater overall variability if the corresponding marginal variances ($\delta^2/(1 - \varphi^2)$) are larger (e.g. L_{∞} and K of 3L males and females). In Table 5, for males and females separately, the correlation coefficients between 3N

Fig. 1 Plot of selectivities for Yankee, Engel and Campelen

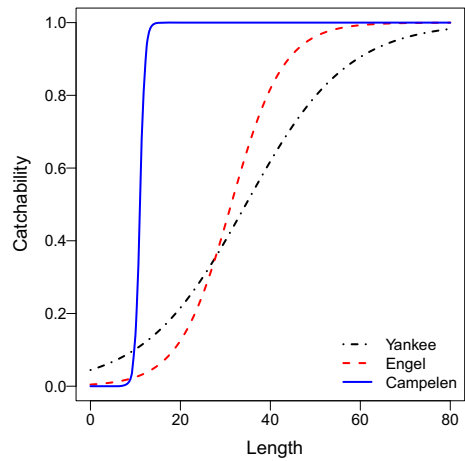


Table 7 The estimated means of $L_{\infty,c,d,g}$ and $K_{c,d,g}$ averaged over all cohorts

		3L	3N	3O
L_{∞}	Female	69.34 (10.02)	68.16 (4.23)	67.97 (8.53)
	Male	57.15 (12.97)	62.53 (6.75)	50.78 (5.52)
K	Female	0.086 (0.022)	0.17 (0.026)	0.040 (0.0072)
	Male	0.60 (0.61)	0.16 (0.028)	0.18 (0.043)

The estimates (Est.) and the corresponding standard errors (SE) are given as “Est. (SE)”

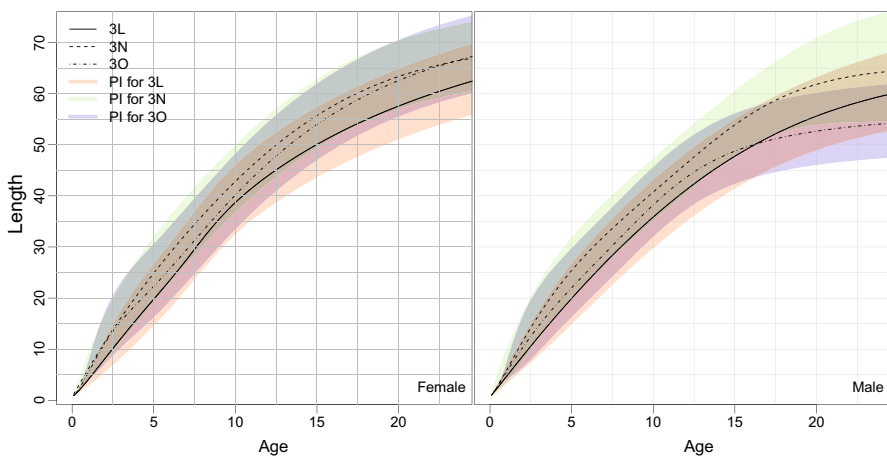


Fig. 2 Cohort-average predicted growth curves and the 95% prediction intervals (PI) with predicted $L_{\infty,c,d,g}$'s and $K_{c,d,g}$'s and other growth parameters in model (4)–(9) for males and females and the NAFO Divisions 3L, 3N and 3O

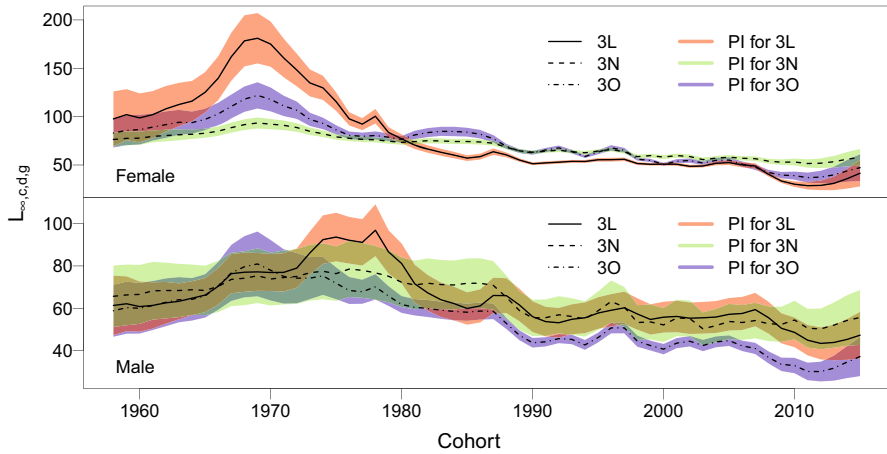


Fig. 3 Predicted male and female cohort asymptotic lengths $L_{\infty,c,d,g}$'s with 95% prediction intervals (PI)

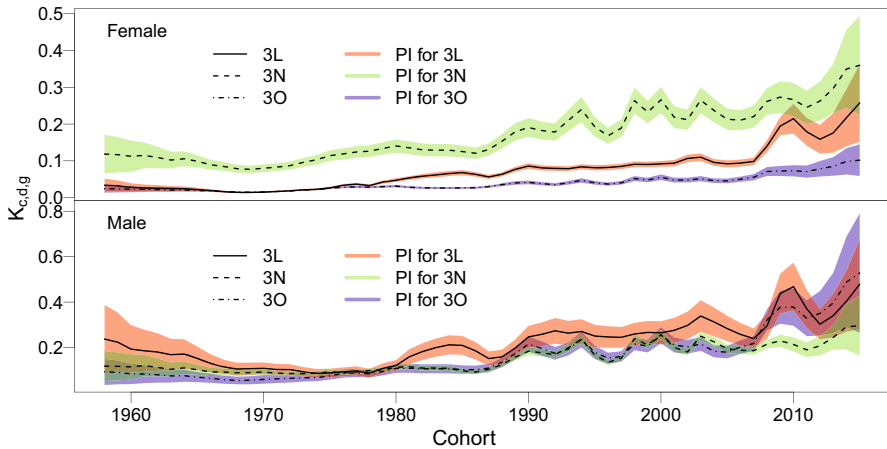


Fig. 4 Predicted male and female cohort growth rates $K_{c,d,g}$'s with 95% prediction intervals (PI)

and 3O are much larger than those between 3L and 3NO, and the SEs of the former are smaller than those of the latter, indicating that the correlations in growth rate between 3N and 3O are much larger than those between 3L and 3N and between 3L and 3O, which is also evident in Fig. 4. Similarly, Table 4 and Fig. 3 also indicate that for males and females separately, the correlations in asymptotic length between 3N and 3O are substantially larger than those between 3L and 3N and between 3L and 3O.

Tables 4 and 5 and Figs. 3 and 4 indicate that males and females in the same division have substantial correlation in both asymptotic length and growth rate. As a result, due to the high correlation of the American plaice of the same sex between 3N and 3O, the correlation between males and females separately in 3N and 3O can also be substantial (e.g. 3N female and 3O male).

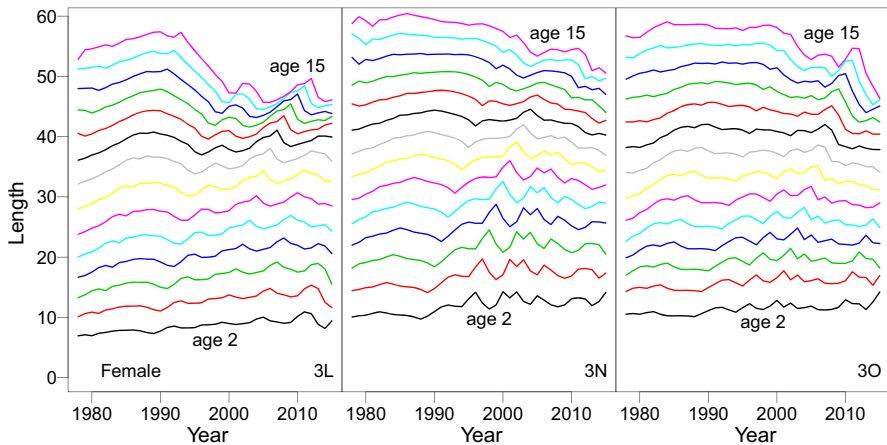


Fig. 5 Estimated female population mean length at age for each year and division

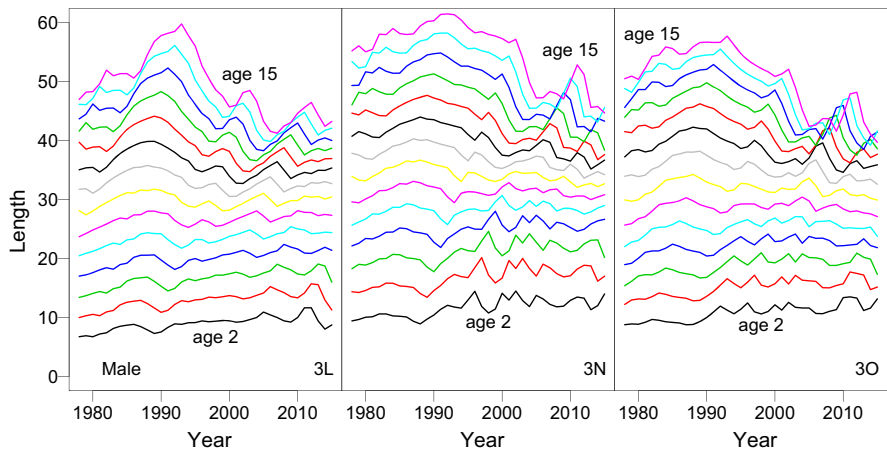


Fig. 6 Estimated male population mean length at age for each year and division

Referring to (8) and Table 3, the linear-in-length components (CVs) of the BI variations are all negligible, and the asymptotic constant parts (ν 's) are about 3.5 cm for both sexes and the 3 divisions. The mean lengths to achieve 95% of ν 's are 24.70 (0.53), 23.34 (1.06) and 19.79 (0.71) respectively for 3L, 3N and 3O females, and 26.25 (0.76), 22.68 (2.10) and 15.39 (0.66) respectively for 3L, 3N and 3O males. The numbers in brackets are standard errors.

Estimated length-at-age time-series (Figs. 5 and 6) suggests that the males and females in 3L are in general smaller than the fish of the same sex in 3N, and also smaller than fish of the same sex in 3O at older ages.

The estimated population probability density function of length-at-age for a series of cohorts of male American plaice in Division 3L are shown in Fig. 7. These densities do not involve the selectivity and LSAS effects, and reflect the population distribution

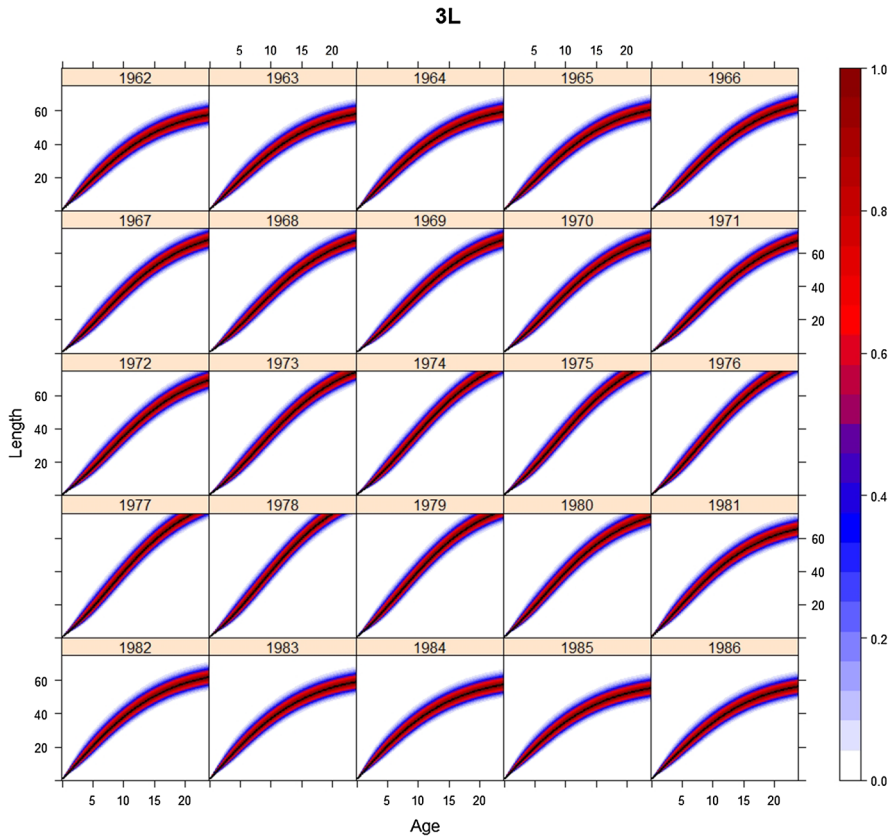


Fig. 7 3L male population length-at-age probability density function by cohort, part 1. The black curves are the corresponding fitted population growth curves, and the number above each panel gives the cohort. Within each panel, at each age the densities are normalized by the maximum density at that age. The density plots of other cohorts and divisions and for females are given in the Supplementary figures

of length at age. The density plots of other cohorts, divisions and sexes are given in Figs. S1–S11 in the Supplementary figures. Fig. 7 and S1–S11 show not only the mean curves, but also the BI variations as defined by (8), namely, the variation in length of the fish at the same age in the same cohort and division.

The transitions between the juvenile and adult growth phases of females and males in the three divisions are illustrated in Fig. 8 and by the a_{50} 's and a_{95} 's in Table 3.

The model-estimated mean length-at-age time series of caught fish (solid curves) and the corresponding model-free generalized ratio estimates (dashed curves. See e.g., p. 236, Hájek 1971; Morgan and Hoenig 1997) are shown in Figs. 9, 10, 11, 12. For LSAS, the generalized ratio estimator for mean length at age of caught fish is

$$\bar{y}_{i \in S_{a,d,g,y}} = \frac{\sum_{i \in S_{a,d,g,y}} w_i y_i}{\sum_{i \in S_{a,d,g,y}} w_i}, \quad (24)$$

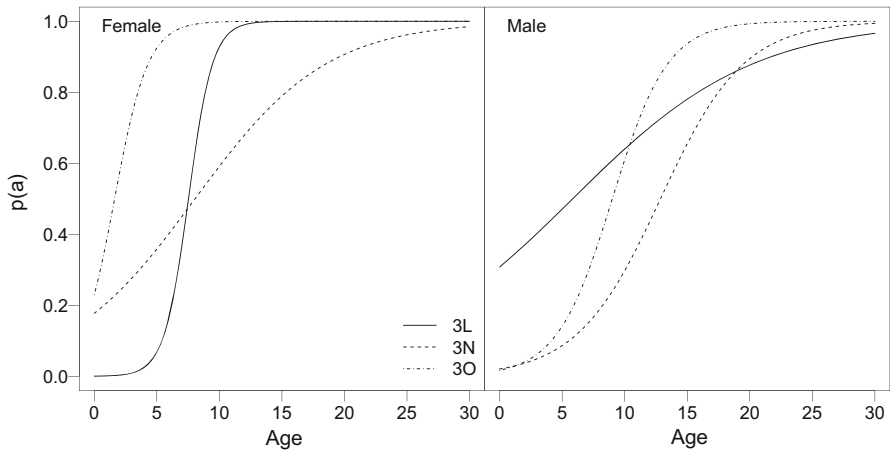


Fig. 8 Plot of the biphasic transition factor $p(a)$ defined by (5) for female and male in the three Divisions 3L, 3N and 3O

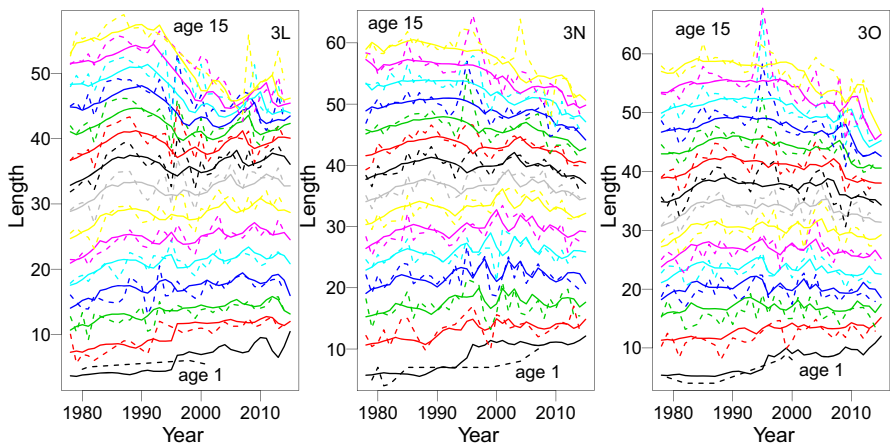


Fig. 9 Estimated mean length at age of caught female American plaice for each year and division (solid curves). The dashed lines are the corresponding generalized ratio estimates of the mean length at age of female American plaice caught in spring surveys

where $S_{a,d,g,y}$ is the collection of all the second phase individuals of sex g with age a in division d and year y , and $w_i = N_h/n_h$ with N_h and n_h being respectively the first phase and second phase sample sizes of the h th stratum to which the i th individual belongs. Due to the age adjustment to fall surveys as explained in Sect. 3.1, we plotted spring means and fall means separately. The estimated adjustment for fall age is 0.42 for male with SE 0.0082 and 0.47 for female with SE 0.0072. The model-based estimates overlap the data-based model-free estimates, indicating that the model fits the data. The sharp jumps at ages 1 and 2 in some figures were due to the change of survey gears from Engle to Campelen in the fall of 1995.

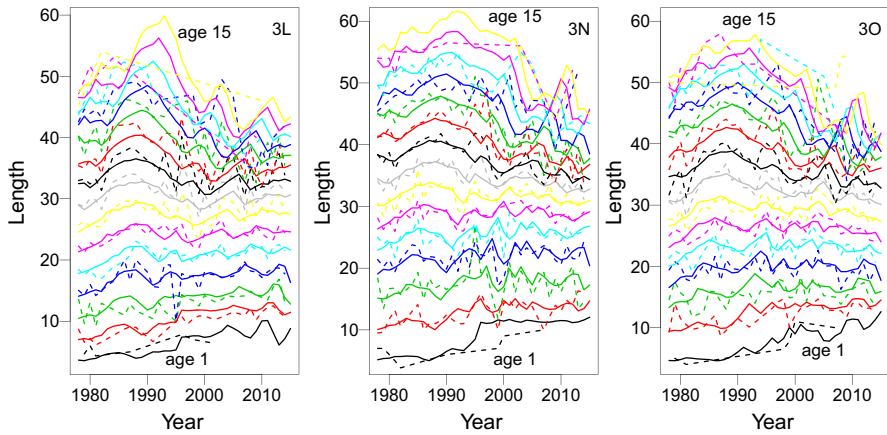


Fig. 10 Estimated mean length at age of caught male American plaice for each year and division (solid curves). The dashed lines are the corresponding generalized ratio estimates of the mean length at age of male American plaice caught in spring surveys

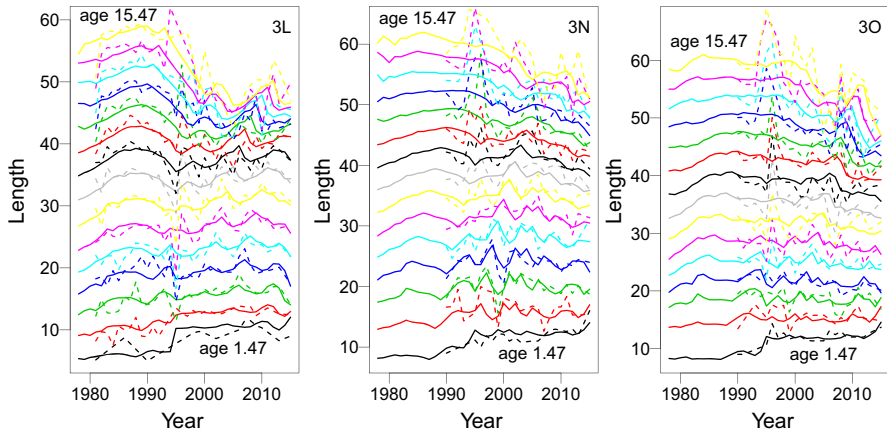


Fig. 11 Estimated mean length at age of caught female American plaice for each year and division (solid curves). The dashed lines are the corresponding generalized ratio estimates of the mean length at age of female American plaice caught in fall surveys

Plots of standardized residuals versus year or age for each division and sex (Figs. S12–S15 in the Supplementary figures) do not indicate lack of model fit overall. At older ages (> 17 for female and > 12 for male), the residual means have higher between-age variability because there are fewer observations at these ages. Residual plots by year and age (Figs. S16–S21 in the Supplementary figures) do not indicate lack of fit either.

The growth rate relative to the asymptotic length, namely the relative growth rate, is defined as

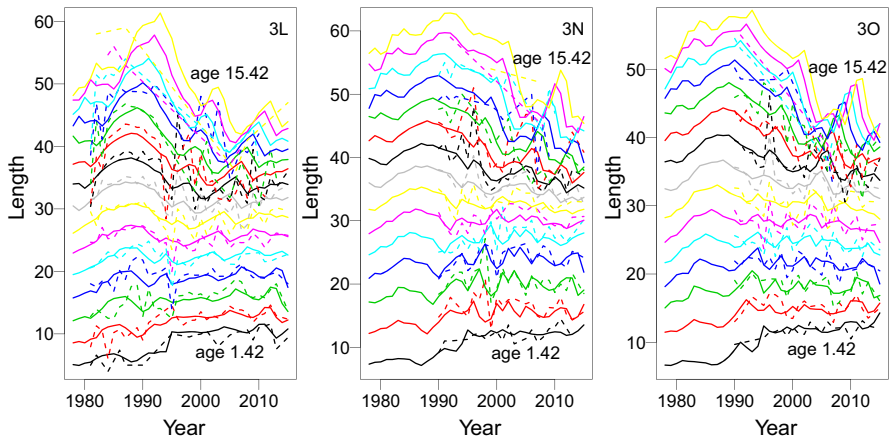


Fig. 12 Estimated mean length at age of caught male American plaice for each year and division (solid curves). The dashed lines are the corresponding generalized ratio estimates of the mean length at age of male American plaice caught in fall surveys

$$V_{c,d,g}(a) = \frac{1}{L_{\infty,c,d,g}} \frac{d\mu_{c,d,g}(a)}{da}, \quad (25)$$

where $\mu_{c,d,g}(a)$ is given by (4). The estimated relative growth rates by cohort for each division and sex are provided in Figs. S22 and S23 in Supplementary figures. The divisional relative growth rates are combined with the divisional abundances at each cohort, age and sex in Fig. S24 of Supplementary figures.

3.3 Simulation study

In this section we investigate the performance of the proposed maximum likelihood estimation approach for the spatiotemporal RS log- L_{∞} growth model (4)–(9) for the length stratified age sampling (LSAS) and when there is research trawl selectivity. To simplify the problem, we consider only the female American plaice in Division 3L during 1996 to 2014 when only one type of research gear (Campelen) was used. To generate similar age samples as the data, we assume that age follows a negative binomial distribution with success probability 0.5 and number of failures equal to 1. Since age 0 can never enter the final sample due to the gear selectivity, we remove age 0 by left-shifting the age distribution, namely, age 1 assumes the probability for age 0, age 2 assumes the probability for age 1, and so on. For each year, we generate 8000 age data following the shifted negative binomial distribution, and to ensure some data of older ages, we increase the count number at each age from 1 to 24 by 4. Corresponding to each age, the length data can be randomly generated with model (4), (5), (8) and (9) as well as the estimated parameters in Table 3 and the predicted $L_{\infty,c,3L,female}$ and $K_{c,3L,female}$. This is the population length at age. The population is randomly selected by the estimated Campelen selectivity given in Table 6 to obtain the first phase sample. LSAS is applied to the first phase sample to get around 430 second phase length at

Table 8 The RelBias and RRMSE for the parameter estimates in the simulation study

	ν	λ	a_{50}	a_{95}	l_{95}	τ	$l_{50, \text{Campelen}}$	$l_{95, \text{Campelen}}$
RelBias	− 0.06	− 0.03	− 0.04	0.02	− 0.11	0.17	0.00	0.01
RRMSE	0.07	0.04	0.06	0.07	0.15	0.26	0.01	0.01

age data for each year, which is then fitted to the model (4), (5), (6), (8) and (9) using the method presented in Sect. 2.2.

This data generation and parameter estimation were repeated 500 times. The estimation performance was measured using the relative root mean squared error (RRMSE),

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{500} (\text{estimate}_i - \text{true})^2}{500}}, \quad \text{RRMSE} = \frac{\text{RMSE}}{|\text{true}|} \cdot 100,$$

and relative bias,

$$\text{Bias} = \frac{\sum_{i=1}^{500} (\text{estimate}_i - \text{true})}{500}, \quad \text{RelBias} = \frac{\text{Bias}}{|\text{true}|} \cdot 100.$$

Here “true” denotes the true parameter value.

The RelBias and RRMSE for the parameter estimates are given in Table 8. Because $L_{\infty, c, 3L, \text{female}}$ and $K_{c, 3L, \text{female}}$ are fixed at the predicted values instead of being generated randomly with (6), the parameters related to the dynamics of L_{∞} and K such as φ_{∞} and δ_{∞} are not relevant and hence not included in Table 8. Instead the boxplots for the mean lengths at age are presented in Figs. S29 and S30 in the Supplementary figures where the red lines are the true values.

Table 8 indicates that the estimating approach proposed in Sect. 2.2 gives good estimation of the model parameters with fairly small bias and MSE. Figs. S29 and S30 illustrate that the mean length at age can also be predicted well. The predictions for the mean length at age 1 may be inferior to those at other ages because there are only a tiny amount of age 1 data due to the gear selectivity.

4 Discussion

Spatiotemporal models are increasingly used in fisheries research to account for spatiotemporal correlation and variation in demographic and dynamic processes over time, space, and among individuals. Meanwhile, stratified sampling designs are widely applied in large-scale sampling programs for reducing sampling expenses but maintaining the population representativeness. Nevertheless, how to make reliable statistical inference for models with spatiotemporal correlations with data collected from length-stratified age sampling (LSAS) has not been sufficiently studied in the literature.

We assumed that individuals (i.e. fish) of the same sex in the same cohort within a division share some similarity in growth, and there are correlations in asymptotic length and growth rate between the temporally neighboring cohorts and the correlations decay with temporal distance between cohorts. The AR(1) type model can represent all these assumptions, and we used this approach to link asymptotic lengths and growth rates using (6). In addition to this temporal aspect of the model, the deviations of the AR(1) series are assumed to follow a multivariate normal distribution with between-division and sex correlations, which represent the spatial component of the model. Furthermore, we assume this AR(1) model with multivariate normal deviations to be stationary so that the marginal distributions and marginal correlations of $\log(L_{\infty,c,d,g})$ and $\log(K_{c,d,g})$, especially $\log(L_{\infty,1,d,g})$ and $\log(K_{1,d,g})$, can be evaluated with no need of introducing extra parameters. The within-cohort between-individual (BI) variations are then modeled through an additive random error to the log mean length at age whose standard deviation represents the BI variation as a function of the mean length. For the LSAS design, we applied the empirical proportion (EP) likelihood approach proposed in Zheng and Cadigan (2019), and we addressed the issue of lacking effective models for the population age distribution by conditioning on age.

One advantage of this spatiotemporal model is to systematically use data collected for many surveys to improve estimation of growth model parameters. For example, length data at young and old ages are important in determining the growth model parameters. For the vonB model (1), lengths at older ages are mostly related to l_{∞} , while lengths at very young ages reflect a_0 . However, when analyzing many years of growth data it is common for some years or cohorts to have a relatively narrow range of ages and in this case the data may be uninformative about all three vonB parameters; that is, the vonB model may be poorly identified. The methods proposed in this paper address this issue by linking the data from many years through autoregressive models connecting asymptotic lengths ($L_{\infty,c,d,g}$) and growth rates ($K_{c,d,g}$) of sequential cohorts, and connecting data of different divisions and sexes through divisional and sex correlations of L_{∞} and K . With this approach, information about length-at-age from all years are accumulated together to improve parameter estimation, and information about L_{∞} and K from all the cohorts, divisions and sexes, especially from the close cohorts, is “borrowed” for cohorts where the data are not informative enough to estimate these parameters and random effects reliably on their own. This model also reveals the spatiotemporal structure of the length-at-age data.

The selectivity parameters b_0 and b_1 in (13), or equivalently l_{50} and l_{95} , for research survey gears are not the primary interest in this study, and the length-at-age data used in this work may not provide precise estimation of these parameters. However, they are an important model component that is necessary to account for the bias in the data introduced by survey selectivity because (i) at younger ages the survey gears tend to catch fast-growers so that the growth curve inferred from the caught fish is different from the true population growth curve, and (ii) three kinds of survey gears with quite different selectivities (see Fig. 1; Gavaris and Brodie 1984; Morgan et al. 1998) have been used in the surveys we examined, and ignoring selectivity amounts to ignoring their differences in selectivity, which can lead to biased model estimation. For point (i), in Figs. S25 and S26 in the Supplementary figures we plotted the estimated population

lengths at age in the fall with the corresponding generalized ratio means of the caught fish (Eq. (24)), which clearly indicate the difference between population lengths and the corresponding lengths of the caught fish, especially ages 1 and 2. Figs. S25 and S26 used the same scales as Figs. 11 and 12 respectively for comparison. If research gear selectivities are neglected, the residuals by age are provided in Figs. S27 and S28 in the Supplementary figures, which indicate substantial underestimation of the mean length at age for caught fish at ages 1 and 2. This underestimation continues somewhat up to age 8 since the l_{50} 's of the Yankee and Engel survey trawls are about 30 cm. For point (ii), many authors (e.g. Gavaris and Brodie 1984; Warren et al. 1997; Stansbury 1997; Morgan et al. 1998) have assessed the different impacts of these survey gears on survey catch rates and their length and age distributions, and tried to find conversion factors among data collected by different survey gears to account for these impacts. Direct implementation of the relevant conversion factors for 3LNO American plaice developed in these studies led to unsatisfactory estimation results. However, when we used these previous studies as prior information for the estimation of selectivities as explained in Sect. 3.1, the selectivities estimated according to Sect. 2.2.1 gave improved estimation of the model parameters.

When applying the vonB model (1) and the Richards–Schnute growth models (Richards 1959; Schnute 1981), the length at age zero, a_0 , was unrealistically positive in some divisions. The problem was identified as a lack of data at younger ages and solved by fixing l_0 , the length at age zero, with data in Region (1995). However, the model fit to the data became much worse primarily due to insufficient number of parameters (only 2, L_∞ and K) to capture the biphasic (juvenile and adult) growth pattern of the large scale data over many years and a wide spatial region. Manabe et al. (2018) surmised that at least 5 parameters are necessary to depict biphasic growth curves. In this work, we combined the VB log- L_∞ model (Minte-Vera et al. 2016) and a simplified version of the Schnute-Richards growth model (Schnute and Richards 1990; Cadigan and Campana 2017) as a 6-parameter RS log- L_∞ growth model, which gives a good fit to the data.

The VB log- L_∞ model was constructed by assuming a smooth transition between the juvenile and adult L_∞ 's without clearly defining the factor $p(a)$ in (5) (Minte-Vera et al. 2016), so it is understandable that $p(a)$ may not exactly reflect the proportion of mature, as can be seen from Fig. 8 and the estimates of a_{50} 's and a_{95} 's in Table 3, which do not exactly agree with the maturity timing of American plaice (Pitt 1966; Region 1995; Morgan and Colbourne 1999). The real data analysis of Minte-Vera et al. (2016) indicates that when assuming that $p(a)$ is maturity-at-age function and using both size-at-age and maturity-at-age data there was a convergence issue. In this study estimating age at maturity was not the aim, but rather this was only included to get better estimates of growth. Minte-Vera et al. (2016) suggests that the more sophisticated growth models derived from the biological first principles and bioenergetics considerations such as the Quince-Boukal model (Quince et al. 2008) and the CoR model (Minte-Vera et al. 2016) can achieve better fits, but require maturity data. We will investigate these alternatives in future research. Nevertheless, these models are substantially more difficult to implement than the VB log- L_∞ model because they are recursive models without closed-form solutions, and the convergence of fitting the Quince-Boukal model is heavily influenced by the correlation among model param-

eters (Minte-Vera et al. 2016), so overall the VB log- L_{∞} model used in this study is still a practically good choice.

BI variation model is especially important when there is gear selectivity. For example, the selectivity of the Campelen trawl has a sharp jump around 11 cm, which is close to the lengths of American plaice at ages 1, 2 and 3. If the growth model is correct but the proposed BI variations at these ages are larger than the true values then the estimated mean length of caught fish will be larger than the actual mean of the caught fish at these ages because the research gear tends to select larger fish. Similarly, if the assumed BI variations are smaller than the true values then the estimated mean lengths of caught fish at these ages will be smaller than the true means. These fitting issues due to an incorrect BI variation model can show up in residual plots and mean length plots for caught fish. In Figs. S14 and S15 in the Supplementary figures, the residual means at ages 1, 2 and 3 are quite close to zero, and in Figs. 9, 10, 11, 12, the estimated mean lengths of caught fish overlap the sample means, indicating the effectiveness of the BI variation model (8). In addition, if at a certain age the modeled BI variation is smaller or greater than the true value, the SD (standard deviation) of the standardized residuals will be greater or smaller than 1 accordingly at this age. In Figs. S14 and S15 the 1-SD lines mostly overlap the reference lines at 1, which provides further validation for the BI variation model (8). Because the estimated CVs are negligible, for almost all the cases where the estimated $\tau > 1$ (except only 3N male), model (8) essentially says that the SD of the BI variation increases first convexly and then concavely with the mean length, and finally converges to a constant. Model (8) is not theoretically derived, but an empirically optimal model selected from many candidate models.

In Figs. S15 and S19–S21 in the Supplementary figures the residuals of male American plaice at older ages around 16 may suggest some larger variability. This can be explained by the small sample sizes at these older ages. For example, There are only 3, 5 and 4 sampled lengths at age at age 16 in 3L, 3N and 3O respectively. However, for the purpose of stock assessment the estimated lengths at ages averaged over male and female are actually used. The proportions of males older than 15 are smaller than 0.014 in 3L, 0.0046 in 3N, and 0.019 in 3O. Therefore, any fitting problem for males at older ages due to lack of data does not influence the application of the model estimates in stock assessment.

The generalized ratio estimators (Eq. (24)) have been widely used in practice to estimate various quantities from data collected through LSAS (see, e.g. Morgan and Hoenig 1997). However, the model-based approach presented in this paper has the following advantages: (i) When the sample size is small, the generalized ratio estimators are biased and may have large standard errors. On the contrary, the model-based approach uses all the data to estimate the model parameters and give smooth estimation and prediction for all the categories despite their sample sizes. (ii) The generalized ratio estimators may not be precisely calculated without a model. For example, Eq. (24) is the generalized ratio estimator for mean length-at-age of caught fish, while the more useful generalized ratio estimator for mean length-at-age of the whole population is

$$\bar{y}_{i \in S_{a,c,d,g}} = \frac{\sum_{i \in S_{a,c,d,g}} \{w_i / s(y_i)\} y_i}{\sum_{i \in S_{a,c,d,g}} w_i / s(y_i)}, \quad (26)$$

where $s(y)$ is the selectivity function defined in (12), which has to be estimated by a model-based approach. (iii) The model-based approach can make predictions and estimate prediction errors for those categories with no data (e.g. year 2006 in Figs. 9 and 10 since there was no spring survey in year 2006), which is not possible with the generalized ratio estimators. (iv) Model parameter estimates can reveal the relationship between responses and explanatory covariates.

Our results indicate an overall decline in asymptotic length L_∞ and sizes at older ages in each of Divisions 3L, 3N and 3O. Growth in fish is affected by many, sometimes interacting, factors. Food availability and temperature are two of the major influences on growth. Metabolic processes are directly affected by temperature, and for each species there is an optimum temperature and food level that results in maximum growth. As food availability declines, the most efficient growth occurs at lower temperatures where requirements for maintenance are less because of a lower metabolic rate (Brett 1979). Temperatures over most of the period examined here have been warm (Colbourne et al. 2016), which could promote growth, but the level of food availability is unknown and may have been limiting. Increased energy allocated to reproduction can result in decreased growth. Division 3LNO American plaice have exhibited a steep decline in the age at which they mature (Morgan and Colbourne 1999; Wheeland et al. 2018) and this may have contributed to the declines in length at age shown in this study. It is also possible that size selective fishing has played a role (Swain et al. 2007; Enberg et al. 2012), although fishing mortality has declined in recent years (Wheeland et al. 2018).

It is reasonable to assume that the maturity of a fish is related to the ratio of its length l to its asymptotic length l_∞ , namely l/l_∞ ; that is, a fish growing to 60% of its asymptotic length is more likely to be mature than a fish growing to 50% of its asymptotic length. If this is true, higher relative growth rates defined by Eq. (25) at younger ages suggest a younger age at 50% maturity, a_{50} . In Fig. 17 of Wheeland et al. (2018), the a_{50} 's of male and female 3LNO American plaice declined steeply around the 1990 cohort, then remained at a substantially lower level compared to the cohorts before 1990. In Fig. S24 of Supplementary figures, the relative growth rates of both males and females at younger ages rose sharply to a higher level near the 1990 cohort, implying a sharp decrease in a_{50} , then remained at the higher level for the cohorts after 1990, implying a lower level of a_{50} . Other details in Fig. S24 also mostly agree with Fig. 17 in Wheeland et al. (2018).

Our residual analyses indicated some small year effects. This may suggest a further improvement to the model. In addition to the cohort effect $K_{c,d,g}$, another year effect $K_{y,d,g}$ could be introduced into the model; that is, a cohort could have a time-varying K parameter (or effect) due to some environmental factors that affect multiple contemporary cohorts. For simplicity, here we neglect the cost of production. The vonB model would be

$$y(a) = L_{\infty,c,d,g} \left\{ 1 - (1 - l_0/L_{\infty,c,d,g}) \exp \left(- \int_0^a k(t) dt \right) \right\}, \quad (27)$$

where l_0 is the age 0 length, and

$$\int_0^a k(t)dt = K_{c,d,g} a + \sum_{y=c}^{c+a-1} K_{y,d,g}.$$

$K_{y,d,g}$ of sequential years can be linked by another AR(1) model similar to (6). In this work we did not pursue this because these year effects could also be related to additional within-division spatial variation in growth rates combined with non-random spatial distribution of samples. We plan to extend the spatial dimension of the model in the near future and this may provide further insights about the small residual year-effects. Such further developments of the model will be the subject of future research.

Growth data consist of measured lengths and ages, which may have some measurement errors (MEs). In this work, because the measurements of lengths and ages were confirmed to be accurate, we did not involve these MEs in our model. If they are nonnegligible, they can be treated as follows. Fish lengths are usually taken using a measuring board or tape, and length ME's may be about the same regardless of the size of the fish; hence, an additive ME model is reasonable. In temperate areas, the age of a fish is often determined by counting growth rings on small ear bones (otoliths) using a microscope. As fish get older the rings may be more difficult to count. This suggests that ME in age increases with age. A multiplicative model can capture this characteristic. Hence, we can use additive ME in length, and multiplicative ME in age,

$$\begin{aligned} Y_i &= Y_i^T + \varepsilon_i, \quad \varepsilon_i \stackrel{i.i.d.}{\sim} N(0, \sigma_\varepsilon^2) \\ A_i &= A_i^T e^{\xi_i}, \quad \xi_i \stackrel{i.i.d.}{\sim} N(0, \sigma_\xi^2), \end{aligned} \quad (28)$$

where Y_i and A_i are respectively the measured length and age for the i th fish with corresponding MEs ε_i and ξ_i , and Y_i^T and A_i^T are their true but unknown values. “i.i.d.” here denotes “independent and identically distributed”. Length measurements can be reliably obtained by trained specialists, so their ME variance can be assumed small and known. Age determination is considerably more expensive and error prone than length measurement. Y_i^T can be modeled as a RS log- L_∞ function of A_i^T (Eq. (4)).

In conclusion, an accurate growth curve modeling and estimation should take into account the sampling strategy, survey gear selectivity, biphasic growth due to reproduction, as well as spatiotemporal and BI variations in growth. We have presented a comprehensive methodology to address all these issues, and demonstrated its good performance through scrutinizing the estimation results.

Acknowledgements We thank all of the people involved in the collection and processing of these data. Research funding to NZ and NC was provided by the Ocean Frontier Institute, through an award from the Canada First Research Excellence Fund. Research funding to NC was also provided by the Ocean Choice International Industry Research Chair program at the Fisheries and Marine Institute of Memorial University of Newfoundland. Many thanks for the comments from the two anonymous reviewers that greatly improved this manuscript.

Appendices

Appendix A: Stationary results

If the AR(1) models in (6) are stationary, the marginal distributions of $\log(L_{\infty,c,d,g})$ and $\log(K_{c,d,g})$ are

$$\begin{aligned}\log(L_{\infty,c,d,g}) &\sim N(\mu_{\infty,d,g}, \omega_{\infty,d,g}^2), \\ \log(K_{c,d,g}) &\sim N(\mu_{k,d,g}, \omega_{k,d,g}^2)\end{aligned}\quad (29)$$

with $\omega_{\infty,d,g}^2 = \delta_{\infty,d,g}^2 / (1 - \varphi_{\infty,d,g}^2)$ and $\omega_{k,d,g}^2 = \delta_{k,d,g}^2 / (1 - \varphi_{k,d,g}^2)$, and the distributions of $\log(L_{\infty,1,d,g})$ and $\log(K_{1,d,g})$ are given by (29).

Based on the stationary assumption, the marginal sex and division correlations are

$$\begin{aligned}\text{corr}(\log(L_{\infty,c,d,g}), \log(L_{\infty,c,d',g'})) &= \frac{\sqrt{1 - \varphi_{\infty,d,g}^2} \sqrt{1 - \varphi_{\infty,d',g'}^2}}{1 - \varphi_{\infty,d,g} \varphi_{\infty,d',g'}} \rho_{\infty,dg,d'g'}, \\ \text{corr}(\log(K_{c,d,g}), \log(K_{c,d',g'})) &= \frac{\sqrt{1 - \varphi_{k,d,g}^2} \sqrt{1 - \varphi_{k,d',g'}^2}}{1 - \varphi_{k,d,g} \varphi_{k,d',g'}} \rho_{k,dg,d'g'}.\end{aligned}\quad (30)$$

Appendix B: Conditional distribution of a captured fish

The conditional distribution of a captured fish (i.e. $C = 1$) is

$$\begin{aligned}f(y | C = 1, a; \theta) &= \frac{\Pr\{Y = y, C = 1 | A = a; \theta\}}{\Pr\{C = 1 | A = a; \theta\}} \\ &= \frac{\Pr\{C = 1 | Y = y, A = a; \theta\} \Pr\{Y = y | A = a; \theta\}}{\Pr\{C = 1 | A = a; \theta\}} \\ &= \frac{s(y) f(y | a; \theta)}{\Pr\{C = 1 | A = a; \theta\}}\end{aligned}$$

In the last step we applied $\Pr\{C = 1 | Y = y, A = a; \theta\} = \Pr\{C = 1 | Y = y\}$; that is, the probability of capture does not depend on age given the size of a fish.

References

- Brett J (1979) Environmental factors and growth. Fish physiology, vol VIII. Bioenergetics and growth. Academic Press, Orlando, pp 599–677
- Brody S (1945) Bioenergetics and growth. Reinhold Publishing Corporation, New York
- Cadigan NG, Campana SE (2017) Hierarchical model-based estimation of population growth curves for redfish (*sebastes mentella* and *sebastes fasciatus*) off the eastern coast of canada. ICES J Mar Sci 74(3):687–697
- Candy SG (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

- Candy SG, Constable AJ, Lamb T, 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 Sci* 14:43–66
- Colbourne E, Holden J, Senciall D, Bailey W, Snook S (2016) Physical oceanographic environment on the newfoundland and labrador shelf in nafo subareas 2 and 3 during 2015. NAFO Scientific Council Research Document 16(07)
- Day T, Taylor PD (1997) Von bertalanffy's growth equation should not be used to model age and size at maturity. *Am Nat* 149(2):381–393
- Enberg K, Jørgensen C, Dunlop ES, Varpe Ø, Boukal DS, Baulier L, Eliassen S, Heino M (2012) Fishing-induced evolution of growth: concepts, mechanisms and the empirical evidence. *Mar Ecol* 33(1):1–25
- Filipe PA, Braumann CA, Roquete CJ (2012) Multiphasic individual growth models in random environments. *Methodol Comput Appl Probab* 14(1):49–56
- Gavaris S, Brodie W (1984) Results of comparative fishing between the at cameron and the wilfred templeman during july-august 1983. *CAFSAC Res Doc* 84(41):16
- Goodyear CP (1995) Mean size at age: an evaluation of sampling strategies with simulated red grouper data. *Trans Am Fish Soc* 124(5):746–755
- Goodyear CP (2019) Modeling growth: consequences from selecting samples by size. *Trans Am Fish Soc* 148:528–551
- Hájek J (1971) Discussion of 'an essay on the logical foundations of survey sampling, part one' by d. basu. Foundations of statistical inference Edited by VP Godambe and DA Sprott Holt, Rinehart, Winston, Toronto, Ont, pp 236
- Kristensen K (2018) Rdocumentation: sdreport. <https://www.rdocumentation.org/packages/TMB/versions/1.7.14/topics/sdreport>
- Kristensen K, Nielsen A, Berg CW, Skaug H, Bell B (2015) Tmb: automatic differentiation and laplace approximation. arXiv preprint [arXiv:1509.00660](https://arxiv.org/abs/1509.00660)
- Laslett GM, Eveson JP, Polacheck T (2002) A flexible maximum likelihood approach for fitting growth curves to tag recapture data. *Can J Fish Aquat Sci* 59(6):976–986
- Lester N, Shuter B, Abrams P (2004) Interpreting the von bertalanffy model of somatic growth in fishes: the cost of reproduction. *Proc R Soc Lond B Biol Sci* 271(1548):1625–1631
- Lester NP, Shuter BJ, Venturelli P, Nadeau D (2014) Life-history plasticity and sustainable exploitation: a theory of growth compensation applied to walleye management. *Ecol Appl* 24(1):38–54
- Manabe A, Yamakawa T, Ohnishi S, Akamine T, Narimatsu Y, Tanaka H, Funamoto T, Ueda Y, Yamamoto T (2018) A novel growth function incorporating the effects of reproductive energy allocation. *PLoS ONE* 13(6):e0199346
- Mills TC, Mills TC (1991) Time series techniques for economists. Cambridge University Press, Cambridge
- Minte-Vera CV, Maunder MN, Casselman JM, Campana SE (2016) Growth functions that incorporate the cost of reproduction. *Fish Res* 180:31–44
- Morgan MJ, Colbourne E (1999) Variation in maturity-at-age and size in three populations of american plaice. *ICES J Mar Sci* 56(5):673–688
- Morgan MJ, Hoenig JM (1997) Estimating maturity-at-age from length stratified sampling. *J Northwest Atl Fish Sci* 21:51–64
- Morgan MJ, Brodie W, Bowering W, Parsons DM, Orr D, (1998) Results of data conversions for american plaice in div. 3 lno from comparative fishing trials between the engel otter trawl and the campelen, (1800) shrimp trawl. *Sci Coun Res Doc NAFO* 98:10
- Morin R, LeBlanc SG, Campana SE (2013) Bomb radiocarbon validates age and long-term growth declines in american plaice in the southern gulf of st. lawrence. *Trans Am Fish Soc* 142(2):458–470
- Perreault AM, Zheng N, Cadigan NG (2019) Estimation of growth parameters based on length-stratified age samples. *Can J Fish Aquat Sci* 999:1–12
- Piner KR, Lee HH, Maunder MN (2016) Evaluation of using random-at-length observations and an equilibrium approximation of the population age structure in fitting the von bertalanffy growth function. *Fish Res* 180:128–137
- Pitt T (1966) Sexual maturity and spawning of the american plaice, *hippoglossoides platessoides* (fabricius), from newfoundland and grand bank areas. *J Fish Board Canada* 23(5):651–672
- Prajneshu Venugopalan R (1999) von bertalanffy growth model in a random environment. *Can J Fish Aquat Sci* 56(6):1026–1030
- Quince C, Abrams PA, Shuter BJ, Lester NP (2008) Biphasic growth in fish i: theoretical foundations. *J Theor Biol* 254(2):197–206

- Quinn TJ, Deriso RB (1999) Quantitative fish dynamics. Oxford University Press, Oxford
- Quist MC, Pegg MA, DeVries DR (2012) Age and growth. Fisheries techniques, 3rd edn. American Fisheries Society, Bethesda, pp 677–731
- R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL <https://www.R-project.org/>
- Region N (1995) American plaice, hippoglossoides platessoides, life history and habitat characteristics. In: Proceedings of the symposium on the potential for development of aquaculture in Massachusetts, vol 15, pp 17
- Richards F (1959) A flexible growth function for empirical use. *J Exp Bot* 10(2):290–301
- Roff DA (1983) An allocation model of growth and reproduction in fish. *Can J Fish Aquat Sci* 40(9):1395–1404
- Roff DA et al (2000) Trade-offs between growth and reproduction: an analysis of the quantitative genetic evidence. *J Evol Biol* 13(3):434–445
- Russo T, Baldi P, Parisi A, Magnifico G, Mariani S, Cataudella S (2009) Lévy processes and stochastic von bertalanffy models of growth, with application to fish population analysis. *J Theor Biol* 258(4):521–529
- Sainsbury K (1980) Effect of individual variability on the von bertalanffy growth equation. *Can J Fish Aquat Sci* 37(2):241–247
- Schnute J (1981) A versatile growth model with statistically stable parameters. *Can J Fish Aquat Sci* 38(9):1128–1140
- Schnute JT, Richards LJ (1990) A unified approach to the analysis of fish growth, maturity, and survivorship data. *Can J Fish Aquat Sci* 47(1):24–40
- Schueller AM, Williams EH, Cheshire RT (2014) A proposed, tested, and applied adjustment to account for bias in growth parameter estimates due to selectivity. *Fish Res* 158:26–39
- Shelton AO, Satterthwaite WH, Beakes MP, Munch SB, Sogard SM, Mangel M (2013) Separating intrinsic and environmental contributions to growth and their population consequences. *Am Nat* 181(6):799–814
- Soriano M, Moreau J, Hoenig JM, Pauly D (1992) New functions for the analysis of two-phase growth of juvenile and adult fishes, with application to nile perch. *Trans Am Fish Soc* 121(4):486–493
- Stansbury D (1997) Conversion factors for cod from comparative fishing trials for engel 145 otter trawl and the campelen 1800 shrimp trawl used on research vessels. NAFO SCR Doc 97:73
- Swain DP, Sinclair AF, Mark Hanson J (2007) Evolutionary response to size-selective mortality in an exploited fish population. *Proc R Soc B Biol Sci* 274(1613):1015–1022
- Von Bertalanffy L (1960) Principles and theory of growth. In: Nowinski W (ed) Fundamental aspects of normal and malignant growth. Elsevier, Amsterdam, pp 137–259
- Walsh SJ (1997) Efficiency of bottom sampling trawls in deriving survey abundance indices. *Oceanogr Lit Rev* 44(7):748–748
- Wang YG, Thomas MR, Somers IF (1995) A maximum likelihood approach for estimating growth from tag-recapture data. *Can J Fish Aquat Sci* 52(2):252–259
- Warren W, Brodie W, Stansbury D, Walsh S, Morgan J, Orr D (1997) Analysis of the 1996 comparative fishing trial between the alfred needler with the engel 145 trawl and the wilfred templeman with the campelen 1800 trawl. NAFO SCR Doc 97:68
- Webber DN, Thorson JT (2016) Variation in growth among individuals and over time: a case study and simulation experiment involving tagged antarctic toothfish. *Fish Res* 180:67–76
- Wheeland L, Dwyer K, Morgan J, Rideout R, Rogers R (2018) Assessment of american plaice in div. 3lno. NAFO SCR Doc 18(039)
- Zheng N, Cadigan N (2019) Likelihood methods for basic stratified sampling, with application to von bertalanffy growth model estimation. *Open J Stat* 9(6):623–642

Nan Zheng joined CFER as an OFI Post-Doctoral Fellow in September 2017. His work is primarily focusing on developing spatial stock assessment models for Grand Banks flatfish. Nan Zheng earned his MSc. in 2013 and PhD in 2017 from the Department of Mathematics and Statistics at MUN. His MSc. work was in financial time series, and his PhD work was in semi-parametric fixed and mixed models for longitudinal discrete data. He was awarded the Dean's Award for 2013–17 as well as the titles of Fellow of the School of Graduate Studies 2013–14 and 2017–18 for his MSc. and PhD studies respectively. Nan Zheng

received his BSc. degree from Zhejiang University in China. His previous work was in general relativity, solid state physics, and optics.

Noel Cadigan is a quantitative fisheries scientist specializing in statistical methods for fish stock assessment and sustainable fisheries. He has extensive experience in the assessment of Newfoundland fish stocks, and experience with other Canadian, American and European stocks. He received a doctor of philosophy in statistics in 1999 at the University of Waterloo, a master of applied statistics in 1993 and a bachelor of science (Honours) in 1990 from Memorial University. He worked with Fisheries and Oceans Canada at the Northwest Atlantic Fisheries Center in Newfoundland from 1990–2011 (with a 2-year break to pursue his studies) and focused on the assessment of Atlantic cod and other groundfish stocks. Dr. Cadigan is a cross-appointed member in the Department of Mathematics and Statistics at Memorial University and has supervised graduate students in statistics whose research involved applications to stock assessment problems.

M. Joanne Morgan Current research and/or projects: My research focuses on factors affecting productivity in groundfish species. Specifically I am attempting to understand the factors that affect growth, condition, maturation and fecundity. These are all critical to determining the level of fishing a population can sustain and in determining the length of time it will take depleted populations to recover. There are also important implications for interactions among different components of the ecosystem. In addition, I am trying to determine how to quantitatively integrate this information into scientific advice to fisheries management. As an important part of this aspect of my work I am actively involved in the stock assessment of a variety of groundfish populations, both nationally and internationally through the Northwest Atlantic Fisheries Organization (NAFO). Education and awards: PhD Behavioural Ecology, Queen's University at Kingston, 1987. Key publications: 1. Morgan, M.J. et al. 2009. The evaluation of reference points and stock productivity in the context of alternative indices of stock reproductive potential. *Can. J. Fish. Aquat. Sci.* 66:404–414 2. Morgan, M.J. 2008. Integrating reproductive biology into scientific advice for fisheries management. *J. Northw. Atl. Fish. Sci.* 41:37–51. 3. Stares, J.C., R.M. Rideout, M.J. Morgan and J. Bratney. 2007. Did population collapse influence individual fecundity of Northwest Atlantic cod? *ICES J. Mar. Sci.* 64:1338–1347. 4. Morgan, M.J., P.A. Shelton and J. Bratney. 2007. Age composition of the spawning stock does not always influence recruitment. *J. Northw. Atl. Fish. Sci.* 38: 1–12. 5. Rideout, R.M. and M.J. Morgan. 2007. Major changes in fecundity and the effect on population egg production for three species of Northwest Atlantic flatfish. *J. Fish Biol.* 70:1759–1779.

Affiliations

Nan Zheng¹  · Noel Cadigan¹ · M. Joanne Morgan²

Noel Cadigan
Noel.Cadigan@mi.mun.ca

M. Joanne Morgan
Joanne.Morgan@dfo-mpo.gc.ca

¹ Centre for Fisheries Ecosystems Research, Fisheries and Marine Institute of Memorial University of Newfoundland, St. John's, NL, Canada

² Fisheries and Oceans Canada, St. John's, NL, Canada