ICES Journal of Marine Science

ICES
CIEM

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models

Luke Batts [1,2,*], Cóilín Minto[1], Hans Gerritsen[2], and Deirdre Brophy[1]

[1]Marine and Freshwater Research Centre, Galway-Mayo Institute of Technology, Galway, H91 T8NW, Ireland
[2]Marine Institute, Rinville, Oranmore, Galway, H91 R673, Ireland

*Corresponding author: tel: +353 91 742525; e-mail: luke.batts@research.gmit.ie.

Analysis of length frequency distributions from surveys is one well-known method for obtaining growth parameter estimates where direct age estimates are not available. We present a likelihood-based procedure that uses mixture models and the expectation–maximization algorithm to estimate growth parameters from length frequency data (LFEM). A basic LFEM model estimates a single set of growth parameters that produce one set of component means and standard deviations that best fits length frequency distributions over all years and surveys. The hierarchical extension incorporates bivariate random effects into the model. A hierarchical framework enables inter-annual or inter-cohort variation in some of the growth parameters to be modelled, thereby accommodating some of the natural variation that occurs in fish growth. Testing on two fish species, haddock (*Melanogrammus aeglefinus*) and white-bellied anglerfish (*Lophius piscatorius*), we were able to obtain reasonable estimates of growth parameters, as well as successfully model growth variability. Estimated growth parameters showed some sensitivity to the starting values and occasionally failed to converge on biologically realistic values. This was dealt with through model selection and was partly addressed by the addition of the hierarchical extension.

Keywords: anglerfish *Lophius piscatorius*, bivariate random effects, EM algorithm, haddock *Melanogrammus aeglefinus*, LFEM, von Bertalanffy growth

Introduction

Fish growth is a widely studied aspect of fish biology (Pardo et al., 2013), reflecting its importance for understanding life histories (Denney et al., 2002), the effects of changing environmental conditions (Jobling, 2002; Baudron et al., 2014) as well as ecosystem and population dynamics (Quinn and Deriso, 1999; Travers et al., 2007; Smith et al., 2015). Growth parameters are a fundamental component of fisheries stock assessments (Quinn and Deriso, 1999) and their accurate estimation can be key to ensuring that assessments are reliable (Hilborn and Walters, 1992). Estimation of other important population dynamics parameters such as fishing mortality and selectivity depend on the accurate estimation of growth (Hoggarth, 2006; Maunder et al., 2016).

For teleost fish, growth information is usually obtained by estimating age from incremental growth layers laid down within the otoliths (fish ear bones), scales or fin rays. When age estimation using calcified structures is too costly to implement, not practical or is suspected to give unreliable estimates then other methods such as tagging or length frequency analysis provide an alternative means of obtaining growth estimates (Hilborn and Walters, 1992). Clear, unambiguous growth increments are often lacking in the calcified structures of tropical or deep sea species that are not exposed to marked seasonal temperature changes, while direct ageing of crustaceans using hard parts still requires further validation (Kilada and Driscoll, 2017). As a consequence, assessments of fisheries for these species often rely on alternative methods for estimating age (Sparre and Venema, 1998; Chang et al., 2012).

The most widely used fish growth model is the von Bertalanffy growth function (VBGF), which relates length to age using three parameters; L_∞ (asymptotic mean length), K (rate of approach to asymptote), and t_0 (the theoretical time at which fish length

equals zero) (von Bertalanffy, 1938). Various modifications of the traditional VBGF appear in the literature. For example, Schnute and Fournier (1980) re-parameterized the VBGF to include more biologically intuitive parameters (i.e. the mean lengths of the first and last ages of the sampled population and a parameter that controls growth between ages). Another example is the incorporation of seasonal oscillation into the VBGF (Pitcher and MacDonald, 1973; Lloyd-Jones *et al.*, 2016). Other growth functions that have been used to model fish growth include the Gompertz (Winsor, 1932), allometric (Quinn and Deriso, 1999), logistic (Cormon *et al.*, 2016), and bi-phasic models (Minte-Vera *et al.*, 2016); the most appropriate choice of model depends on the species (Katsanevakis and Maravelias, 2008). Here, the re-parameterized VBGF was chosen as the underlying model because of the widespread use of the VBGF and its appropriateness for length frequency analysis (Schnute and Fournier, 1980).

Growth can often vary spatially, temporally, between cohorts, between individuals, and within individuals (Quinn and Deriso, 1999; Wang, 1999; Morrongiello and Thresher, 2015). Cohort effects can vary to a lesser or greater extent depending on the species and have been linked to environmental changes (Baudron *et al.*, 2014; Morrongiello and Thresher, 2015). Accounting for cohort growth variability can also improve stock assessment models (Whitten *et al.*, 2013). Methods for accounting for individual growth variation in length-at-age data have also been explored, for example Pilling *et al.* (2002) used a non-linear random effects model to model growth parameter variability between individuals length-at-age data. Cadigan and Campana (2016) developed a hierarchical mixed-effects model to account for between stock and between individual variability in length-at-age data. Growth models incorporating random effects have also been applied to analysis of tagging data, where between-individual growth variability is often considerable (Eveson *et al.*, 2015).

In situations where fish cannot be aged directly, length frequency analysis offers an alternative means of estimating fish growth parameters, provided there is a distinct spawning season (Hasselblad, 1966; Bhattacharya, 1967; Fournier *et al.*, 1990; Taylor and Mildenberger, 2017). Perhaps the most prominent length frequency method in the literature is the robust maximum likelihood-based estimation procedure MULTIFAN, which is now incorporated into a length-based statistical catch at age stock assessment model (Fournier *et al.*, 1990, 1998). The non-parametric ELEFAN procedure (Pauly, 1987) is also widely used and has recently been extended into the R environment where modern optimization algorithms are used to improve its performance (Mildenberger *et al.*, 2017; Taylor and Mildenberger, 2017).

Currently available approaches to length frequency analysis present some limitations. Both MULTIFAN and ELEFAN require that length bins are defined prior to the analysis, potentially introducing subjectivity to the identification of cohorts. For ELEFAN it is recommended that L_{∞} be fixed in the initial analysis and is then re-calculated post-optimization. Both models are also largely constrained to a single growth curve and do not incorporate inter-annual or inter-cohort variability in growth (although this is addressed to some extent in the development of MULTIFAN-CL) (Fournier *et al.*, 1998). Current limitations could be addressed by developing a model that uses an expectation–maximization (EM) algorithm to estimate growth parameters (Dempster *et al.*, 1977). This approach would remove the

need to predefine length classes and offers the flexibility to include variability in growth.

Whilst the EM algorithm is widely used to model length frequency distributions, it has rarely been used to estimate mixture model parameters with an incorporated von Bertalanffy growth structure. However, the methodology of MULTIFAN is similar to the EM algorithm in its application of maximum likelihood theory (Fournier *et al.*, 1990). The "mixdist" package in the R environment enables the user to fit a mixture model with component means constrained to a growth curve and this utilizes an EM algorithm (Macdonald and Du, 2011). In addition, a recent paper by Lloyd-Jones *et al.* (2016) used a minorization–maximization (MM) algorithm (the EM algorithm is a form of MM algorithm) to estimate seasonally oscillating von Bertalanffy growth over a year on monthly length frequency data of a crab species (Lloyd-Jones *et al.*, 2016), where L_{∞} is considered a random variable, although it is not modelled explicitly. These examples leave much scope for expanding this area of research further.

This study aims to develop a new method, based on the EM algorithm, to model fish growth from length frequency data whilst incorporating von Bertalanffy growth structure and accounting for variability in growth between cohorts or years. Gaussian mixture models are developed and fit to length frequency distributions, where means of components within a cohort (represented by normal distributions in the mixture models) are restricted to von Bertalanffy growth. First, a basic version (analogous to MULTIFAN) of the methodology is introduced and tested using length frequency data for two fish species; haddock *Melanogrammus aeglefinus* and white-bellied anglerfish *Lophius piscatorius*. Second, hierarchical models that introduce various bivariate random effects, are introduced and tested with the intention of modelling some of the variation in the recruitment and growth of the fish species. The overall aim of the study is to introduce an alternative approach to obtaining a range of credible estimates of growth parameters from length frequency data and account for variability in growth by cohort or year.

Methods
A description of the growth function used within the model is first introduced, followed by a description of the length frequency expectation-maximization (LFEM) model itself. A summary of model parameters, inputs, and outputs is then given, as well as a description of the model selection criteria and software used for development. Finally, the two data sources that the LFEM model is applied to are described.

Growth function
Length frequency distributions were modelled using mixture models, where the means of the cohorts (normal distributions within the mixture models) follow a VBGF. Schnute and Fournier (1980) re-parameterized the classical von Bertalanffy growth (CVBG) equation into one more appropriate to length frequency analysis, represented by

$$\mu_i = l + (L - l)\frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \ldots, M, \quad (1)$$

where μ_i is the mean length of fish at age a_i; l and L are the first and final component or cohort mean lengths, i.e. μ_1 and μ_M, where M is the maximum age class; k is equivalent to $\exp(-K)$

and controls the distance between two successive component means [Equation (1)]. The CVBG parameters can then be calculated from these parameters using

$$L_\infty = \frac{L - lk^{M-1}}{1 - k^{M-1}}, \tag{2}$$

$$K = -\ln(k), \tag{3}$$

$$t_0 = a_1 - \frac{1}{\ln(k)}\ln\left\{\frac{L-l}{L-lk^{M-1}}\right\}, \tag{4}$$

where L_∞ is the asymptotic length of the fish, K controls the rate at which the asymptote is approached, and t_0 is the theoretical time when an individual would have length zero. a_1 is the assumed age of the l component. Standard deviation (*SD*) of the normal distribution of the component (σ_i) is either estimated as constant (CSD) or a linear (LSD) function of component means (Schnute and Fournier, 1980). The equation for LSD is

$$\sigma_i = s + (S - s)\frac{\mu_i - l}{L - l}; \quad i = 1, \ldots, M, \tag{5}$$

where s is the *SD* of the first component and S is the *SD* of the final component.

LFEM model and algorithm

This methodology summarizes the basic model developed in this study where a single set of growth parameters that produce one set of component means and *SD*s are estimated (i.e. if l, L, and k are identical for every cohort then component means in each annual mixture for a particular survey model are identical).

Based on the observed data log-likelihood for a single mixture model (McLachlan and Peel, 2004), the observed data log-likelihood for a set of finite mixture models on length frequency distributions of fish where cohort component means are restricted to the re-parameterized VBGF (RVBG) is

$$\ln L(\Psi) = \sum_{v=1}^{V}\sum_{r=1}^{R}\sum_{j=1}^{n_{vr}}\ln\left\{\sum_{i=1}^{M}\lambda_{vri}f_{vri}(y_{jvr}; \theta_{vri}, a_{1v})\right\}, \tag{6}$$

$$f_{vri}(y_{jvr}; \theta_{vri}, a_{1v}) = N(y_{jvr}; \mu_{vri}, \sigma_{vri}, a_{1v}), \tag{7}$$

$$\mu_{vri} = l_v + (L_v - l_v)\frac{1 - k^{i-1}}{1 - k^{M-1}}; \quad i = 1, \ldots, M; v = 1, \ldots, V, \tag{8}$$

where the components of the mixture models are normally distributed [Equation (7)] and their means, μ_{vri}, are given by Equation (8). Component *SD*s are either assumed constant or constrained by Equation (5). Indices denote survey (v), year (r), component (i), and observation (j). Ψ is a vector that contains all unknown parameters in the model, V is the number of surveys, R is the number of years, M is the number of components, and n_{vr} is the total number of fish in any given survey and year. λ_{vri} is the mixing proportion for the ith component in the vth survey in the rth year, y_{jvr} is the length of fish j in the vth survey in the rth year, and θ_{vri} is the parameters of the ith component in the vth survey in the rth year. a_{1v} determines the assumed age of first component of the vth survey and is used to incorporate surveys of differing timings by back or forward projecting component means using Equations (2)–(4).

The complete data log-likelihood can be given as

$$\log L_c(\Psi) = \sum_{v=1}^{V}\sum_{r=1}^{R}\sum_{i=1}^{M}\sum_{j=1}^{n_{vr}}z_{jvri}\{\log\lambda_{vri} + \log f_{vri}(y_{jvr}; \theta_{vri}, a_{1v})\}, \tag{9}$$

where Equations (7) and (8) apply. z_{jvr} is a data label vector distributed according to a multinomial distribution of length of M with $z_{jvri} = (z_{jvr})_i = 1$ or 0, according to whether that particular fish or y_j belongs to the ith component in the vth survey in the rth year (McLachlan and Peel, 2004). However, as we do not know what component each fish belongs to the EM algorithm was used to treat z_{jvri} as missing data and estimate Ψ iteratively. Direct estimation of the observed log-likelihood is possible but not practical in this case and the EM algorithm offered a simple framework to implement the model in Minto *et al.* (2018). For details on the EM algorithm see Supplementary Appendix S1.

Hierarchical extensions

Hierarchical models were developed, incorporating bivariate random effects and allowing inter-annual or inter-cohort variation to be modelled. In contrast to the basic model that applies a single set of component means to the data, these models allow cohorts to be modelled through years and surveys. For the hierarchical models both "CSD and LSD" models were implemented. To avoid over-parameterization these hierarchical models were not extended beyond bivariate random effects and in each model either the key parameter for the asymptotic length (L) or the growth parameter (k) is random, not both. The logic behind this is also discussed by Eveson *et al.* (2007) and Lloyd-Jones *et al.* (2016), suggesting that a random effect on either K or L_∞ incorporates sufficient variability and has less bias than if both parameters were random.

Hierarchical models were fitted using the same number of components as the best fitting basic model (based on the model selection criteria described below). Three different hierarchical model structures were tested: (i) bivariate random effects on cohort specific l and L; (ii) bivariate random effects on cohort specific l and k; (iii) bivariate random effects on cohort specific l and yearly k. Correlation between the bivariate random effects in each scenario is also modelled. A random effect on l was specifically chosen to account for variability in the mean length of the first component in each cohort. Other random effects were chosen to investigate and incorporate growth variability. For details on the model formulation see Supplementary Appendix S1.

Model parameters, inputs, and outputs

A wide range of starting values for parameters of the basic model were tested (Table 1, see Table 2 for parameter descriptions), amounting to over 1000 model runs for each species and each *SD* type. To test the suitability of the model selection criteria for identifying the number of identifiable age classes in the length frequency data, the number of components tested ranged from 3 to 14 for haddock and 6 to 14 for white-bellied anglerfish. *SD*s for the random effects of l and k for the haddock hierarchical model were fixed at $\exp(-5)$ to give the model stability but still retain the random effect. More information on model stability in the haddock hierarchical models and general model inputs can found in Supplementary Appendix S1.

Table 1. Summary of the starting values tested for the model parameters of the basic model.

| Species | Parameter | Starting values tested | | |
		From	To	By
Haddock	l	10	10	NA
	L	40	70	10
	k	0.6	0.9	0.1
	s	6	6	NA
	S or σ	10	10	NA
	No. components	3	14	1
	λ_i	1/No. components	NA	NA
White-bellied anglerfish	l	16	16	NA
	L	100	130	10
	k	0.7	0.99	0.1
	s	6	6	NA
	S or σ	10	10	NA
	No. components	6	14	1
	λ_i	1/No. components	NA	NA

Model selection criteria and standard errors

Through exploratory simulation analysis it was found that Akaike information criterion (AIC) performed best from a range of model selection criteria. Model selection for this study was performed by selecting the model with highest observed log-likelihood value for each "number of components" tested. A variation on AIC we call "sub-AIC" was used to select the best fitting model. "sub-AIC" is similar to AIC, however the observed log-likelihood used is only contributed to by a subset of the length frequency data where reasonable numbers of fish were observed (i.e. haddock $< = 40$ cm & anglerfish $< = 100$ cm). This method ensures model selection is based on the majority of the data and not confounded by low fish numbers at high lengths. A similar method where model deviance is calculated from a subset of length classes when numbers in some length classes are low was used in model selection for fish selectivity curves (Millar and Fryer, 1999).

The EM algorithm does not directly provide estimates of uncertainty of parameters (McLachlan and Peel, 2004). Standard errors were approximated by first fitting the model to convergence with the EM algorithm, then using the maximum likelihood parameter estimates to optimize the observed data log-likelihood function [Equation (6)] (Minto *et al.*, 2018). Standard errors were approximated via the delta method with the R package "TMB" (Kristensen *et al.*, 2016).

Validation for haddock

Age data on the North Sea International Bottom Trawl Survey (NS-IBTS) are collected on a length-stratified basis. To obtain unbiased length-at-age data for haddock the length-stratified age data was used to construct annual age length keys (ALKs). Annual length frequency distributions (i.e. the same used in the length frequency analysis) were applied to their respective ALKs using the method described by Isermann and Knight (2005), resulting in an assigned age for each individual fish of the length frequency distribution. To enable comparison to the growth parameters estimated from the length frequency data alone, a VBGF was fit to this length-at-age data using maximum likelihood and mean length-at-age was also calculated.

Table 2. Summary of model parameters, and other relevant inputs/outputs of models.

Notation	Type	Description
Basic model		
L	Parameter	Mean of final component
l	Parameter	Mean of first component
k	Parameter	Controls rate of growth
s	Parameter	SD of first component
S	Parameter	SD of final component
j	Index	Individual fish length
i	Index	Component number
v	Index	Survey number
r	Index	Number of year
λ_{vri}	Parameter(s)	Mixing proportion of ith component in survey(v) and year(r)
μ_i	Output	Mean of ith component
σ_i	Output	SD ith component
L_∞	Output	CVBG parameter
K	Output	CVBG parameter
t_0	Output	CVBG parameter
a_{1v}	Input	Used to calculate t_0
M	Input	Number of components/ages assumed present in length frequency data
n_{vr}	Input	Number of observations(fish) in a given survey(v) and year(r)
Hierarchical models		
L_c	Parameter	Mean of final component of a cohort
l_c	Parameter	Mean of first component of a cohort
k_c	Parameter	Mean of random effect that controls rate of growth for a cohort
k_r	Parameter	Mean of random effect that controls rate of growth for a particular year
σ_L	Parameter	SD of random effect on L
σ_l	Parameter	SD of random effect on l
σ_k	Parameter	SD of random effect on k
ρ_{lk}	Parameter	Correlation of bivariate random effects on l, k for example
μ_{vri}	Output	Mean of ith component in survey(v) and year(r)

Not all hierarchical model parameters are used in each model.

Software

We develop a hybrid EM algorithm that uses the R package "TMB" within the "maximization" step (Kristensen *et al.*, 2016). Specifically, complete data log-likelihood functions were written in C++ using the TMB library class, compiled and automatically differentiated (Kristensen *et al.*, 2016). These objective functions, including gradients, were then optimized at each iteration with the statistical software R using the nlminb function within base R (R Core Team, 2018). The remaining computations of the algorithm and other analysis were conducted with R (R Core Team, 2018). The "FSA" R package (Ogle, 2018) was used to apply ALKs to the un-aged length frequency data, using the semi-random method from Isermann and Knight (2005).

Data

Haddock was chosen to validate the method as otolith age readings and growth parameters derived from lengths at age are reliable. White-bellied anglerfish was chosen because age readings are considered unreliable (ICES, 2011) and there is need for a cohort analysis to build an assessment model on. Both species are characterized by relatively fast growth and a discrete recruitment season, leading to
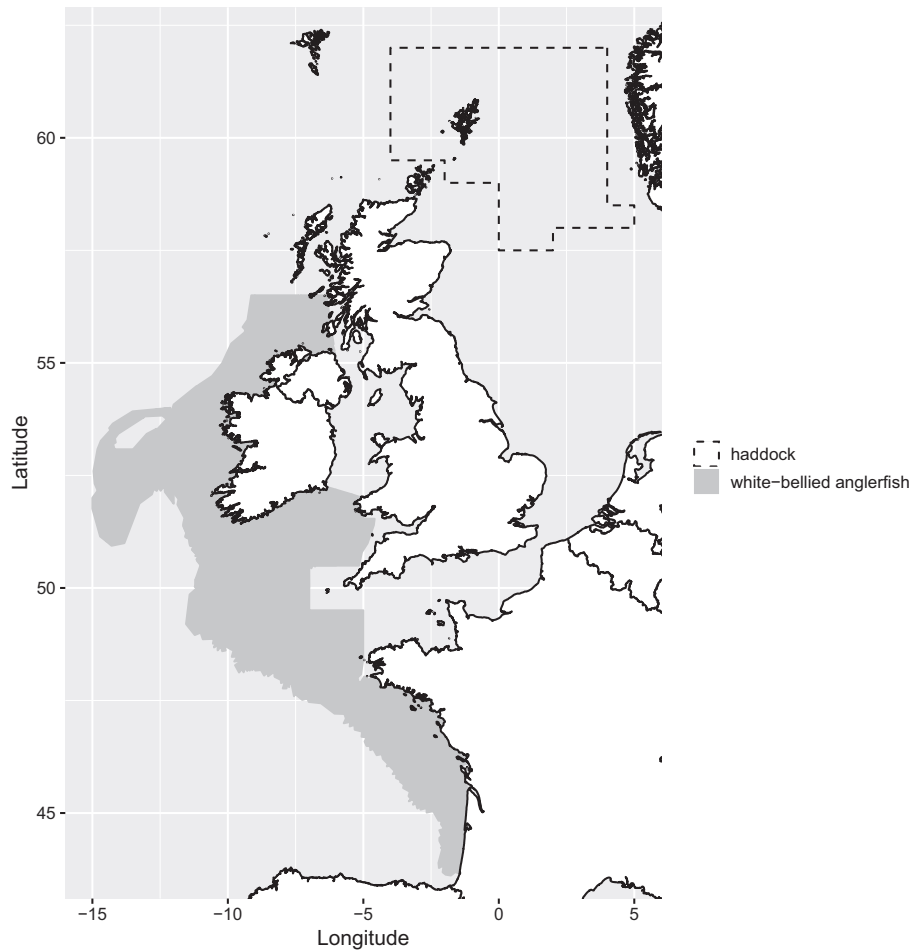
Figure 1. Areas of survey coverage from where length frequency data were obtained for haddock (standard roundfish area one of the NS-IBTS) and white-bellied anglerfish (EVHOE, IE-IGFS, and SP-PORC).

distinct length cohorts that can be tracked over time. Combined sex data was used for both species as is convention for haddock and white-bellied anglerfish in the most recent ICES assessment of the stock (ICES, 2018c). Survey catch data were obtained from the ICES DATRAS database (ICES, 2018b). Haddock length frequency data were from standard roundfish area one of the NS-IBTS (Figure 1). Only data collected in the third quarter of the year (Q3) were used. The corresponding age–length key for this area and time of year was also extracted from DATRAS. Length frequency data for white-bellied anglerfish were obtained from three surveys that cover the anglerfish stock in ICES areas 7.b-k, 8.a-b, and 8.d; the French EVHOE groundfish survey, the Irish groundfish survey (IE-IGFS) and the Spanish Porcupine groundfish survey (SP-PORC) (Figure 1). The French and Irish surveys were conducted mainly in the fourth quarter of the year and the Spanish survey in September. A Grande Ouverture Verticale fishing gear was used on all surveys with the exception of the Spanish Porcupine bank survey which used "porcupine baca" fishing gear (ICES, 2018a).

Results
Haddock
Haddock basic LFEM model
The basic model (single set of growth parameters estimated) was sensitive to starting parameters. Some of the model runs tended to

converge on k parameter values very close to 1. Whilst these gave good or even the best model fits according to the sub-AIC model selection criteria, the classical VBGF parameters were biologically unreasonable (i.e. L_∞ tending to very large values). Some model runs were discarded (406 of 1488 CSD models and 42 of 1488 LSD models) due to this issue and were identified by choosing an arbitrary value of 500 as the cut off for "biologically unreasonable" L_∞ estimates. To further filter out any spurious models the results were further refined by identifying and removing any models that had a second component mean $< = 15$cm. An example of why refining is important when a wide range of starting parameters are tested can be found in Supplementary Appendix S2.

A nine component model was the best fitting (lowest sub-AIC) for CSD and no models above nine components had reasonable fits for this *SD* type. LSD models gave generally lower sub-AIC values than CSD models and the best fitting model overall had 12 components and LSD (Table 3). Within both *SD* types, but particularly in LSD models, parameter estimates, and component means were consistent across the range of number of components tested (Tables 3 and 4). When compared to the VBGF estimated from length-at-age data ($\hat{K} = 0.499$, $\hat{L}_\infty = 39.5$, and $\hat{t}_0 \approx 0$), the VBGF from the best fitting basic LSD model is similar, although it does slightly underestimate early age mean lengths (Figure 2). The best fitting basic CSD model follows a similar growth pattern until the fifth

Table 3. von Bertalanffy growth parameters from (lowest sub-AIC) model runs on haddock length frequency data after refining valid model runs.

SD type	No. components	RVBG parameters			CVBG parameters			
		$\hat{l}$ (cm)	$\hat{L}$ (cm)	$\hat{k}$	$\hat{L}_\infty$ (cm)	$\hat{K}$	$\hat{t}_0$	sub-AIC
CSD	3	10.5 (0.008)	32.4 (0.007)	0.816 (0.0012)	76	0.20	−0.10	6 167 506
	4	10.3 (0.006)	37.1 (0.011)	0.780 (0.0007)	61	0.25	−0.12	6 031 048
	5	10.3 (0.005)	39.7 (0.014)	0.754 (0.0006)	53	0.28	−0.13	5 989 126
	6	10.3 (0.005)	41.8 (0.018)	0.750 (0.0005)	51	0.29	−0.15	5 974 158
	7	10.3 (0.005)	43.0 (0.021)	0.752 (0.0005)	50	0.29	−0.18	5 957 923
	8	10.4 (0.004)	44.0 (0.022)	0.761 (0.0004)	49	0.27	−0.23	5 946 292
	9	10.4 (0.004)	45.0 (0.026)	0.773 (0.0004)	50	0.26	−0.27	5 941 101
LSD	3	10.1 (0.003)	31.3 (0.008)	0.854 (0.0013)	88	0.16	−0.14	5 982 480
	4	10.1 (0.003)	34.9 (0.012)	0.720 (0.0007)	49	0.33	−0.07	5 918 972
	5	10.1 (0.003)	37.0 (0.015)	0.689 (0.0006)	44	0.37	−0.06	5 901 698
	6	10.1 (0.003)	37.5 (0.016)	0.665 (0.0005)	41	0.41	−0.06	5 889 986
	7	10.1 (0.003)	38.4 (0.018)	0.661 (0.0005)	40	0.41	−0.06	5 886 624
	8	10.1 (0.003)	39.0 (0.020)	0.658 (0.0004)	40	0.42	−0.06	5 885 749
	9	10.1 (0.003)	39.5 (0.021)	0.658 (0.0004)	40	0.42	−0.06	5 885 240
	10	10.1 (0.003)	40.0 (0.023)	0.661 (0.0004)	40	0.41	−0.06	5 885 145
	11	10.1 (0.003)	40.4 (0.024)	0.663 (0.0004)	40	0.41	−0.07	5 885 031
	12	10.1 (0.003)	40.6 (0.025)	0.665 (0.0004)	40	0.41	−0.07	5 885 004
	13	10.1 (0.003)	40.8 (0.026)	0.666 (0.0004)	41	0.41	−0.07	5 885 005
	14	10.1 (0.003)	40.9 (0.027)	0.667 (0.0004)	41	0.41	−0.07	5 885 023

Number of components is the number of ages assumed to be present in the data. Rows highlighted in grey are the models with the lowest sub-AIC model selection criteria values. Estimated standard errors for the corresponding RVBG parameters are given in parentheses. RVBG parameter estimates are given to 3 s.f. and CVBG to 2 s.f.

Table 4. Component/cohort means and their SD parameters for haddock, corresponding to the basic models presented in Table 3.

SD type	No. components	Component													$\hat{s}$	$\hat{S}$
		1	2	3	4	5	6	7	8	9	10	11				
CSD	3	10.5	22.6	32.4												3.66
	4	10.3	21.5	30.2	37.1											2.89
	5	10.3	21.0	29.0	35.1	39.7										2.65
	6	10.3	20.6	28.4	34.2	38.5	41.8									2.50
	7	10.3	20.2	27.7	33.3	37.5	40.7	43.0								2.39
	8	10.3	19.8	27.0	32.4	36.6	39.8	42.2	44.0							2.29
	9	10.4	19.4	26.4	31.9	36.1	39.3	41.8	43.8	45.3						2.21
LSD	3	10.0	21.5	31.3											1.63	4.93
	4	10.1	21.1	29.1	34.9										1.65	4.59
	5	10.1	20.9	28.3	33.5	37.0									1.66	4.42
	6	10.1	20.7	27.7	32.4	35.5	37.5								1.67	4.28
	7	10.1	20.6	27.5	32.1	35.1	37.1	38.4							1.67	4.23
	8	10.1	20.6	27.4	32.0	34.9	36.9	38.2	39.0						1.68	4.22
	9	10.1	20.5	27.4	31.9	34.9	36.8	38.1	39.0	39.5					1.68	4.21
	10	10.1	20.5	27.4	31.9	34.9	36.9	38.2	39.1	39.6	40.0				1.68	4.20
	11	10.1	20.5	27.3	31.9	34.9	36.9	38.3	39.1	39.7	40.1	40.4			1.68	4.19
	12	10.1	20.4	27.3	31.9	34.9	37.0	38.3	39.2	39.8	40.2	40.4	40.6		1.68	4.18
	13	10.1	20.4	27.3	31.9	34.9	37.0	38.3	39.2	39.8	40.2	40.5	40.7	40.8	1.68	4.18
	14	10.1	20.4	27.3	31.9	35.0	37.0	38.3	39.3	39.9	40.3	40.5	40.7	40.8 40.9	1.69	4.18

In the case of CSD models S = s but values are only given in the final (S) column.

component, the following component means then appeared to overestimate the mean length of fish in comparison to the LSD model and the length-at-age estimated model fit (Figure 2). Estimated SD was very similar within SD types (Table 4).

Haddock hierarchical LFEM model with cohort specific *l* and *k*

Overall CVBG parameters for the hierarchical CSD model ($\hat{K}$ = 0.241, $\hat{L}_\infty$ = 52.9 cm, and $\hat{t}_0$ = −0.31) were similar to those given by the basic CSD model and followed a similar trajectory. Hierarchical LSD model parameters ($\hat{K}$ = 0.394, $\hat{L}_\infty$ = 42.1 cm− and $\hat{t}_0$ = −0.06) were also similar to estimates from its corresponding basic model (Table 3 and Figure 2).

The performance of the hierarchical *l* and *k* LSD model (corresponding to the best fitting basic model) was examined in more detail. Ranges of the cohort specific L_∞ estimates were 41.8–42.5 cm for $\hat{L}_\infty$ and 0.334–0.463 for $\hat{K}$. Correlation
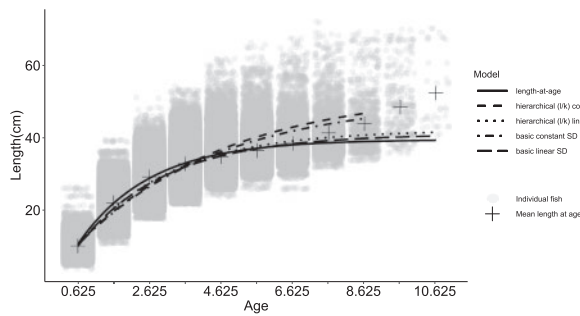
Figure 2. Length-at-age data and estimated von Bertalanffy growth curves for haddock. Showing mean length-at-age, the von Bertalanffy model fit (maximum likelihood estimated) on the length-at-age data, growth curves from best fitting basic length frequency models, and overall growth curves (i.e. means of RE parameters) from hierarchical *l*/*k* models. Ages are jittered for visualization.

parameter $\hat{\rho}$ for the random effects l and k was estimated as −0.23. Cohort-specific random effects gave the model flexibility in its estimated growth trajectories for each cohort. When examining the cohorts that were observed at least until the ninth component over the period analysed (2000–2009) the estimated cohort growth curves from this study's hierarchical model appear to fit reasonably well in some cohorts (e.g. cohorts 2002, 2003, and 2004) and very well in others (e.g. cohorts 2000, 2005, and 2009) (Figure 3). When cross-referenced with Figure 4 the hierarchical model cohort curves that are very similar to length-at-age estimated cohort curves are those cohorts where a strong signal can be seen through a number of years (i.e. cohort 2000) (Figures 3 and 4). Deviations (difference between LFEM model length-at-age and ML estimated length-at-age) of these cohorts showed that in the early ages of some cohorts the hierarchical LFEM model underestimated the mean length-at-age in comparison to the ML fit on the length-at-age data (Supplementary Appendix S2). In the later years of the length frequency distributions the hierarchical model begins to severely underestimate growth (Figure 4). Also of note is the flexibility the random effect on the mean of the first component (*l*) gives when modelling the shifting position of the first component over the years, such as in 2006 or 2013 (Figure 4).

White-bellied anglerfish

White-bellied anglerfish basic LFEM model

As with haddock, the basic LFEM model was sensitive to starting parameters, particularly *k* and some of the model runs tended to converge on *k* parameter values very close to 1. These models often gave good or even the best model fits according to the sub-AIC model selection criteria, however the classical VBGF parameters were biologically unreasonable (i.e. L_∞ tending to very large values). A large number of model runs was discarded (720 of 1080 CSD models and 740 of 1080 LSD models) due to this issue and were identified by choosing an arbitrary value of 500 as the cut off for "biologically unreasonable" L_∞ estimates.

Overall, the nine component model for LSD was best fitting (lowest sub-AIC) and LSD models consistently had lower sub-AIC values than their counterpart CSD models. For LSD models as the number of components assumed was increased the final
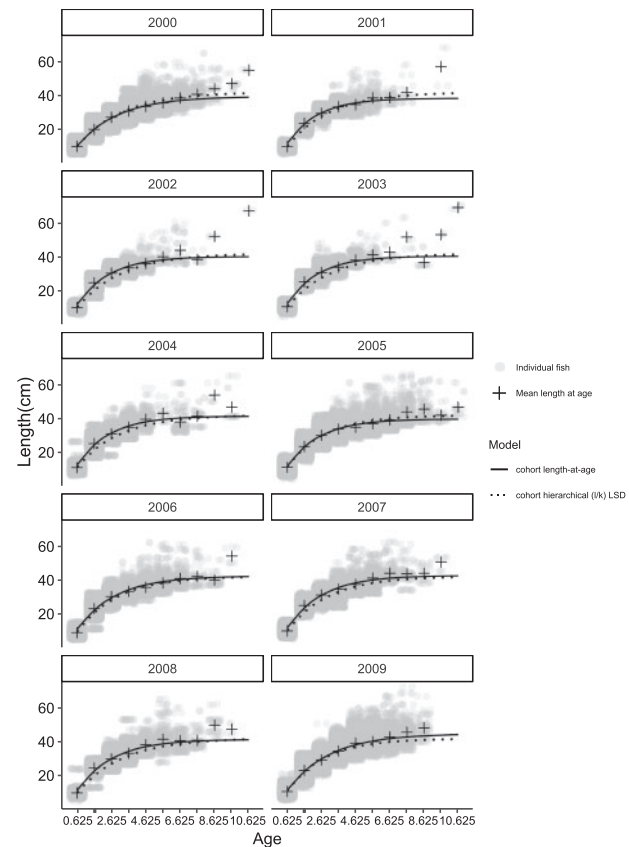

Figure 3. Length-at-age data and estimated von Bertalanffy growth curves for cohorts of haddock observed up to at least their 9th component. Showing mean length-at-age, the von Bertalanffy model fit (maximum likelihood estimated) on the length-at-age data of the cohort and cohort-specific growth curves estimated within the LSD hierarchical model that varies *l* and *k*. Ages are jittered for visualization.

component means (L) increased, as did the CVBG parameter K, while L_∞ decreased. CSD models showed no such pattern (Tables 5 and 6).

Component means were consistent between and within *SD* types up to the fourth component particular for models with an assumed number of components of 11 or above. LSD model component means did not change with increasing number of components assumed and the best fitting model across all models (LSD, nine components) had very similar component means to the best fitting CSD model (Table 6). When plotted on the raw length frequency data the best fitting basic model appears to fit the data reasonably well in some years but not in others and cohort progression can be seen up until the seventh component in some cohorts (Figure 6a).

The growth curve of the parameters estimated from the best fitting basic LFEM model differs somewhat from the growth curves estimated in a previous study (Landa *et al.*, 2013). Growth curves were reasonably similar up to age six, although this study's VBGF shows marginally faster growth. After age six the growth curves begin to diverge more, with this study estimates of length-at-age markedly higher (Figure 5a).
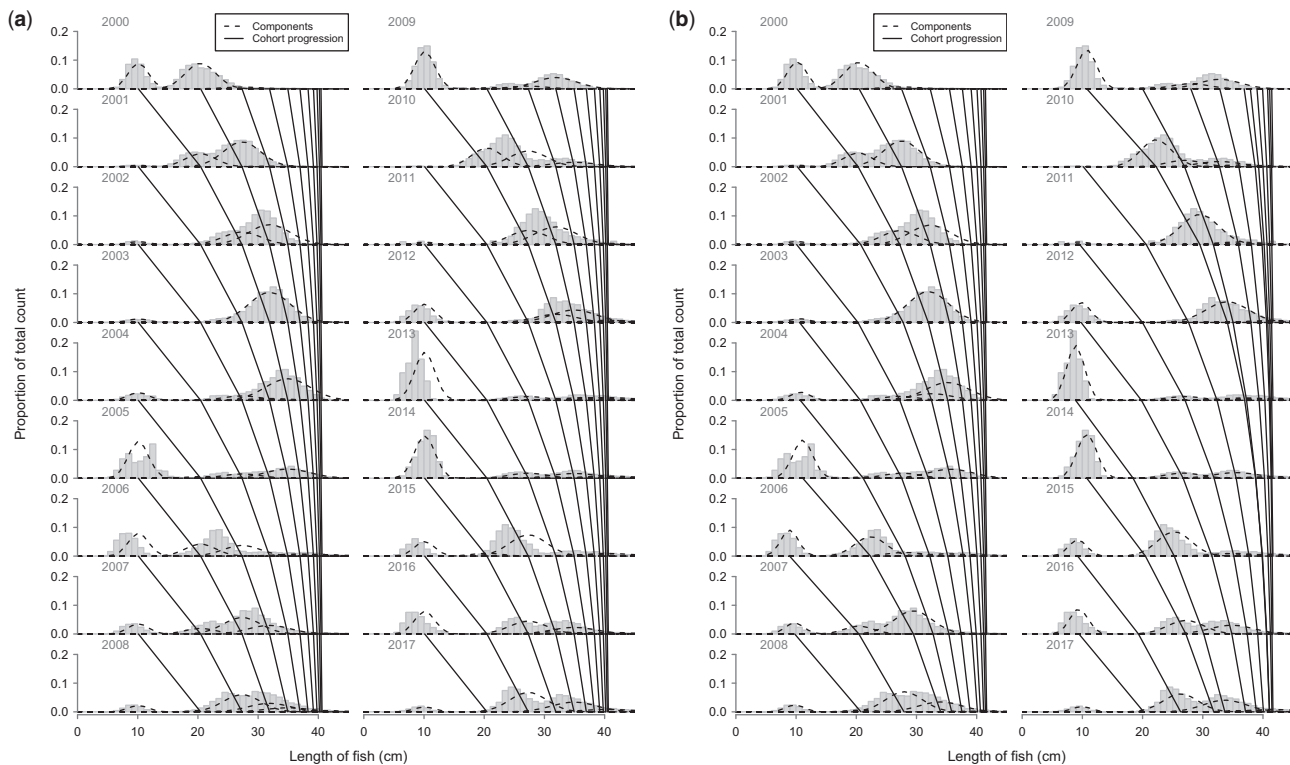
Figure 4. Length frequency distributions of haddock with fitted distributions from: (a) the lowest sub-AIC basic model (12 components and LSD) and (b) the corresponding hierarchical model with bivariate random effects on cohort specific *l* and *k*. Also shown is the progression of the cohort mean values. Note that the maximum length of fish was 72 cm but only fish up 45 cm were plotted here as there was little data above this length, all estimated cohorts are displayed.

White-bellied anglerfish hierarchical LFEM model with cohort specific *l* and *L*

Overall CVBG parameters for the hierarchical LSD model with cohort specific *l* and *L* corresponding to the best fitting basic model ($\hat{K} = 0.089$, $\hat{L}_\infty = 215.2$ cm, and $\hat{t}_0 = -0.14$) differed from those given by the basic model. $\hat{K}$ was increased and $\hat{L}_\infty$ reduced dramatically, $\hat{t}_0$ was also slightly reduced. The resulting VBGF curve was different from the basic model curve as well as the growth curves from a previous study (Landa *et al.*, 2013). The hierarchical model showed faster growth than the other curves and estimated component means somewhat different from the best fitting basic LFEM model (Figure 5a). Cohort-specific growth curves from the hierarchical LFEM model showed a range of trajectories and cohort $\hat{L}_\infty$ estimates ranged from 184.9 to 246.6 cm (Figure 5b). The slowest growing cohort with the lowest estimated L_∞ (2014 cohort) followed a similar trajectory to the basic LFEM model (Figure 5).

There was a negative correlation estimated ($\hat{\rho} = -0.69$) between the bivariate random effects *l* and *L* of each cohort (Figure 7). The bivariate cohort-specific random effects show a moderate degree of variation over a few centimetres for $\hat{l}$ and $\sim$30 cm for $\hat{L}$. There is also a negative trend in *L* (Figure 7b). There is a high degree of uncertainty around estimates of *L* for the last few cohorts (Figure 7b).

Other white-bellied anglerfish hierarchical models

Hierarchical models with a cohort specific *k* or yearly varying *k* were also fit to the anglerfish data. In both cases the RVBG

parameter *K* varies to a large extent between cohorts, however L_∞ also varies and the resulting cohort curves differ but not dramatically. Parameter estimates can be found in Supplementary Appendix S3.

Discussion
Overview of LFEM

This study aimed to improve upon currently available approaches for growth estimation of fish through length frequency analysis. A basic method utilizing the EM algorithm and analogous to MULTIFAN (Fournier *et al.*, 1990) was produced. The methodology was then extended further to incorporate bivariate random effects to capture some of the growth variation in length frequency data.

Few studies have modelled growth variation in length frequency data (Fournier *et al.*, 1998; Roa-Ureta, 2010; Lloyd-Jones *et al.*, 2016). The approach of using a VBGF to constrain mixture model component means through a time series (i.e. along a cohort) whilst explicitly estimating cohort or year specific random effects on two growth parameters is a worthwhile extension to this area of research [see Lloyd-Jones *et al.* (2016) for an alternative]. The aim of this integrated approach was to aid identification of growth parameters in variable length frequency distributions obtained from multiple annual surveys. Furthermore, the associated uncertainty of the parameter estimates is also approximated. Overall both versions of the methodology produced reasonable estimates of von Bertalanffy growth parameters and their variability between cohorts/years. However,

Table 5. von Bertalanffy growth parameters from LFEM model runs that gave the lowest sub-AIC value on white-bellied anglerfish length frequency data.

		RVBG parameters			CVBG parameters			
SD type	No. components	$\hat{l}$ (cm)	$\hat{L}$ (cm)	$\hat{k}$	$\hat{L}_\infty$ (cm)	$\hat{K}$	$\hat{t}_0$	sub-AIC
CSD	8	17.1 (0.090)	114.4 (0.572)	0.966 (0.003)	469	0.035	−0.20	99 430
	9	17.0 (0.087)	118.6 (0.621)	0.945 (0.003)	294	0.057	−0.17	99 397
	10	17.0 (0.085)	119.9 (0.621)	0.930 (0.003)	232	0.072	−0.17	99 419
	11	16.6 (0.082)	125.3 (0.653)	0.974 (0.002)	489	0.026	−0.45	99 341
	12	16.6 (0.081)	128.4 (0.699)	0.969 (0.002)	401	0.031	−0.48	99 367
	13	16.6 (0.080)	130.1 (0.734)	0.951 (0.002)	267	0.050	−0.40	99 419
	14	16.5 (0.078)	126.8 (0.653)	0.940 (0.002)	215	0.062	−0.41	99 444
LSD	8	16.5 (0.076)	107.3 (0.919)	0.941 (0.005)	276	0.061	−0.13	99 112
	9	16.3 (0.075)	108.7 (0.997)	0.967 (0.005)	412	0.033	−0.34	99 102
	10	16.3 (0.076)	113.1 (0.999)	0.974 (0.004)	471	0.027	−0.45	99 110
	11	16.4 (0.076)	114.9 (1.017)	0.951 (0.004)	264	0.051	−0.39	99 173
	12	16.4 (0.076)	116.3 (0.992)	0.938 (0.003)	213	0.064	−0.37	99 237
	13	16.4 (0.076)	120.1 (1.022)	0.930 (0.003)	195	0.072	−0.34	99 312
	14	16.4 (0.076)	122.1 (1.077)	0.921 (0.003)	177	0.082	−0.31	99 393

Number of components is the number of ages assumed to be present in the data. Parameters *l* and *L* are those estimated for the surveys that occurred in the fourth quarter of the year (EVHOE and IE-IGFS). Estimated standard errors for the corresponding RVBG parameters are given in parentheses. RVBG parameters estimates are given to 3 s.f. and CVBG to 2 s.f. (with the exception of L_∞). The Spanish survey estimates of can be found in Supplementary Appendix S3.

Table 6. Component/cohort means and their *SD* parameters for white-bellied anglerfish, corresponding to the basic models presented in Table 5.

		Component															
SD type	No. components	1	2	3	4	5	6	7	8	9	10	11	12	13	14	$\hat{s}$	$\hat{S}$
CSD	8	17.1	32.5	47.3	61.7	75.5	88.9	101.9	114.4								4.91
	9	17.0	32.4	46.9	60.6	73.6	85.8	97.4	108.3	118.6							4.75
	10	16.9	32.0	45.9	58.9	71.0	82.3	92.7	102.5	111.5	119.9						4.62
	11	16.6	28.8	40.7	52.3	63.5	74.5	85.2	95.6	105.8	115.7	125.3					4.34
	12	16.6	28.4	39.8	50.9	61.7	72.1	82.3	92.1	101.6	110.8	119.7	128.4				4.23
	13	16.6	28.9	40.6	51.7	62.2	72.3	81.9	90.9	99.6	107.8	115.6	123.0	130.1			4.20
	14	16.5	28.5	39.8	50.4	60.3	69.7	78.5	86.7	94.5	101.8	108.6	115.1	121.1	126.8		4.12
LSD	8	16.5	32.0	46.5	60.2	73.1	85.2	96.6	107.3							3.96	11.15
	9	16.3	29.3	41.8	53.9	65.6	76.9	87.9	98.5	108.7						3.83	12.03
	10	16.3	28.2	39.9	51.2	62.2	72.9	83.4	93.5	103.5	113.1					3.85	9.90
	11	16.3	28.6	40.2	51.3	61.8	71.8	81.4	90.4	99.0	107.2	114.9				3.86	9.28
	12	16.4	28.6	40.1	50.9	61.1	70.5	79.4	87.8	95.6	103.0	109.9	116.3			3.87	8.55
	13	16.4	28.9	40.5	51.3	61.3	70.6	79.3	87.4	94.9	101.9	108.4	114.5	120.1		3.90	7.72
	14	16.4	29.1	40.8	51.6	61.5	70.6	79.1	86.8	94.0	100.5	106.6	112.2	117.3	122.1	3.91	7.60

In the case of CSD models S = s but values are only given in the final (S) column.

there were problems with convergence to biologically unreasonable estimates of CVBG parameters particularly in the basic models, as well as issues with model stability in the hierarchical models where information was intermittently sparse (i.e. weak or low abundance cohorts).

Haddock

Basic and overall hierarchical parameter estimates for haddock were within or close to previously reported ranges ($K \approx 0.1$–0.5 and $L_\infty \approx 45$– 55 cm) (Baudron *et al.*, 2011). Parameter estimates from the best fitting basic model (12 components and LSD) and overall parameter estimates from the *l*/*k* hierarchical LSD model gave very similar growth curves to that of a maximum likelihood fit on length-at-age data for fish from the same region and time

of year (Figure 2). CSD models overestimated growth somewhat in older fish indicating the importance of identifying the correct *SD*. When applying their method to data for pike and abalone Schnute and Fournier (1980) also found that *SD* treatment was important when deriving growth from observed length frequencies. The number of components identified by the best fitting basic LFEM model (12 components) is close to the number of ages present in the ALK (11 ages). The model selection criterion was also particularly useful for identifying the more suitable *SD* type. The effect of misidentifying component number is also negated somewhat by the consistency of the component means with differing number of components (Table 4).

From the hierarchical model (random effects on *l* and *k*) we can see a similar variation in cohort-specific growth parameters as Baudron *et al.* (2011) found in their study. Hierarchical
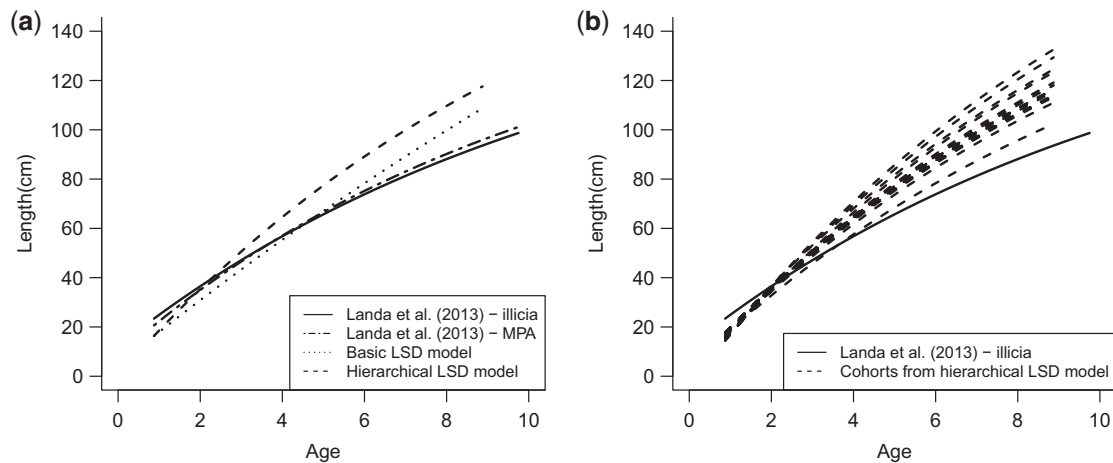
Figure 5. von Bertalanffy growth curves estimated for white-bellied anglerfish from: (a) the best fitting basic model and overall parameters from the corresponding hierarchical model with a cohort-specific random effect on l and L, as well as growth curves [estimated from illicia and modal progression analysis (MPA)] from a previous study (Landa *et al.,* 2013), (b) estimated cohort growth curves from the hierarchical LSD model with cohort-specific random effects on l and L.
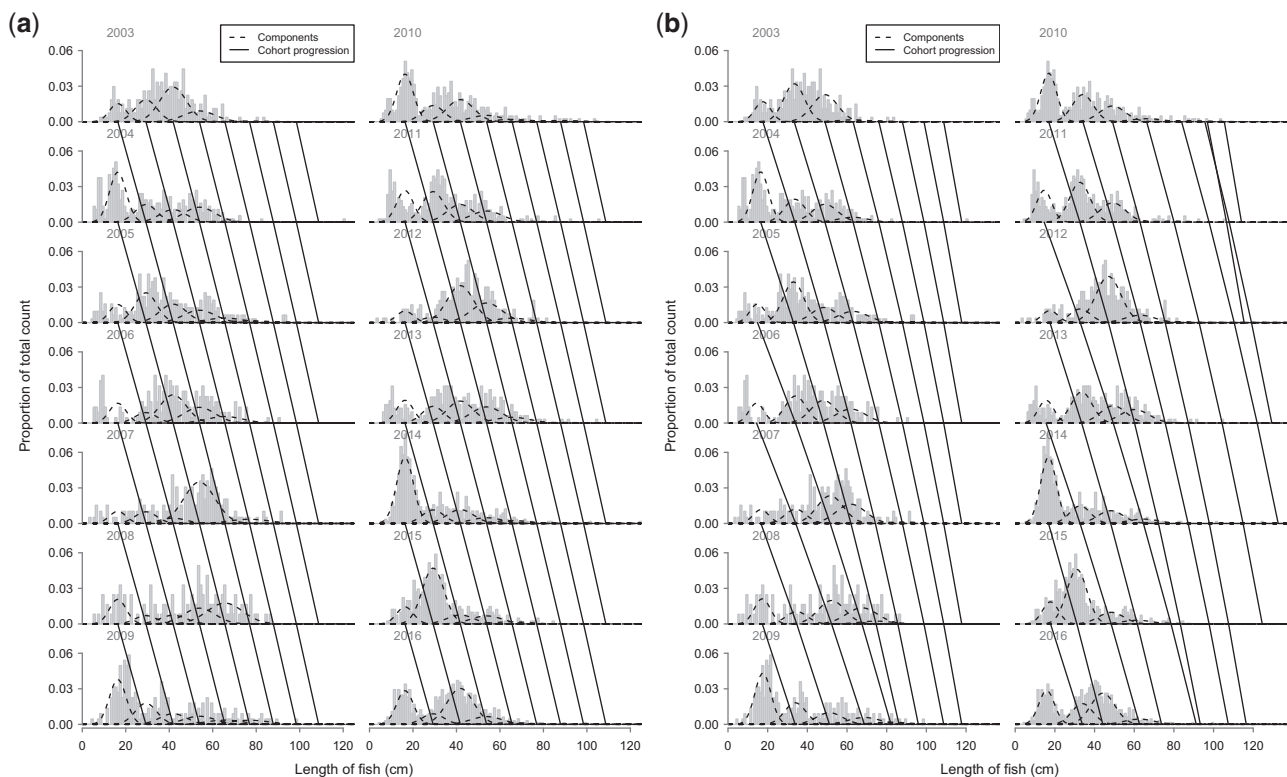

Figure 6. Length frequency distributions of white-bellied anglerfish from the quarter four IE-IGFS with fitted distributions from: (a) the lowest sub-AIC basic LFEM model (9 components and LSD) and (b) the corresponding hierarchical model with bivariate random effects on cohort specific l and L. Also shown is the progression of the cohort mean values (black solid lines). Similar plots for the Spanish Porcupine Bank survey (SP-PORC) on the French groundfish survey (EVHOE) can be found in Supplementary Appendix S3.

model and length-at-age estimated cohort growth curves were similar in cohorts observed up until the ninth component. However outside of these cohorts growth was often underestimated, indicating the effect of the lack of information in later

ages. Cohort growth curves that showed the most similar fits to the data between the two methods also tended to be for cohorts where strong signals were present in sequential years (Figure 3 and 4).
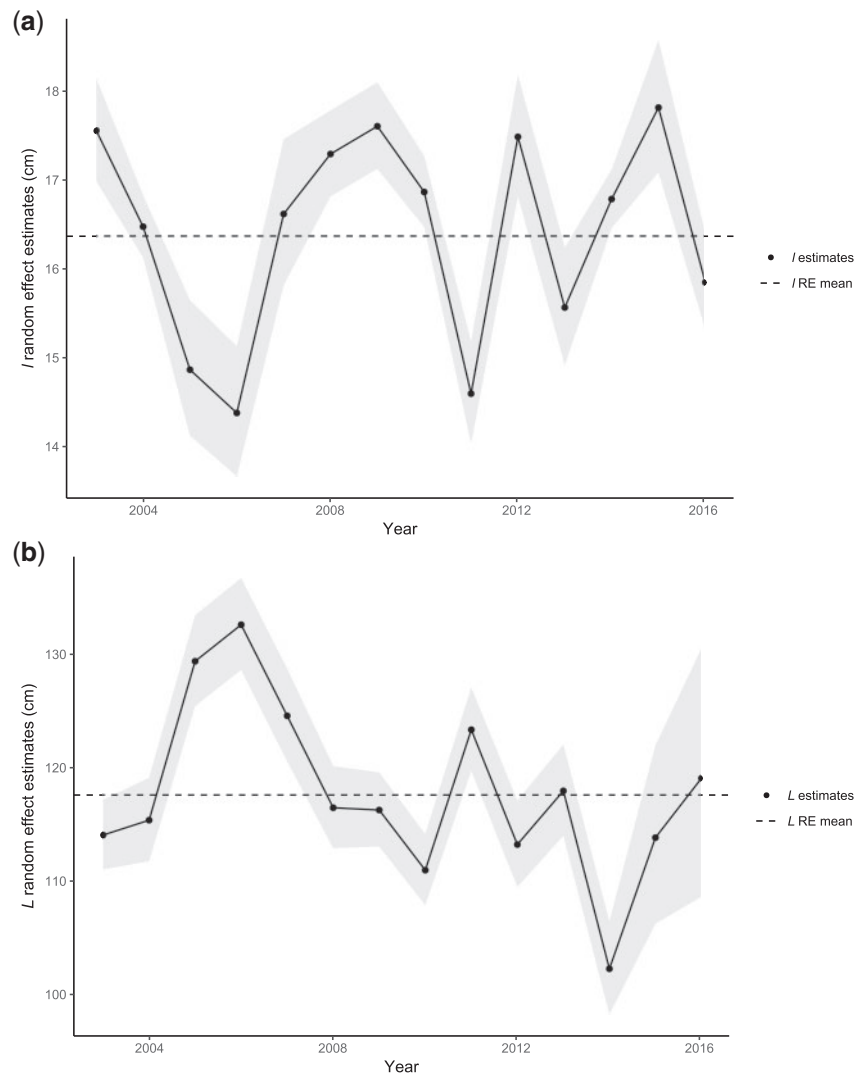
Figure 7. Estimated random effects on first (*I*) and last (*L*) component means of cohorts from a nine component hierarchical LSD model on length frequency distributions of white-bellied anglerfish. Shaded area denotes 2*standard error of the random effect estimates and dashed lines denote the mean of the random effect. Parameter estimates are for quarter four surveys (EVHOE and IE-IGFS).

White-bellied anglerfish

Parameter estimates for white-bellied anglerfish (both sexes combined) were broadly similar to those in the literature, although differences are more pronounced in the later ages and estimates of L_∞ from this study were consistently higher than previously reported (Farina *et al.*, 2008; Landa *et al.*, 2013).

The basic and hierarchical models produced different growth curves, despite having the same starting parameters. The 2014 cohort from the hierarchical model follows a similar trajectory to the best fitting basic LFEM model growth curve, indicating that 2014–2016 (the years 2014 cohort is observed) may be dominating in the estimation of the basic model. Flexibility within the hierarchical model estimated a slower growth for the 2014 cohort, whilst estimating a slightly faster growth for other cohorts. Further investigations showed that when years 2014–2016 (these years had a higher number of fish than other years) were removed from the data the basic LFEM model component means were more alike to the hierarchical model estimates for the majority of years. Subjectively, fitted distributions of the hierarchical model

seem to fit better than the basic LFEM model (Figure 6). Estimated growth curves predicted faster growth particularly in the older fish compared to anglerfish age validation studies (Figure 5) (Farina *et al.*, 2008; Landa *et al.*, 2013). Faster growth in the later ages was less pronounced in models that assumed a higher number of components, however these were not selected as the best fitting model (Table 5).

Model selection based on sub-AIC indicated that a nine component model with LSD provided the best fit to the data. Although *L. piscatorius* is thought to live frequently to over ten years of age (Farina *et al.*, 2008), fish of this age are not sufficiently well represented in the length frequency data to support the selection of a large number of components. The nine component model produced a higher estimate of L_∞ and a lower estimate of *K* compared to other studies (Farina *et al.*, 2008; Landa *et al.*, 2013), however, the estimates of components means and *SD*s fit the length frequency distributions reasonably well (Figure 6).

As with haddock *SD* treatment was important. Including a linear increasing component *SD* produced components with very

wide length distributions and mixing proportions appeared to be biased towards a single component in some cases. CSD allowed for more overlap between components and more balanced mixing proportions but not necessarily a better fitting model (Table 5).

Overall, discrepancies between parameter estimates from this study and those from studies that use direct ageing methods may reflect modelling challenges with the data or the difficulties of obtaining reliable age estimates for anglerfish (ICES, 2011). Landa *et al.* (2013) used length frequency analysis techniques to validate direct ageing of illicia and found good agreement between the parameter estimates produced by both methods. However, that study was based on survey data from the Porcupine Bank while this study also included data from two other surveys as model inputs. Further difficulty in estimating growth may have been introduced here by combining data for males and females as white-bellied anglerfish show sexual dimorphism in growth (Farina *et al.*, 2008).

Caveats

The EM algorithm is an accepted method for estimating latent variables but is sensitive to the choice of starting parameters (Biernacki *et al.*, 2003), as was evident from the model outputs presented here. This was accounted for by conducting sensitivity runs on some of the main parameters (k, L and number of components), producing a wide range of values for each parameter estimate. Some other starting parameters were fixed, such as the mean of the first component (l) and the starting mixing proportions, which was arbitrarily set as $1/\text{No. components}$. It is likely that the starting mixing proportions in conjunction with a k growth rate parameter drove the variation in parameter estimates at convergence. Lloyd-Jones *et al.* (2016) modelled length frequency distributions using a methodology similar to the one employed here, but used a MM algorithm to estimate parameters.

The number of sensitivity runs was large for this study (>1000 for each SD type) and a wide range of starting parameter values was used. Many k parameter starting values were very different from known estimates, which may partially account for the number of model runs that were excluded before further analysis. This study's approach to parameter sensitivity by multiple model runs followed by exclusion of models with "biologically unreasonable" parameter estimates is a robust method for dealing with the sensitivity, although caution should be used when setting a threshold (e.g. this study used a L_∞ of 500 cm, so as not to introduce subjectivity into model selection). Haddock required a further refining of models by identifying and removing any models with a second component mean $<=15$ cm. Second component means $<=15$ cm were caused by certain (slow growth) starting parameters and the second component mixing proportion tending to zero.

In this study, the sub-AIC model selection criteria appeared to work reasonably well (details of sub-AIC are given in "Model selection criteria and standard errors" section). sub-AIC was chosen as the model selection criteria in part due to the performance of AIC in exploratory simulation analysis with the LFEM model. AIC was able to identify the correct number of components when the SD of components was small (~1), as did Bayesian information criterion (BIC). With increasing SD of length frequency data AIC was more robust, still under-fitting the data somewhat but much less severely than other model selection criteria. However, model selection is often problematic in studies such as this, hence

should be conducted cautiously using expert knowledge of the species in question (Schnute and Fournier, 1980; Fournier *et al.*, 1990; Taylor and Mildenberger, 2017).

Correlation between parameters could contribute to the observed variation in parameter estimates. CVBG parameters (K and L_∞) have been shown to covary strongly (Gallucci and Quinn, 1979; Pilling *et al.*, 2002). Here, the re-parameterized von Bertalanffy growth parameters appeared strongly correlated, with many combinations of k and L parameters resulting in very similar model fits. The same issue arises with other length frequency analysis methods such as MULTIFAN and ELEFAN and is usually dealt with by fixing either K or L_∞ in the initial stages of the analysis (Fournier *et al.*, 1990; Taylor and Mildenberger, 2017). Variation in estimated maximum likelihood parameters could also be driven by the substantial variability that exists in the input data.

Hierarchical models appeared to be less sensitive to the choice of starting parameters, although due to computational demands these models were not subjected to the same level of sensitivity testing. Model testing using the haddock data showed that when length frequency data indicates very variable recruitment (i.e. in some years the first component is almost not present) the hierarchical model initially struggled to estimate these components, shifting the first component to the second component position and thus confounding estimates of cohort progression. It seemed that there was not sufficient information within the estimated distribution of the random effect to prevent this. This issue is a form of the label switching problem that is well known in mixture models (Yao, 2015). Fixing the SD of the bivariate random effects at a suitably low value [i.e. $\exp(-5)$] as detailed in "Model selection criteria and standard errors" section, retained enough flexibility in the model to estimate variable random effects but sufficient constraint to keep l and k random effects estimates within reasonable bounds, thus preventing label switching.

Further work

A comprehensive comparison of the basic and hierarchical models with other methodologies such as MULTIFAN, the updated ELEFAN procedure, and the more subjective Bhattacharya method would be a useful avenue for further research (Bhattacharya, 1967; Fournier *et al.*, 1990; Taylor and Mildenberger, 2017). In MULTIFAN-CL density dependence can be included in the structure of the model as abundance of a cohort can effect growth rate in some fish species (Fournier *et al.*, 1998). The hierarchical model presented in this study is currently designed to model growth variability between cohorts independently of cohort abundance but could be modified to include cohort abundance as a covariate, increasing complexity and modelling the connection between mixing probabilities of a cohort. Laslett *et al.* (2004) and Roa-Ureta (2010) devised stage-based approaches to model fitting. A comparison of these stage-based approaches with hierarchical model presented here would be useful to the assess the effectiveness of integrating constraints and random effects into the mixture model parameter estimation.

Currently, the components of the mixture models are assumed to be normally distributed. Further development of the method could enable the user to specify alternative distributions such as log normal, gamma, or other, thus allowing non-normality and skewness to be modelled (Macdonald and Pitcher, 1979; Lloyd-Jones *et al.*, 2016).

Conclusions

Use of length frequency analysis to obtain growth parameters is a well-developed area of research (Hasselblad, 1966; Fournier *et al.*, 1990, 1998; Taylor and Mildenberger, 2017). These techniques provide growth information when age of individuals cannot be determined otherwise. Parameter estimates, their associated uncertainty and model fits presented here demonstrate the usefulness of our method for obtaining growth information from length frequency data.

Growth parameters are estimated whilst simultaneously fitting finite mixture models using the EM algorithm. The use of the EM algorithm removes the need to define length bins, allowing a more objective set up of the modelling procedure. Models can take inputs from multiple surveys from different times of year and offer an alternative framework to other well-known approaches for estimating growth parameters from length frequency data. Furthermore, within the hierarchical framework we explicitly model bivariate random effects on growth parameters, successfully modelling growth variability and allowing trends to be identified in the estimates.

Where age data are not readily available modelling cohort growth variability through this study's methodology can improve information available for stock assessments (i.e. cohort-specific growth curves for age slicing). In addition, estimates from these length frequency models give useful credible estimates of the boundaries of growth parameters that could be tested for stock assessments in a simulation context, such as in a management strategy evaluation (Punt *et al.*, 2013). Finally, length frequency distribution derived cohort growth parameter estimates could also be used to investigate environmental covariates such as sea temperature (Baudron *et al.*, 2014; Barrow *et al.*, 2018).

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Acknowledgements

We'd also like to thank ICES, the reviewers and editor for their comments and those that provided and collated the data that was used in the study.

References

Barrow, J., Ford, J., Day, R., and Morrongiello, J. 2018. Environmental drivers of growth and predicted effects of climate change on a commercially important fish, *Platycephalus laevigatus*. Marine Ecology Progress Series, 598: 201–212.

Baudron, A. R., Needle, C. L., and Marshall, C. 2011. Implications of a warming North Sea for the growth of haddock *Melanogrammus aeglefinus*. Journal of Fish Biology, 78: 1874–1889.

Baudron, A. R., Needle, C. L., Rijnsdorp, A. D., and Tara Marshall, C. 2014. Warming temperatures and smaller body sizes: synchronous changes in growth of North Sea fishes. Global Change Biology, 20: 1023–1031.

Bhattacharya, C. 1967. A simple method of resolution of a distribution into Gaussian components. Biometrics, 23: 115–135.

Biernacki, C., Celeux, G., and Govaert, G. 2003. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. Computational Statistics & Data Analysis, 41: 561–575.

Cadigan, N. G., and Campana, S. E. 2016. Hierarchical model-based estimation of population growth curves for redfish (*Sebastes mentella* and *Sebastes fasciatus*) off the eastern coast of Canada. ICES Journal of Marine Science, 74: 687–697.

Chang, Y.-J., Sun, C.-L., Chen, Y., and Yeh, S.-Z. 2012. Modelling the growth of crustacean species. Reviews in Fish Biology and Fisheries, 22: 157–187.

Cormon, X., Ernande, B., Kempf, A., Vermard, Y., and Marchal, P. 2016. North Sea saithe *Pollachius virens* growth in relation to food availability, density dependence and temperature. Marine Ecology Progress Series, 542: 141–151.

Dempster, A. P., Laird, N. M., and Rubin, D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological), 39: 1–38.

Denney, N. H., Jennings, S., and Reynolds, J. D. 2002. Life-history correlates of maximum population growth rates in marine fishes. Proceedings of the Royal Society of London B: Biological Sciences, 269: 2229–2237.

Eveson, J. P., Million, J., Sardenne, F., and Le Croizier, G. 2015. Estimating growth of tropical tunas in the Indian Ocean using tag-recapture data and otolith-based age estimates. Fisheries Research, 163: 58–68.

Eveson, J. P., Polacheck, T., and Laslett, G. M. 2007. Consequences of assuming an incorrect error structure in von Bertalanffy growth models: a simulation study. Canadian Journal of Fisheries and Aquatic Sciences, 64: 602–617.

Farina, A., Azevedo, M., Landa, J., Duarte, R., Sampedro, P., Costas, G., Torres, M. *et al.* 2008. Lophius in the world: a synthesis on the common features and life strategies. ICES Journal of Marine Science, 65: 1272–1280.

Fournier, D. A., Hampton, J., and Sibert, J. R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Canadian Journal of Fisheries and Aquatic Sciences, 55: 2105–2116.

Fournier, D. A., Sibert, J. R., Majkowski, J., and Hampton, J. 1990. MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). Canadian Journal of Fisheries and Aquatic Sciences, 47: 301–317.

Gallucci, V. F., and Quinn, T. J. 1979. Reparameterizing, fitting, and testing a simple growth model. Transactions of the American Fisheries Society, 108: 14–25.

Hasselblad, V. 1966. Estimation of parameters for a mixture of normal distributions. Technometrics, 8: 431–444.

Hilborn, R., and Walters, C. J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Kluwer Academic Publishers, Boston/Dordrecht/London.

Hoggarth, D. D. 2006. Stock assessment for fishery management: a framework guide to the stock assessment tools of the fisheries management and science programme. FAO Fisheries Technical Paper. No. 487. Food & Agriculture Organisation, Rome.

ICES. 2011. Report of the anglerfish (*Lophius piscatorius*) illicia and otoliths exchange 2011, ICES Planning Group on Commercial Catch, Discards and Biological Sampling (PGCCDBS).

ICES. 2018a. International Bottom Trawl Survey (2000-2017). http://datras.ices.dk (last accessed 9 May 2018).

ICES. 2018b. DATRAS Survey Descriptions. http://datras.ices.dk (last accessed 9 May 2018).

ICES. 2018c. ICES Advice on Fishing Opportunities, Catch, and Effort Bay of Biscay and the Iberian Coast, Celtic Seas, Greater North Sea, and Oceanic Northeast Atlantic Ecoregions—mon.27.78abd.

Isermann, D. A., and Knight, C. T. 2005. A computer program for age–length keys incorporating age assignment to individual fish. North American Journal of Fisheries Management, 25: 1153–1160.

Jobling, M. 2002. Environmental factors and rates of development and growth. Handbook of fish biology and fisheries, 1: 97–122.

Katsanevakis, S., and Maravelias, C. D. 2008. Modelling fish growth: multi-model inference as a better alternative to a priori using von Bertalanffy equation. Fish and Fisheries, 9: 178–187.

Kilada, R., and Driscoll, J. G. 2017. Age determination in crustaceans: a review. Hydrobiologia, 799: 21–36.

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. M. 2016. TMB: automatic differentiation and Laplace approximation. Journal of Statistical Software, 70: 1–21.

Landa, J., Barrado, J., and Velasco, F. 2013. Age and growth of angler-fish (*Lophius piscatorius*) on the porcupine bank (west of Ireland) based on illicia age estimation. Fisheries Research, 137: 30–40.

Laslett, G. M., Eveson, J. P., and Polacheck, T. 2004. Fitting growth models to length frequency data. ICES Journal of Marine Science, 61: 218–230.

Lloyd-Jones, L. R., Nguyen, H. D., McLachlan, G. J., Sumpton, W., and Wang, Y.-G. 2016. Mixture of time-dependent growth models with an application to blue swimmer crab length-frequency data. Biometrics, 72: 1255–1265.

Macdonald, P., and Du, J. 2011. mixdist: Finite Mixture Distribution Models. R Package Version, 0.5-4. https://CRAN.R-project.org/package=mixdist.

Macdonald, P., and Pitcher, T. 1979. Age-groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. Journal of the Fisheries Board of Canada, 36: 987–1001.

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

McLachlan, G., and Peel, D. 2004. Finite Mixture Models. John Wiley & Sons, New York.

Mildenberger, T. K., Taylor, M. H., and Wolff, M. 2017. TropFishR: an R package for fisheries analysis with length-frequency data. Methods in Ecology and Evolution, 8: 1520–1527.

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

Minte-Vera, C. V., Maunder, M. N., Casselman, J. M., and Campana, S. E. 2016. Growth functions that incorporate the cost of reproduction. Fisheries Research, 180: 31–44.

Minto, C., Hinde, J., and Coelho, R. 2018. Including unsexed individuals in sex-specific growth models. Canadian Journal of Fisheries and Aquatic Sciences, 75: 282–292.

Morrongiello, J. R., and Thresher, R. E. 2015. A statistical framework to explore ontogenetic growth variation among individuals and populations: a marine fish example. Ecological Monographs, 85: 93–115.

Ogle, D. H. 2018. FSA: Fisheries Stock Analysis. R Package Version 0.8.20. https://CRAN.R-project.org/package=FSA.

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

Pauly, D. 1987. A review of the ELEFAN system for analysis of length-freculency data in fish and aquatic invertebrates. *In* Length-based Methods in Fisheries Research. Ed. by D. Pauly and G. R. Morgan. ICLARM Conf Proc, International Center for Living Aquatic Resources Management, Manila, Philippines, and Kuwait Institute for Scientific Research, Safat, Kuwait, pp. 7–34.

Pilling, G. M., Kirkwood, G. P., and Walker, S. G. 2002. An improved method for estimating individual growth variability in fish, and the correlation between von Bertalanffy growth parameters. Canadian Journal of Fisheries and Aquatic Sciences, 59: 424–432.

Pitcher, T., and MacDonald, P. 1973. Two models for seasonal growth in fishes. Journal of Applied Ecology, 10: 599–606.

Punt, A. E., Trinnie, F., Walker, T. I., McGarvey, R., Feenstra, J., Linnane, A., and Hartmann, K. 2013. The performance of a management procedure for rock lobsters, *Jasus edwardsii*, off western Victoria, Australia in the face of non-stationary dynamics. Fisheries Research, 137: 116–128.

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

R Core Team. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Roa-Ureta, R. H. 2010. A likelihood-based model of fish growth with multiple length frequency data. Journal of Agricultural, Biological, and Environmental Statistics, 15: 416–429.

Schnute, J., and Fournier, D. 1980. A new approach to length–frequency analysis: growth structure. Canadian Journal of Fisheries and Aquatic Sciences, 37: 1337–1351.

Smith, M., Fulton, E., Day, R., Shannon, L., and Shin, Y-J. 2015. Ecosystem modelling in the southern Benguela: comparisons of Atlantis, Ecopath with Ecosim, and OSMOSE under fishing scenarios. African Journal of Marine Science, 37: 65–78.

Sparre, P., and Venema, S. C. 1998. Introduction to Tropical Fish Stock Assessment—Part 1: Manual. FAO, Rome.

Taylor, M., and Mildenberger, T. K. 2017. Extending electronic length frequency analysis in R. Fisheries Management and Ecology, 24: 330–338.

Travers, M., Shin, Y.-J., Jennings, S., and Cury, P. 2007. Towards end-to-end models for investigating the effects of climate and fishing in marine ecosystems. Progress in Oceanography, 75: 751–770.

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

Wang, Y.-G. 1999. Estimating equations for parameters in stochastic growth models from tag–recapture data. Biometrics, 55: 900–903.

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

Winsor, C. P. 1932. The Gompertz curve as a growth curve. Proceedings of the National Academy of Sciences, 18: 1–8.

Yao, W. 2015. Label switching and its solutions for frequentist mixture models. Journal of Statistical Computation and Simulation, 85: 1000–1012.

Handling editor: Jan Jaap Poos