

Title

Accounting for temporal and individual variation in the estimation of Von Bertalanffy growth curves

Author details

J. C. Croll (Corresponding author); Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam,
P.O. Box 94240, 1090GE Amsterdam, The Netherlands. J.C.Croll@UvA.nl
T. van Kooten; Wageningen Marine Research, Wageningen University and Research.

Corresponding author

J. C. Croll; Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, P.O. Box 94240, 1090GE
Amsterdam, The Netherlands. J.C.Croll@UvA.nl

Running headline

Von Bertalanffy growth with variation

Data availability

The R package to perform the analysis is available on bitbucket: <https://bitbucket.org/JCroll/VBGfit/>. The data
used for the North Sea plaice example is publicly available on ICES DATRAS data portal: <https://datras.ices.dk>.

Accounting for temporal and individual variation in the estimation of Von Bertalanffy growth curves

Jasper C. Croll and Tobias van Kooten

April 20, 2022

Abstract

Growth and growth limitation are important indicators of density dependence and environmental limitation of populations. Estimating individual growth trajectories is therefore an important aspect of understanding and predicting the life history and dynamics of a population. Variation in individual growth trajectories arises due to variation in the environmental factors limiting individual growth. This environmental limitation can vary over time, between cohorts and between individuals within a cohort. For a complete and accurate understanding of individual growth in a population, it is important to include all these sources of variation. So far, statistical models only accounted for a subset of these factors or required an extensive growth history of individuals. Here we present a novel model describing the growth curves of cohorts in a population. This model is derived from a stochastic form of the Von Bertalanffy growth equation describing individual growth. The model is specifically tailored for use on length-at-age data in which the growth trajectory of an individual is unknown and every individual is only measured once. The presented method can also be used if growth limitation differs strongly between age or length classes. We demonstrate the use of the model for length-at-age data of North Sea plaice (*Pleuronectes platessa*) from the last thirty years. Fitting this model to length-at-age data can provide new insights in the dynamics of the environmental factors limiting individual growth and provides a useful tool for ecological research and management.

keywords: Environmental limitation, Individual variation, Life history, North Sea plaice, Population structure, length-at-age data, Von Bertalanffy growth curve

Introduction

Body length is shown to be an important indicator of many life history traits such as maturation, consumption, mortality and reproduction rates (Kooijman, 2010; Calder, 1984; Peters, 1983; Pauly, 1980; Beverton and Holt, 1959).

39 Estimating the growth trajectory of individuals is therefore a general aspect of understanding the life history and
40 dynamics of a population. The Von Bertalanffy growth equation is one of the most commonly used models to describe
41 growth of individuals. It is already used to describe the growth of a wide range of species (Ramirez et al., 2021;
42 Teleken et al., 2017; Narinc et al., 2017; Kingsley, 1979) and is especially often used for fish (Flinn and Midway, 2021;
43 Lorenzen and Enberg, 2002). The Von Bertalanffy growth equation describes growth in terms of individual energy
44 assimilation and catabolism rates (Bertalanffy, 1938). In the resulting growth curve, individuals grow towards an
45 asymptotic length at which the catabolism rate is equal to the assimilation rate and no energy is available for growth.
46 If we follow Dynamic Energy Budget theory (Kooijman, 2010) and assume that assimilation scales with surface area
47 while catabolism scales with body volume, we obtain the most commonly used form of the Von Bertalanffy growth
48 equation:

$$\frac{d\ell_{(t,a,i)}}{dt} = r_B (f_t \ell_\infty - \ell_{(t,a,i)}) \quad (1)$$

49 Herein we indicate the length of individual i at age a and time t with $\ell_{(t,a,i)}$, which emphasizes that equation (1)
50 describes the growth rate of a single individual. In this equation, individual growth is proportional to the difference
51 between the asymptotic length ($f_t \ell_\infty$) and the current length, scaled with the Von Bertalanffy growth rate scalar
52 (r_B). Following Dynamic Energy Budget theory (Kooijman, 2010), the asymptotic length consists of an maximum
53 asymptotic length (ℓ_∞) scaled by the limitation of the environment through assimilation (f_t). It is commonly assumed
54 that the asymptotic length is the only parameter in the Von Bertalanffy growth equation which depends on the
55 environment (Kooijman, 2010; Lorenzen and Enberg, 2002). As such, the Von Bertalanffy growth equation provides
56 an opportunity to estimate the environmental limitation on the individual growth rate.

57 Research on the dynamics of individual growth is generally based on one of two types of data containing age and
58 length measurements of individuals. The first type of data contains multiple measurements of the same individual,
59 for example obtained through controlled experiments, mark-recapture methods or back calculation from otoliths or
60 year-rings. This type of data generally allows for extensive correction for variation between individuals and cohorts
61 because the growth history of individuals is known (Vincenzi et al., 2016, 2014; Graaf and Prein, 2005; Rafail, 1973).
62 Often however, such rich individual-level data is unavailable, because individuals cannot be tagged or retrieved or
63 because back calculation of otoliths and year-rings is often imprecise (Eveson et al., 2007). Even if these methods are
64 successful they only result in a relative relationship between age and length. Much more common is the second type
65 of data, which only contains a single measurement per individual. To obtain this kind of data, individual only have
66 to be measured once and the age of individuals can be determined based on hard or internal body structures such as
67 year-rings, scales, bones, teeth, and chemical composition. This is a common method for fish (Maceina et al., 2007),
68 amphibians (Smirina, 1994), reptiles (Castanet, 1994), mammals (Read et al., 2018) and insects (Robson and Crozier,

69 2009). In this study we focus on the estimation of growth curves and variation herein based on data consisting of a
70 single observation per individual.

71 In data with a single age and length observation per individual, the growth history of individuals is unknown and
72 it is difficult to deal with the different overlapping sources of variation in the length individuals have at a specified age.
73 In addition to variation between measurements due to sampling errors (Piner et al., 2016; Taylor et al., 2005), variation
74 in the growth rate can be separated into variation as a result of changes in the environment over time, variation in
75 the growth history of cohorts and variation between individuals within a cohort. So far, statistical methods dealing
76 with single observations of individual ages and lengths only deal with a subset of these sources of variation.

77 Variation due to changes in the environment over time is most likely to affect the asymptotic length in the Von
78 Bertalanffy growth equation and therefore affect all cohorts at a given time equally. This can be used to estimate
79 the effect of an environmental factor on the growth of individuals. This is generally done by directly substituting the
80 asymptotic length in the Von Bertalanffy growth equation with a linear dependency on the environmental factor of
81 interest (Graaf and Prein, 2005; Lorenzen and Enberg, 2002; Lorenzen, 1996; Cloern and Nichols, 1978). Although
82 this can be useful to prove a general relationship between the growth rate and an environmental factor, the a priori
83 assumption of linearity is questionable.

84 Variation between cohorts arises due to differences in the growth history of cohorts. It is evident that cohorts in
85 a given year differ in length due to the difference in age, but the length at a given age is likely to vary over time
86 as well. This variation between cohorts might arise due to variation in the length at birth, but might also occur
87 because cohorts lived at different times and therefore differ in the experienced environment (He and Bence, 2007;
88 Wang and Thomas, 1995). A common way to correct for the growth history of individuals is to consider the average
89 growth increment between two time points, instead of the actual length-at-age (Wang and Thomas, 1995; Lipinski
90 and Roeleveld, 1990; Rafail, 1973). As we consider datasets that consist of independent length-at-age observations
91 throughout years, this method can only be applied to the average length-at-age in every sampling instance and as such
92 neglects individual variation in length-at-age and environmental limitation. In addition, this has been shown to yield
93 less accurate estimates of the Von Bertalanffy growth parameters with a larger uncertainty (Vaughan and Kanciruk,
94 1982).

95 Similar to variation between cohorts, variation in length-at-age within cohorts arises due to differences in length
96 at birth and differences in the experienced environment between individuals. Although individuals in the same cohort
97 are not separated in time, they might be separated spatially or due to other ecological factors leading to variation
98 in the experienced environment. It is relatively easy to account for this type of variation with datasets in which the
99 entire growth history of an individual is recorded. In contrast, the only option to account for individual variation in
100 datasets with single measurements is by considering the length-at-age at the population level as a distribution rather

101 than a single value (Eveson et al., 2007; Pilling et al., 2002; Prajeshy and Venugopalan, 1999).

102 The different sources of variation in the length-at-age are entangled due to the auto-regressive nature of individual
 103 growth processes in which the current growth rate depends on the growth history of an individual. As a consequence,
 104 variation in growth arises between individuals and cohorts and could fluctuate over time. All these sources of variation
 105 should be considered to obtain an accurate estimate of the Von Bertalaffy growth parameters of a specific species, even
 106 if we are only interested in one of the sources or variation or average growth parameters. such a method is currently
 107 unavailable. To fill this gap, we derive a model that describes the length distribution at a given age for every cohort,
 108 which can be used for datasets with a single length and age observation per individual. Because we derived this
 109 model from a stochastic version of the Von Bertanffy growth equation for single individuals (eq. (1)) it simultaneously
 110 includes variation due to changes over time, variation due to differences in the growth history of cohorts and variation
 111 between individuals within a cohort. Here, we derive the model and apply it to length-at-age data of North Sea plaice
 112 (*Pleuronectes platessa*).

113 Methods

114 Model formulation

115 We start with the equation describing the growth trajectory of a single individual (eq. 1). Individuals will differ in the
 116 experienced environmental limitation and this limitation might vary over time. We therefore assume that the limiting
 117 effect of the environment at a given point in time follows a gaussian distribution of which the mean (μ_t) and variance
 118 (σ_t^2) are allowed to vary over time:

$$f_t \sim \mathcal{N}(\mu_t, \sigma_t^2) \quad (2)$$

119 By substituting this distribution in equation (1) we obtain a stochastic differential equation describing the growth of
 120 an individual born at time T_b . We can solve this equation by separation of variables and integration:

$$\begin{aligned} d\ell_{(t,a,i)} &= r_B (\mu_t \ell_\infty - \ell_{(t,a,i)}) dt + r_B \sigma_t \ell_\infty dW_t \\ \ell_{(T,a,i)} &= \ell_{(T_b,0,i)} e^{-r_B(T-T_b)} + \int_{T_b}^T k \mu_t \ell_\infty e^{-r_B(T-t)} dt + \int_{T_b}^T r_B \sigma_t \ell_\infty e^{-r_B(T-t)} dW_t \end{aligned} \quad (3)$$

121 The parameter $\ell_{(T_b,0,i)}$ represents the length at birth of an individual. In addition, W_t represents a Wiener process,
 122 which describes the outcome of a continuous process with independent gaussian increments ($W_{t+u} - W_t \sim \mathcal{N}(0, u)$).
 123 The integrals in this expression cannot be solved explicitly because the dynamics of the mean and variance of the
 124 environmental limitation (μ_t, σ_t^2) are not defined. If the environmental limitation was constant over time and space
 125 ($\mu_t = \mu, \sigma_t^2 = 0$), individuals would follow a Von Bertalanffy growth curve towards a constant asymptotic length.

126 In this method we consider datasets in which every individual is only measured once. In the ideal situation,
 127 these individuals are selected randomly from the population. In such datasets, it is not possible to follow the growth
 128 trajectory of a single individual and fit the derived growth curve on single individuals. Instead we describe the
 129 distribution of the length-at-age for a cohort. Because we assumed that the environmental limitation of growth follows
 130 a gaussian distribution, the length of individuals in a given cohort at time T follows a gaussian distribution as well.
 131 Because the expected value of the Wiener process is equal to zero, we can derive an expression for the expected mean
 132 length at time T of a cohort born at time T_b .

$$\begin{aligned} E[\ell_{(T,a)}] &= E[\ell_{(T_b,0,i)} e^{-r_B(T-T_b)}] + E\left[\int_{T_b}^T r_B \mu_t \ell_\infty e^{-r_B(T-t)} dt\right] + E\left[\int_{T_b}^T r_B \ell_\infty \sigma_t e^{-r_B(T-t)} dW_t\right] \\ &= E[\ell_{(T_b,a)}] e^{-r_B(T-T_b)} + \int_{T_b}^T r_B \mu_t \ell_\infty e^{-r_B(T-t)} dt \end{aligned} \quad (4)$$

133 We omitted the indices referring to single individuals in the expression of the expected value of the length-at-age
 134 ($E[\ell_{(T,a)}]$), to make clear that this expected value is a statistic of the length-at-age distribution of a cohort rather than
 135 the length of single individuals. By using equations (3) and (4) and applying Ito's isometry rule we can also derive an
 136 expression for the expected variance in length at time T for a cohort born at time T_b :

$$\begin{aligned} V[\ell_{(T,a)}] &= E[(\ell_{(T,a,i)} - E[\ell_{(T,a)}])^2] = E\left[\left(\int_{T_b}^T r_B \sigma_t \ell_\infty e^{-r_B(T-t)} dW_t\right)^2\right] \\ &= E\left[\int_{T_b}^T r_B^2 \sigma_t^2 \ell_\infty^2 e^{-2r_B(T-t)} dt\right] = \int_{T_b}^T r_B^2 \sigma_t^2 \ell_\infty^2 e^{-2r_B(T-t)} dt \end{aligned} \quad (5)$$

137 We assume that samples are taken with an approximately constant time interval and therefore discretise the
 138 equations characterising the length distribution of a given cohort at time T . Under this assumption, the equations
 139 become independent of the length distribution at birth and can be applied without knowledge about the full growth
 140 history of a cohort.

$$E[\ell_{(T+1,a+1)}] = E[\ell_{(T_b,0)}] e^{-r_B(T-T_b+1)} + \int_{T_b}^{T+1} r_B \mu_t \ell_\infty e^{-r_B(T+1-t)} dt \quad (6a)$$

$$\begin{aligned} &= E[\ell_{(T,a)}] e^{-r_B} + \int_T^{T+1} r_B \mu_t \ell_\infty e^{-r_B(T+1-t)} dt \\ V[\ell_{(T+1,a+1)}] &= \int_{T_b}^{T+1} r_B^2 \sigma_t^2 \ell_\infty^2 e^{-2r_B(T+1-t)} dt \\ &= V[\ell_{(T,a)}] e^{-2r_B} + \int_T^{T+1} r_B^2 \sigma_t^2 \ell_\infty^2 e^{-2r_B(T+1-t)} dt \end{aligned} \quad (6b)$$

141 To make these equations usable, we need to make assumptions about the dynamics of the mean and the variance
 142 of the environmental limitation (μ_t and σ_t^2). These quantities only appear within the integral from time T up to time
 143 $T + 1$. We therefore only have to make assumptions about the mean and variance of the environmental limitation
 144 between consecutive timepoints or measurements. We assume that the mean and variance of the environmental
 145 limitation between times T and $T + 1$ are well approximated by the average value of these quantities in the given

time interval ($\bar{\mu}_T$ and $\bar{\sigma}_T^2$). Under these assumptions, the model will approximate the growth dynamics if the interval between measurements becomes small relative to the average lifetime of an individual. Because we substitute the mean and variance of the environmental limitation by the average of these quantities over a time interval, they become independent of time in the domain of integration and we can solve the integrals in equation (6), which results in the final form of our model:

$$E[\ell_{(T+1,a+1)}] = E[\ell_{(T,a)}] e^{-r_B} + \bar{\mu}_T \ell_\infty (1 - e^{-r_B}) \quad (7a)$$

$$V[\ell_{(T+1,a+1)}] = V[\ell_{(T,a)}] e^{-2r_B} + \frac{1}{2} r_B \bar{\sigma}_T^2 \ell_\infty^2 (1 - e^{-2r_B}) \quad (7b)$$

Interesting to note from this formulation is that the variance in environmental limitation over a given period ($\bar{\sigma}_T^2$) has the same unit as the time constant (T). This arises because we model the length of an individual as a Brownian processes which is a process with random increments. The variance of a Brownian process increases due to the random nature of the process and therefore depends on the length of the time between consecutive measurements in our model. In other words, the total variance in environmental limitation experienced by an individual increases (decreases) if the actual variation of the environment increases (decreases) or the individual experiences the environment for a longer (shorter) period of time.

To obtain a time independent measurement of the environmental variation, we can consider the long term asymptotic variation in individual length ($V[\ell_{(T,\infty)}]$). This represents the variation in length that individuals would have after spending an infinity large time in an environment with a given amount of variation in growth limitation ($\bar{\sigma}_T^2$). At this asymptotic variation in length, the loss of variation in length due to growth ($V[\ell_{(T,a)}] (1 - e^{-2r_B})$) is equal to the gain in variation in length due to variation in the environment ($\frac{1}{2} r_B \bar{\sigma}_T^2 \ell_\infty^2 (1 - e^{-2r_B})$). In other words, when a cohort reaches the asymptotic variation in length, the variation in length of a cohort does not change any further over time. From equation (7b) we can therefore derive the expression of the long term asymptotic variance in length:

$$V[\ell_{(T,\infty)}] = \frac{1}{2} r_B \bar{\sigma}_T^2 \ell_\infty^2 \quad (8)$$

Model application

The model proposed in equation (7) predicts an independent Gaussian length distribution for every cohort at every discrete age and time value. Therefore the model can be fitted to datasets containing pairs of age and length measurements using maximum likelihood estimates. The best results are obtained if measurements represent a random sample of the population. This is especially important for individuals in the same age class and year. Assigning weights to the measurements allows to correct for biases in the sample, if biases are known. In addition, sample instances should approximately be evenly distributed in time and individual ages and cohorts should be characterised on the same discrete scale as sample instances. For example, measurements could be taken yearly on randomly selected individuals

at a specified date. The age of individuals is consequently measured in year classes and individuals born between two measurements belong to the same cohort. Fitting the model described by equation (7) to a dataset with pairs of length and age measurements is done by optimising the log likelihood through altering the value of the Von Bertalanffy scalar (r_B), the length distribution at the youngest age at every time point ($E[\ell_{T,a_{min}}], V[\ell_{T,a_{min}}]$), the length distribution of all other cohorts at the first time point ($E[\ell_{T_{min},a}], V[\ell_{T_{min},a}]$) and the distribution of the environmental limitation between all time points ($\bar{\mu}_T, \bar{\sigma}_T^2$) (Table 1). In the proposed model, the mean and variance of the environmental limitation always occur as a product with the maximum asymptotic length ($\bar{\mu}_T \ell_\infty, \bar{\sigma}_T^2 \ell_\infty^2$). Therefore the maximum asymptotic length cannot be estimated separately with this method and is incorporated as a species specific scalar of the environmental limitation. We provided an R-package (Croll, 2022) that includes a procedure for fitting the model to a dataset with pairs of age and length measurements using maximum likelihood optimisation through optimisation methods available in the NLOptR-package (Johnson, 2022).

The R-package for fitting the described Von Bertalanffy growth model contains some additional features to tailor the model to specific populations. The first feature deals with the dynamics of the mean length at age when the environmental limitation is very variable. In the model described in equation (7) the mean length of a cohort ($E[\ell_{T,a}]$) decreases if it exceeds the asymptotic length at some time step ($\bar{\mu}_T \ell_\infty$). Some species do not decrease in length and it might therefore be unrealistic for the average length of a cohort to decrease. For these populations the package also includes a version of the model in which the average length of a cohort does not decrease if the average cohort length exceeds the maximum asymptotic length. This version of the model should be used with care and only with reasonable arguments, because this method inflates the impact of small and younger cohorts on the estimated environmental limitation. In any case we advice to first fit the model without this additional assumption, to check whether this indeed predicts large decreases in mean length of some cohorts.

The second extension available in the R-package deals with differences in environmental limitation between age and length classes. Differences in environmental limitation between age or length classes can arise if age or length classes show spatial segregation or differ in diet. The model allows to specify age or length classes and estimates separate means and variances in environmental limitation ($\bar{\mu}_{T,c}, \bar{\sigma}_{T,c}^2$) for every age or length class at every time step. It is important to note that this extension of the model only uses the mean length of a cohort to identify the length class and therefore all individuals in a cohort are always placed in the same length class. In addition, the number of observations per class decreases with an increase in the number of age or length classes. The incorporation of age or length classes can therefore make the model fit less accurate if there are no true differences in the environmental limitation of the selected classes.

Table 1: Description of the parameters that are estimated during the model fitting procedure.

Parameter	Description	Type	Number of parameters
n_T	number of sampling instances	discrete	0 (fixed value)
n_a	number of sampled age classes	discrete	0 (fixed value)
T_{min}	Time of the first sampling instance	discrete	0 (fixed value)
a_{min}	First age class in the dataset	discrete	0 (fixed value)
r_B	Von Bertalanffy growth rate scalar	continuous	1
$E[\ell_{(T,a_{min})}]$	Expected mean length at the first age class	continuous	n_T
$V[\ell_{(T,a_{min})}]$	Variance in length at the first age class	continuous	n_T
$E[\ell_{(T_{min},a)}]$	Expected mean length at the first sampling instance	continuous	$n_a - 1$
$V[\ell_{(T_{min},a)}]$	Expected variance in length at the first sampling instance	continuous	$n_a - 1$
$\bar{\mu}_T \ell_\infty$	Mean asymptotic length	continuous	$n_T - 1$
$\bar{\sigma}_T^2 \ell_\infty$	Variance in asymptotic length	continuous	$n_T - 1$

Application to North Sea plaice

To illustrate the use of the proposed model we fit the model to a dataset with age and length measurements of plaice (*Pleuronectes platessa*) obtained from the Beam Trawl survey (BTS). This survey is designed to monitor plaice in the North Sea and is consistently conducted in the third quarter (July to September) from 1990 onwards. The length of individuals is measured with at least 5 mm accuracy and the age of sampled individuals is obtained through otolith readings. We downloaded the datasets with individual ages and lengths recorded during the third quarter of 1990 to 2021 from the online ICES DATRAS data portal on the 1st of November 2021 (ICES, 2021). The age-length observations were weighted by the inverse of the catch per unit effort (CPUE) of the observed length. The CPUE per length indicates the probability that an individual of a given length is caught in the survey of a given year. Weighting the observations with the inverse of the CPUE corrects for any factor that affects the catch probability of a given length in a given year. After weighing of the samples, all lengths approximately had the same contribution to the dataset and visual inspection confirmed that the data approximated the assumption that length at age in a given year follows a Gaussian distribution (judged by eye, fig. A3).

Starting values for the expected length at the first age in the dataset and the expected asymptotic length and the growth scalar were estimated by fitting a Von Bertalanffy growth model without considering differences between years and cohorts ($E[\ell_{(T,a_{min})}] = 120mm$, $\bar{\mu}_t \ell_\infty = 380mm$ and $r_B = 0.303y^{-1}$). Starting values for the variance in length at the lowest age class and the variance in asymptotic length were set to the variance in the youngest en oldest age class respectively ($V[\ell_{(T,a_{min})}] = 801mm^2y$ and $\bar{\sigma}_t^2 \ell_\infty^2 = 5789mm^2y$).

221 We used the Sbplx algorithm of the NloptR package (Johnson, 2022), which is a variant of the Nelder-Mead
 222 optimisation method, with a relative tolerance of 10^{-10} to optimise the likelihood of our model. The optimisation was
 223 performed using a log-transformed parameter space to account for the magnitudinal difference between parameters.
 224 For comparison we fitted two versions of the model. In the first version the environmental limitation was constant
 225 over years and therefore the mean and variance of the asymptotic length were estimated as a single parameter. In
 226 the second version the environmental limitation was allowed to vary between years and the mean and variance of the
 227 asymptotic length were estimated separately for every year.

228 To assess the robustness of the model with yearly varying asymptotic length, we used a jackknife approach in
 229 which we repeated the analysis 31 times with data from one entire year omitted every time. This shows the impact
 230 of the samples from a given year on the model fit and gives an indication of the robustness of the method to years in
 231 which no data could be collected. Lastly, we demonstrate the use of separate age groups with different environmental
 232 limitation in the model by splitting the plaice population in three ecological groups by age.

233 We used a estimation of the maximum asymptotic length ($\ell_{\infty} = 780$) estimated by Van der Veer et al. (2001)
 234 and scaled the estimated mean and variance in asymptotic length with this value to obtain the mean and variance in
 235 environmental limitation $(\bar{\mu}_t, \bar{\sigma}_t^2)$.

236 Results

237 A model with a constant environmental limitation and a model with yearly varying environmental limitation were
 238 fitted to a length-at-age dataset for North sea plaice. The model with yearly varying mean and variance of the
 239 environmental limitation fitted the data better compared to the model with constant mean and variance of the
 240 environmental limitation (AIC of respectively 16076750 and 16048031, likelihood ratio test: $p < 0.001$). This suggests
 241 that the mean and variance of the environmental limitation are likely to fluctuate between years. More precisely, the
 242 model with yearly varying environmental limitation suggests a weak downward trend in this limitation, indicating
 243 that the environmental limitation became stronger over time (Fig. 1, solid line). The estimated variance in the
 244 environmental limitation is slightly larger if the environmental limitation is fixed compared to the model in which the
 245 environmental limitation is allowed to fluctuate ($0.0335y$ and on average $0.0268y$ respectively). This overestimation
 246 of the variance in environmental limitation arises because the fixed limitation model accounts for the variation in
 247 the asymptotic length between years in addition to the variation in asymptotic length within a year. As expected,
 248 the models also differ slightly in the estimated parameters defining the length distribution at the youngest age. The
 249 estimates of the Von Bertalanffy growth scalar of the models ($0.2553y^{-1}$ and $0.2977y^{-1}$ respectively) are relatively
 250 close to estimates based on individual energy expenses ($0.2955y^{-1}$, (Van der Veer et al., 2001)).

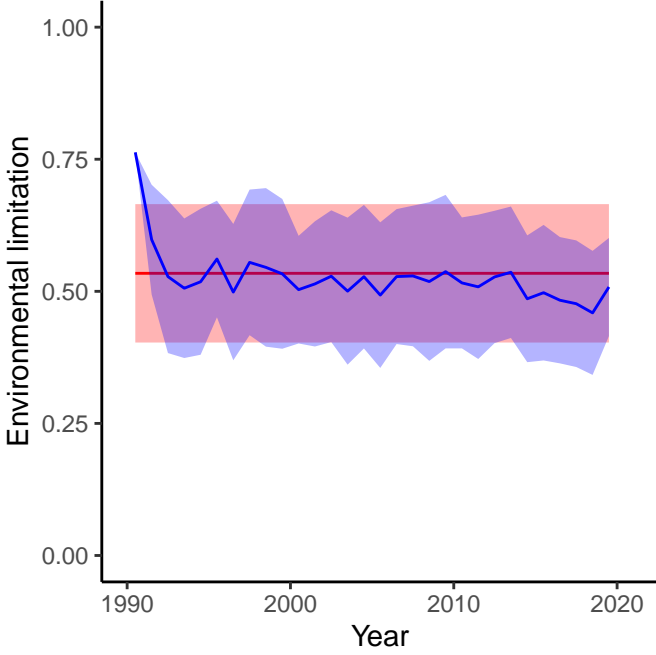


Figure 1: Fitted environmental limitation ($\bar{\mu}_T$) for the model with constant environmental limitation (red) and the model with yearly varying environmental limitation (blue). Shaded areas indicate the mean plus or minus the two times standard deviation derived from the estimated asymptotic variances ($\frac{1}{2}r_B\bar{\sigma}_T^2$).

The estimated model parameters lead to predictions of the length-at-age distribution for every cohort, which differ most strongly for the older age classes (Fig. 2). On visual inspection both the model with a constant environmental limitation and a yearly varying environment limitation appear to fit the datapoints well (Fig. 2c-f). Note that the expected length of individuals in a cohort can shrink in the model with yearly varying environmental limitation. This occurs if the estimated mean asymptotic length falls below the expected length of individuals in a cohort ($\bar{\mu}_T\ell_\infty < E[\ell_{(T,a)}]$). While such a decrease can realistically occur, it is sometimes a biologically impossible result. Repeating the analysis on this dataset with the restriction that the expected length of a cohort can not decrease, yields very similar results (not shown). Nonetheless this additional restriction should be handled with care, because early tests on simulated data showed that this restriction makes the model more dependent on the data points of young age classes.

To demonstrate the robustness of the model we used a jackknife approach in which we repeated the analysis with the samples from one year omitted (Fig. 3). It is not unlikely that actual datasets will contain years for which there is no data, for example due to sampling problems. This analysis showed that missing samples mainly affect the estimate of the mean environmental limitation in the time step directly before and directly after the sample instance with missing data. At one of these time steps the mean environmental limitation will be overestimated while it will be underestimated in the other time step. In addition, it seems that this over- and underestimation of the mean asymptotic length becomes larger towards the start and end of the time period included in the model. A possible cause for this pattern is that these time steps include cohorts which partly fall outside the time-period covered by

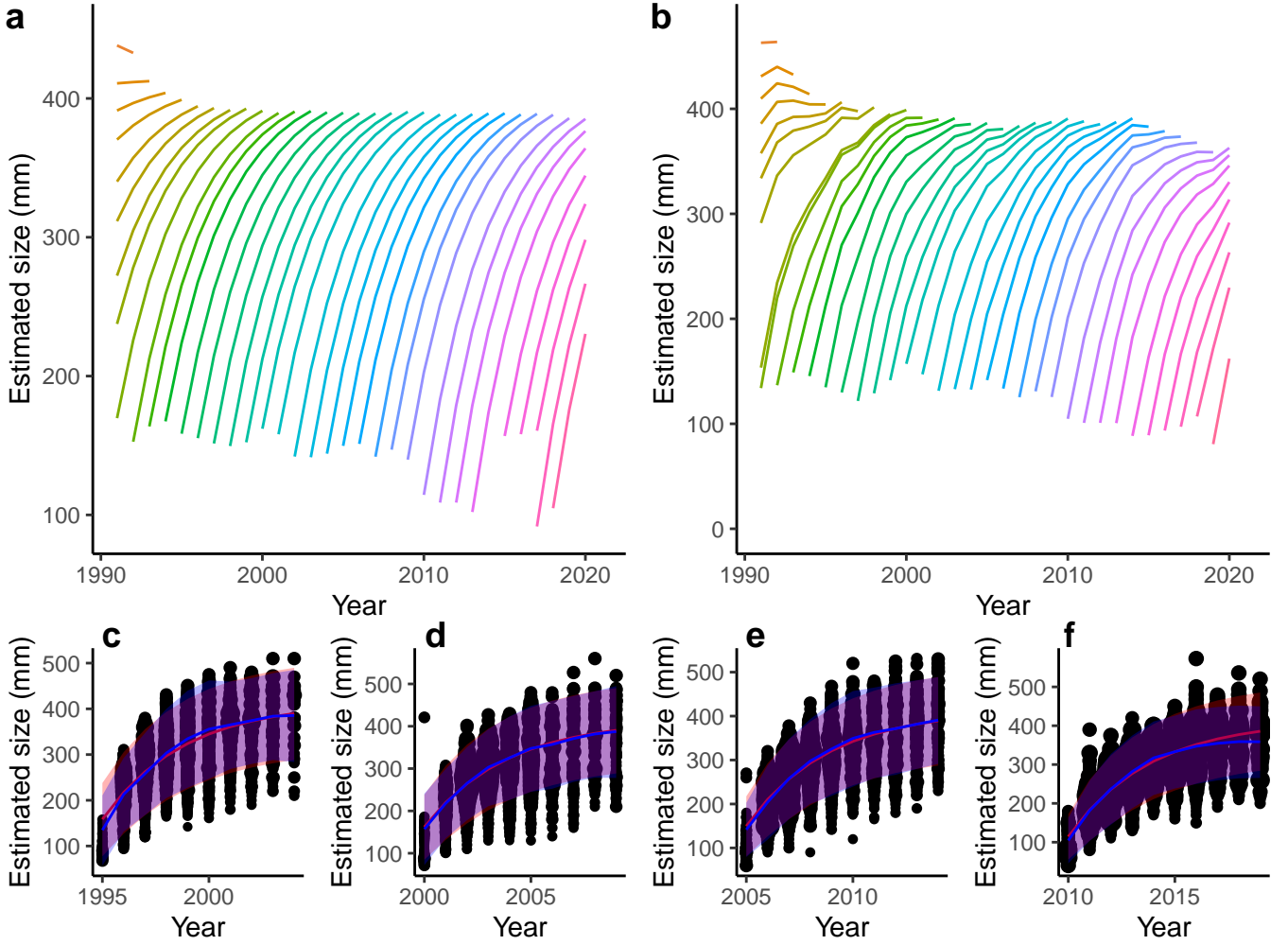


Figure 2: Expected value of the length-at-age ($E[\ell_{(T,a)}]$) for the model with a constant environmental limitation (a) and a yearly varying environmental limitation (b). Line colours correspond to the year of birth of the cohort. The expected length-at-age for the cohorts born in the years 1995 (c), 2000 (d), 2005 (e) and 2010 (f) are plotted separately for the model with a constant environmental limitation (red) and the model with a yearly varying environmental limitation (blue), together with the data points corresponding to the specific cohort. Shaded areas indicate the expected length plus or minus two times the standard deviation from the estimated length distribution. The length of the data points indicates the weighted number of observations of a specific age-length combination in the given cohort ranging from 1 to 60 times.

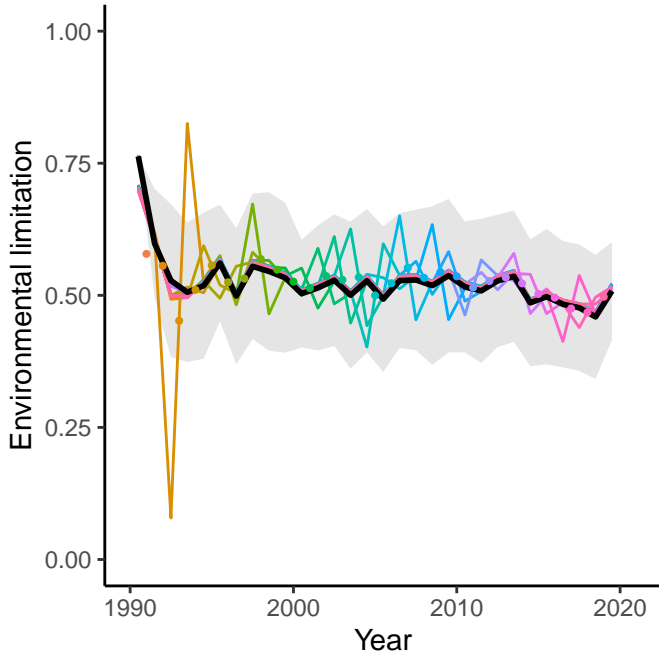


Figure 3: Fitted environmental limitation ($\bar{\mu}_T$) as predicted during a jackknife approach. Colours correspond to different model fits. For every model fit, the data in the year indicated by the dot is omitted from the analysis. The black line is the estimated mean environmental limitation without omitted data and shaded areas indicate the mean plus or minus the two times standard deviation derived from the estimated asymptotic variances without omitted data ($\frac{1}{2}r_B\bar{\sigma}_T^2$).

the data and therefore are estimated on a restricted number of ages. This could make estimates of the growth curves of these cohorts more vulnerable to missing data, which is reflected in the larger over- and underestimations of the asymptotic length in the years these cohorts are in. Indeed, the early years included in the analysis include significantly less observation than later years. Lastly, it is clear that the over- and underestimation of the mean asymptotic length due to omitted data is small compared to the variation between individuals within a cohort.

Our estimation method can also be used on populations which consist of separate ecological groups with, for example, differences in diet or habitat use. We demonstrate this using plaice, the distribution of which has been shown to shift away from the coast with increasing length or age (Braber and De Groot, 1973; Basimi and Grove, 1985; Rijnsdorp and Vingerhoed, 2001). We divide the plaice population into three arbitrary age groups to represent this spatial shift with age, respectively a group up to 3 years old, a group from 4 to 7 years old and a group with individuals over 7 years old. The model fit yields an estimate of the mean and the variance of the environmental limitation for every year and group (Fig. 4). In this way the environmental limitation can be estimated if it differs between age or length groups. If the chosen division in age groups is an accurate representation of the ecology of North Sea plaice, this model should fit the data better compared to the model using a common feeding level for all groups. In addition, it is expected that the estimated variance of the environmental limitation is lower in the model

with groups compared to the model without groups because individuals within a group are logically expected to be more similar than the individuals in the entire population. The model with three age groups does indeed fit the data better compared to the model without age groups (AIC of respectively 15982917 and 16048036). Despite a similar trend, year to year changes in some years differ substantially between length groups, both in magnitude and direction. Such differences could indicate relevant ecological differences between the groups. The average estimated variance in environmental limitation for the youngest age group (on average $0.0217y$) is smaller compared to the average estimated variance in environmental limitation in the model without age groups (on average $0.0268y$). This might suggest that the environmental limitation of individuals in the youngest age group is more similar to the environmental limitation of individuals in the same age group compared to the environmental limitation of individuals in other age groups. This could explain why the model with three age groups fits the data better compared to the model without age groups. In contrast, the average estimated variance in environmental limitation for the two oldest age groups (on average $0.0399y$ for age group 4-7 years and $0.0801y$ for age group 7-10 years) is larger compared to the average estimated variance in environmental limitation in the model without age groups (on average $0.0268y$). This might suggest that the environmental limitation of some individuals in the two oldest age groups is more similar to the environmental limitation of individuals in other age groups compared to the environmental limitation of individuals in the same age groups. The division of the population in age groups was made under the assumption that individuals within an age group are more similar in the experienced environmental limitation compared to individuals in another age group. The increase in variance in environmental limitation in the two oldest age groups might therefore be an indication that the incorporation of the two oldest age groups in the model is not an accurate representation of the ecology of North Sea plaice. The possibility to define age or length groups is used here only to illustrate how the method can be used as a research tool to explore differences in growth limitation across a population's life history.

Discussion

We presented a new method to estimate Von Bertalanffy growth parameters from datasets with pairs of age and length measurements and provide an R package called VBGfit (Croll, 2022) to apply this method. The method is based on a model that describes the length distribution of cohorts in a population under the assumption that cohorts partly overlap in time and experience a fluctuating environment (eq. (7)). The model is derived from a stochastic differential equation describing the growth of a single individual in a fluctuating environment (eq. (1)) and therefore accounts for variation due to changes in the environment over time, variation in the growth history of cohorts and variation between individuals within a cohort. Because the model is described in a discretised form, it is easy to fit on pairs of length and age measurements taken with a regular interval, which is one of the most common forms of data on

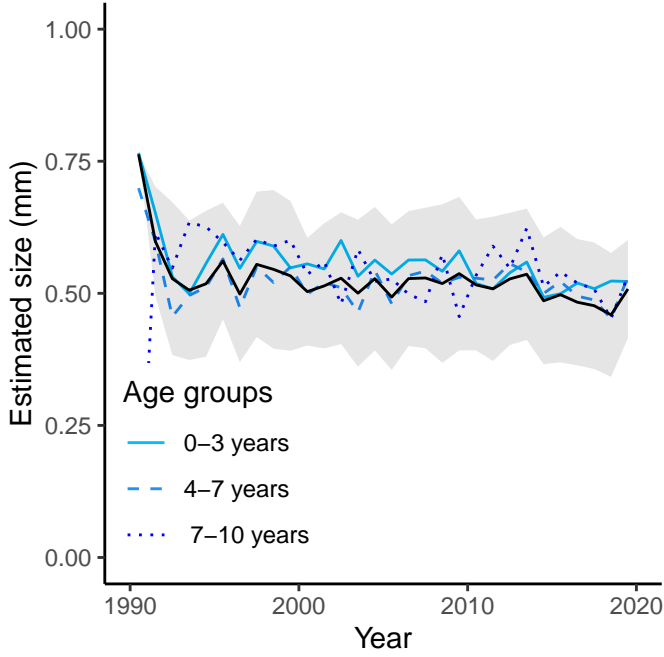


Figure 4: Estimated environmental limitation if three groups with separate environmental limitations are defined based on age. The black line and Shaded areas indicate the mean plus or minus the two times standard deviation derived from the estimated asymptotic variances if all ages are part of the same group ($\frac{1}{2}r_B\bar{\sigma}_T^2$).

population structure (Eveson et al., 2007).

Our model makes several assumptions about the underlying population structure to obtain a model applicable to datasets with random observation pairs of individual lengths and ages. First of all, we assume individuals follow a Von Bertalanffy growth curve in which only the asymptotic length fluctuates over time and between individuals (eq. (1)). This is the most common and first proposed form of the Von Bertalanffy growth equation (Bertalanffy, 1938). Nonetheless, it is sometimes assumed that both the asymptotic length and the Von Bertalanffy growth rate scalar fluctuate (Eveson et al., 2007; Pilling et al., 2002). It has been shown that estimates of the asymptotic length and the Von Bertalanffy growth rate scalar are strongly correlated if both are allowed to fluctuate. Due to this correlation it might be difficult to obtain correct parameter estimates, because different sets of parameters are likely to fit the dataset equally well (Eveson et al., 2007; Pilling et al., 2002). In addition, it has been shown that only the asymptotic length or the Von Bertalanffy growth rate scalar has to fluctuate to obtain a very good prediction of the population structure and an accurate estimate of the environmental limitations, when data consists of independently observed pairs of individual age and length (Eveson et al., 2007). We therefore chose to only make the asymptotic length dependent on the environmental limitation, as this has the most comprehensive substantiation in energetic theory (Kooijman, 2010). Secondly we assume that the dynamics of the environmental limitation between two measurements can be described accurately by the average environmental limitation in this period. This is a very convenient assumption borne from the discrete nature of most datasets with length-at-age data. Nonetheless, it is possible to substitute a more complex,

time dependent formulation for the environmental limitation in equation (6) and work out the more complex model through integration. This would lead to a more specific and less generally applicable form of the model. Thirdly we assume that the environmental limitation experienced by an individual at a given moment is drawn from a gaussian distribution. The central limit theorem states that if a variable is influenced by many additive random factors, it will approach a gaussian distribution. As the environmental limitation emerges from a complex ecological, chemical or physical system, it is likely to be influenced by many random factors and therefore is likely to approach a gaussian distribution. In conclusion, most of the assumptions in this method are made to ensure that the model is as generic as possible but still applicable to the currently available datasets with length-at-age data.

Just like other methods, our model assumes that individual length is normally distributed in a cohort and therefore in the obtained samples. Deviations from this normal distribution can occur for example due to sampling biases or a link between individual mortality rate and individual length. The value of mean environmental limitation is slightly overestimated if larger individuals are more likely to contribute to the dataset, while the value of the mean environmental limitation is underestimated if smaller individuals are more likely to contribute to the dataset. Model fits on simulated data show that the effects of length bias in the dataset might be relatively small (supplementary materials). Nonetheless, it is important to correct for skewness in the individual length distribution in the data when possible. One way to do so, is to add relative weights to the samples. In our example with North sea plaice we weighted the samples by the inverse of the catch per unit effort (CPUE) per length. The CPUE is a measure of the relative presence of a length class in the dataset in a given year. In this way, we corrected for the impact of length on the catchability of an individual, which can arise for example due very strong length specific mortality or harvesting probabilities. This resulted in a dataset in which individual length approximates a normal distribution of every age class in every year. (Fig. A3)

A novel and very important aspect of our method is that it accounts for variation caused by environmental changes over time, variation between the growth history of cohorts and variation between individuals within a cohort simultaneously. Earlier methods only account for variation due to changes over time by fitting a growth curve separately for every sampling instance, or only account for the growth history of a cohort by fitting a growth curve separately for every cohort. Similarly, more recent methods only accounted for variation between individuals (Vincenzi et al., 2014; He and Bence, 2007; Pilling et al., 2002; Prajneshy and Venugopalan, 1999; Wang and Thomas, 1995; Rafail, 1973) or variation caused by changes through time (Lorenzen and Enberg, 2002; Lorenzen, 1996; Cloern and Nichols, 1978). Due to the autoregressive nature of individual growth rates, these sources of variation are strongly intertwined and should not be considered separately. We account for this by fitting the Von Bertalanffy growth equation for all cohorts and sampling instances simultaneously. Because we derived our model from a stochastic differential equation describing individual growth, our model also accounts for variation in environmental limitation between individuals.

362 In addition, we show that the variation between individuals is overestimated if a model does not account for changes
363 is environmental limitation over time.

364 Because our model simultaneously accounts for variation between individuals and cohorts and allows variation
365 from the environment to fluctuate over time, the model can be used for a wide range of applications. First of
366 all, the estimated mean and variance of the asymptotic length estimated by our model can be used as a summary
367 statistic for environmental limitation. Growth of individuals is likely limited by numerous factors, which are often
368 unknown. Our method offers a summary statistic for the cumulative distribution of all these factors. Our method is
369 especially appropriate to estimate individual limitation in growth due to limitation through food availability. General
370 theory about individual energy allocation links the asymptotic length of this Von Bertalanffy growth equation to the
371 energy ingestion by individuals (Kooijman, 2010; Bertalanffy, 1938). The distribution of the environmental limitation
372 estimated by our model could therefore be used as a proxy for the distribution of food availability among the individuals
373 in a population. Estimates of individual food availability are scarce, because they commonly have to be obtained from
374 intensive observations or analysis of stomach samples. Because our model provides a proxy of the individual food
375 availability throughout the entire population, it can be used for more detailed analyses of the dynamics of food
376 availability. For example, linking the estimated environmental limitation to the consumer density might reveal density
377 dependent feedbacks in the growth rate of individuals. The environmental limitation as a proxy for individual food
378 availability might also provide insight into feeding links between species. The environmental limitation in our model
379 always appears as a product with the maximum asymptotic length ($f_t \ell_\infty$), it should therefore first be scaled by an
380 estimate of the species specific maximum asymptotic length before it can be compared between species. Comparison
381 of this scaled proxy for individual food availability between species might then reveal links such as shared resources
382 or competition.

383 In our example with North sea plaice we showed that our model can be used to explore environmental segregation
384 between length or age groups as well. The R-package (Croll, 2022) allows to split a population into a priori defined
385 length or age groups and fits an environmental limitation separately for every length or age group. With a realistic
386 division in length or age groups, our model could provide valuable information about growth limitation in different
387 life stages. If for example length or age groups show very diverse patterns in environmental limitation, it is likely that
388 the length or age groups are environmentally separated either through segregation in space or differences in diet. In
389 this way the model could therefore yield additional understanding in the growth dynamics during various life stages.

390 It is often difficult to judge whether a certain division in length or age groups is valid. A way to assess the suitability
391 of a division in age or length groups is to look at the estimated variance in environmental limitation for every group
392 compared to the variance in environmental limitation estimated in a fit without groups. Without groups, Our general
393 model combines all length and age groups and fits a single environmental limitation for all groups. This method

lumps together the variation in environmental limitation within and between age or length groups which will lead to a high estimate of the variance in environmental limitation if groups strongly differ in environmental limitation. In general, one can assume that individuals within a group are more similar to each other compared to individuals within another group. The variance in environmental limitation estimated for a single group is therefore expected to be lower compared to the variance in environmental limitation estimated for the entire population. Our model can therefore be used to verify whether a suspected division in ecological groups is likely by comparing the estimated variance in environmental limitation of a fit without ecological groups with a fit with ecological groups as is done in the example for north sea plaice.

Lastly our method might also be applicable to management as it is able to model variation in growth through time and between individuals based on only a limited number of parameters (Flinn and Midway, 2021). Many management models, in particular those used to estimate reference points for fish stock managements, assume a fixed length-at-age relationship, while variation in growth rates is shown to be important for the response of populations to exploitation (Lorenzen and Enberg, 2002). With our model the variation in growth between years can be easily quantified, resulting in a more accurate prediction of the length at age for every cohort. The age-length relationships from our model then can be used to calculate a more precise estimate of the needed reference points.

In conclusion, our model provides a way to estimate growth curves and length distributions of individual cohorts based on single individual length and age observations. In our model growth is allowed to vary over time, while our model also accounts for variation between individuals and variation between cohorts. So far, these factors could only be estimated simultaneously if the growth history of individuals was known. Models for single observations of individual length and age only accounted of a subset of these factors. Our model does account for all these factors and in this way estimates a proxy for the limitation in individual growth, which may vary over time. This estimate of the limitation in individual growth is a new step in understanding patterns in individual growth based on individual field observations.

Acknowledgement

This research is funded by the European Union's Horizon 2020 research and innovation programme under the grant agreement No. 773713, also known as the Pandora Project.

Conflict of interest statement

The authors declare that there is no conflict of interest.

422 **Author contributions**

423 The authors conceived the ideas for the proposed method together. Jasper Croll preformed the mathematical deriva-
424 tions of the model and the examples given for North Sea plaice. Both authors critically evaluated the results and
425 contributed significantly to the drafts. Both authors gave final approval for publication.

References

- Basimi, R. A. and Grove, D. J. (1985). Estimates of daily food intake by an inshore population of *Pleuronectes platessa* L. off eastern Anglesey, North Wales. *J. Fish Biol.*, 27:505–520.
- Bertalanffy, L. v. (1938). A quantitative theory of organic growth (inquiries on growth laws. ii.). *Human Biology*, 10(2):181–213.
- Beverton, R. J. H. and Holt, S. J. (1959). *A Review of the Lifespans and Mortality Rates of Fish in Nature, and Their Relation to Growth and Other Physiological Characteristics*. Novartis Foundation Symposia.
- Braber, K. and De Groot, S. J. (1973). The food of five flatfish species (pleuronectiformes) in the southern north sea. *Netherlands Journal of Sea Research*, 6(1-2):163–172.
- Calder, W. I. (1984). *Size, Function and Life History*. Harvard University Press.
- Castanet, J. (1994). Age estimation and longevity in reptiles. *Gerontology*, 40(174-192).
- Cloern, J. E. and Nichols, F. H. (1978). A von bertalanffy growth model with a seasonally varying coefficient. *J. Fish. Res. Bd. Canada*, 35(479-1482).
- Croll, J. C. (2022). Vbgfit package.
- 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(4):602–617.
- Flinn, S. A. and Midway, S. R. (2021). Trends in growth modeling in fisheries science. *Fishes*, 6(1):1–18.
- Graaf, G. D. and Prein, M. (2005). Fitting growth with the von bertalanffy growth function: a comparison of three approaches of multivariate analysis of fish growth in aquaculture experiments. *Aquaculture Research*, 36(1):100–109.
- He, J. X. and Bence, J. R. (2007). Modeling annual growth variation using a hierarchical bayesian approach and the von bertalanffy growth function, with application to lake trout in southern lake huron. *Transactions of the American Fisheries Society*, 136(2):318–330.
- ICES (2021). ICES BTS (Beam Trawl Survey) dataset.
- Johnson, S. G. (2022). The nlopt nonlinear-optimization package.
- Kingsley, M. C. S. (1979). Fitting the von bertalanffy growth equation to polar bear age-weight data. *Can. J. Zool.*, 57:1020–1025.
- Kooijman, S. A. L. M. (2010). *Dynamic Energy Budget theory*. Cambridge university press.

- 453 Lipinski, M. and Roeleveld, M. (1990). Minor extension of the von bertalanffy growth theory. *Fisheries Research*,
454 9:367–371.
- 455 Lorenzen, K. (1996). A simple van bertalanffy model for density-dependent growth in extensive aquaculture, with an
456 application to common carp (*Cyprinus carpio*). *Aquaculture*, 142:191–205.
- 457 Lorenzen, K. and Enberg, K. (2002). Density-dependent growth as a key mechanism in the regulation of fish pop-
458 ulations: evidence from among-population comparisons. *Proceedings of the Royal Society of London. Series B:*
459 *Biological Sciences*, 269(1486):49–54.
- 460 Maceina, M. J., Boxrucker, J., Buckmeier, D. L., Gangl, R. S., Lucchesi, D. O., Isermann, D. A., Jackson, J. R., and
461 Martinez, P. J. (2007). Current status and review of freshwater fish aging procedures used by state and provincial
462 fisheries agencies with recommendations for future directions,. *Fisheries*, 32(7):329–340.
- 463 Narinc, D., Narinc, N. Ö., and Aygün, A. (2017). Growth curve analyses in poultry science. *World’s Poultry Science*
464 *Journal*, 73(2):395–408.
- 465 Pauly, D. (1980). On the interrelationships between natural mortality, growth parameters, and mean environmental
466 temperature in 175 fish stocks. *J. Cons. int. Explor. Mer.*, 39(2):175–192.
- 467 Peters, R. (1983). *The Ecological Implications of Body Size*. Cambridge University Press.
- 468 Pilling, G. M., Kirkwood, G. P., and Walker, S. G. (2002). An improved method for estimating individual growth
469 variability in fish, and the correlation between von bertalanffy growth parameters. *Canadian Journal of Fisheries*
470 *and Aquatic Sciences*, 59(3):424–432.
- 471 Piner, K. R., Lee, H.-H., and Maunder, M. N. (2016). Evaluation of using random-at-length observations and anequi-
472 librium approximation of the population age structurein fitting the von bertalanffy growth function. *Fisheries*
473 *Research*, 180:128–138.
- 474 Prajneshy and Venugopalan, R. (1999). von bertalanffy growth model in a random environment. *Can. J. Fish. Aquat.*
475 *Sci*, 56:1026–1030.
- 476 Rafail, S. Z. (1973). A simple and precisemethod for fitting a von bertalanffy growth curve. *Marine Biology*, 19:354–358.
- 477 Ramirez, M. D., Popvska, T., and Bbcock, E. A. (2021). Global synthesis of sea turtle von bertalanffy growth
478 parameters through bayesian hierarchical modeling. *Marine Ecology Progress Series*, 657:191–207.
- 479 Read, F. L., Hohn, A. A., and Lockyer, C. H. (2018). A reveiw of age estimation methods in marine mammals with
480 special referene to monodontids. *NAMMCO*, 10.

- 481 Rijnsdorp, A. D. and Vingerhoed, B. (2001). Feeding of plaice (*Pleuronectes platessa* L.) and sole (*Solea solea* L.) in
482 relation to the effects of bottom trawling. *Journal of Sea Research*, 45:219–229.
- 483 Robson, S. K. A. and Crozier, R. H. (2009). An evaluation of two biochemical methods of age determination in insects
484 (pteridines and lipofuscins) using the ant*Polyrhachis sexpinosa*Latrielle (hymenoptera: Formicidae). *Australian*
485 *Journal of Entomology*, 48(2):102–106.
- 486 Smirina, E. M. (1994). Age determination and longevity in amphibians. *Gerontology*, 40:133–146.
- 487 Taylor, N. G., Walters, C. J., and Martell, S. J. (2005). A new likelihood for simultaneously estimating von bertalanffy
488 growth parameters, gear selectivity, and natural and fishing mortality. *Canadian Journal of Fisheries and Aquatic*
489 *Sciences*, 62(1):215–223.
- 490 Teleken, J. T., Galvao, A. C., and Robazza, W. D. S. (2017). Comparing non-linear mathematical models to describe
491 growth of different animals. *Acta Scientiarum. Animal Sciences*, 39(1):73.
- 492 Van der Veer, H. W., Kooijman, S. A. L. M., and Van der Meer, J. (2001). Intra- and interspecies comparison of energy
493 flow in north atlantic flatfish species by means of dynamic energy budgets. *Journal of Sea Research*, 45:303–320.
- 494 Vaughan, D. S. and Kanciruk, P. (1982). An empirical comparison of estimation procedures for the von bertalanffy
495 growth euqation. *J. Cons. int. Explor. Mer.*, 40:211–219.
- 496 Vincenzi, S., Crivelli, A. J., Munch, S., Skaug, H. J., and Mangel, M. (2016). Trade-offs between accuracy and
497 interpretability in von bertalanffy random-effects models of growth. *Ecological applications*, 26(5).
- 498 Vincenzi, S., Mangel, M., Crivelli, A. J., Munch, S., and Skaug, H. J. (2014). Determining individual variation
499 in growth and its implication for life-history and population processes using the empirical bayes method. *PLoS*
500 *Computational Biology*, 10(9):e1003828.
- 501 Wang, Y.-G. and Thomas, M. R. (1995). Accounting for individual variability in the von ertalanffy growth model.
502 *Can. J. Fish. Aquat. Sci*, 52(1368-1375).

503 Appendix

504 Model test on simulated data

505 We fitted the model to simulated data to test the sensibility of the estimated parameters to the number of observations
506 in the data and possible biases in length in the data. To do so, we generated the population structure of a population

Table A1: Parameters used for simulating population structure

Parameter definition	Symbol	Value	Unit
Von Bertalanffy growth rate scalar	r_B	0.001	d^{-1}
Maximum asymptotic length	ℓ_∞	1000	mm
Mean environmental limitation	μ_f	0.5	-
Variance in environmental limitation	σ_f^2	0.0002	-
Minimum length at birth	ℓ_{0min}	5	mm
Maximum length at birth	ℓ_{0max}	15	mm
Maximum cohort age	-	3650	d
Time between cohort starts	-	365	d
Time between cohort start and sampling	-	180	d

with a constant mean and variance of the environmental limitation. To do so we used the following continuous model:

$$\frac{\partial E(\ell)}{\partial t} + \frac{\partial E(\ell)}{\partial a} = r_B (\mu_f \ell_\infty - E(\ell)) \quad (9)$$

$$\frac{\partial V(\ell)}{\partial t} + \frac{\partial V(\ell)}{\partial a} = 2r_B (\sigma_f^2 \ell_\infty^2 - V(\ell)) \quad (10)$$

The average length at birth was determined by drawing a random value from a predetermined interval ($\ell_{0min} - \ell_{0max}$). Other parameters of the simulation were fixed (Tab. A1). The population structure was simulated over a period of 30 years and every cohort was simulated from age 0 to 10. A sample was drawn from the population structure for every year and cohort at 180 days after a new cohort with age 0 is introduced to the population.

To test the influence of the number of parameters on the estimated environmental limitation, we draw 100 data sets from the simulated population structure. The number of samples varied between 10 and 100 samples per cohort per year (Fig. A1). The estimated mean and variance of the environmental limitation are distributed evenly around the actual mean and variance in the simulation. With an increasing number of samples, the standard error of the estimated mean and variance of the environmental limitation become smaller. In other words, the estimated mean and variance in environmental limitation become on average closer to the real mean and variance if the number of samples increases.

Another concern when estimating environmental limitations from sampled length structures is a possible bias in the sample method. It is possible that the samples are biased towards individuals with a larger or smaller length. To test the effect of this bias on the estimated environmental limitation we sampled 100 data sets with 50 observations per cohort per year from the simulated population structure. When fitting our model we weighed the samples by individual length. The weight of a sample changed linearly with individual length ($w_i = a\ell_i + b$). The slope of this relationship determined the strength of the bias, ranging from a sampling bias towards larger individuals ($a > 0$) to

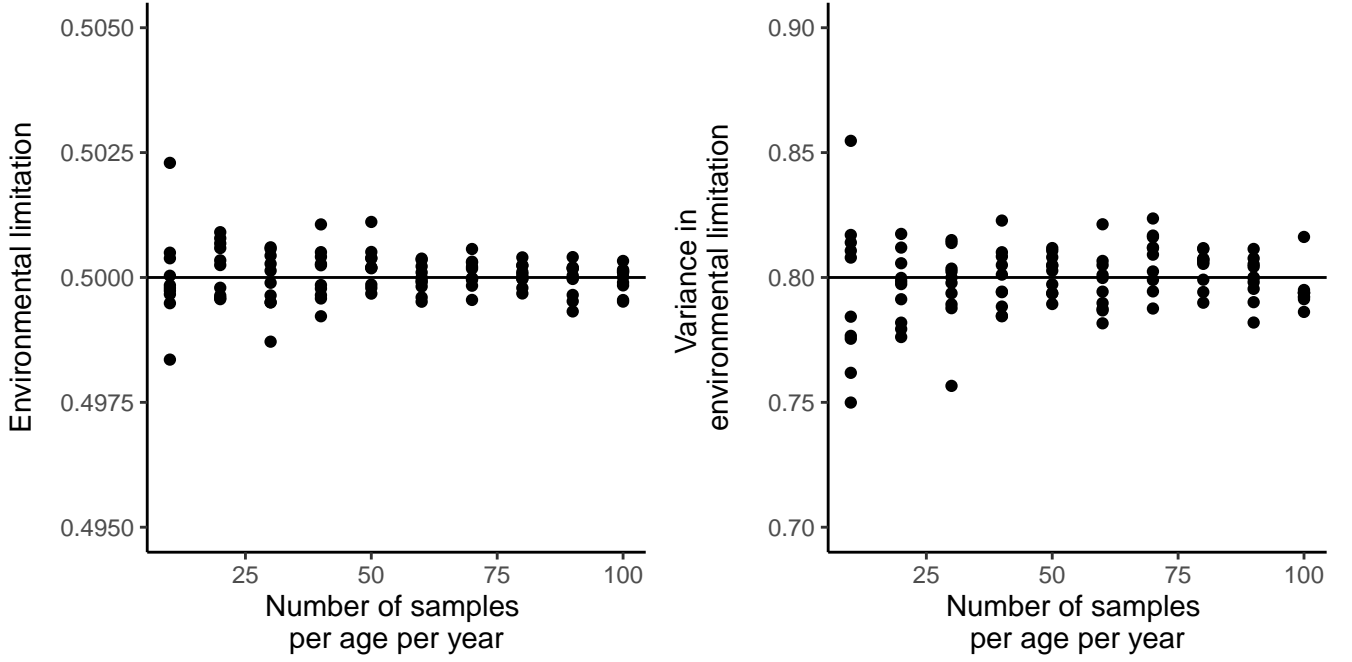


Figure A1: Average of the estimated mean and variance of the environmental limitation for different number of data points drawn from a simulated population structure. The black line indicates the actual values of the mean and variance in environmental limitation in the simulation.

a sampling bias towards smaller individuals ($a < 0$). The intercept of the relation between bias weight and length is determined in such a way that the average weight of samples is equal to one ($b = 1 - \frac{\sum_{i=1}^n a \ell_i}{n}$). This will ease comparison between model fits, because the weighted number of samples is equal for all datasets.

A sample bias linked to weight might result in an error in the estimation of the mean environmental limitation, but does not affect the estimation of the variance in environmental limitation (Fig. A2). The mean environmental limitation is slightly overestimated if larger individuals are more likely to end up in the samples, while the mean environmental limitation is slightly underestimated when smaller individuals are more likely to end up in the sample. Although this error seems to be small, it is important to keep this bias in mind and correct for this bias when possible.

Supplementary figures

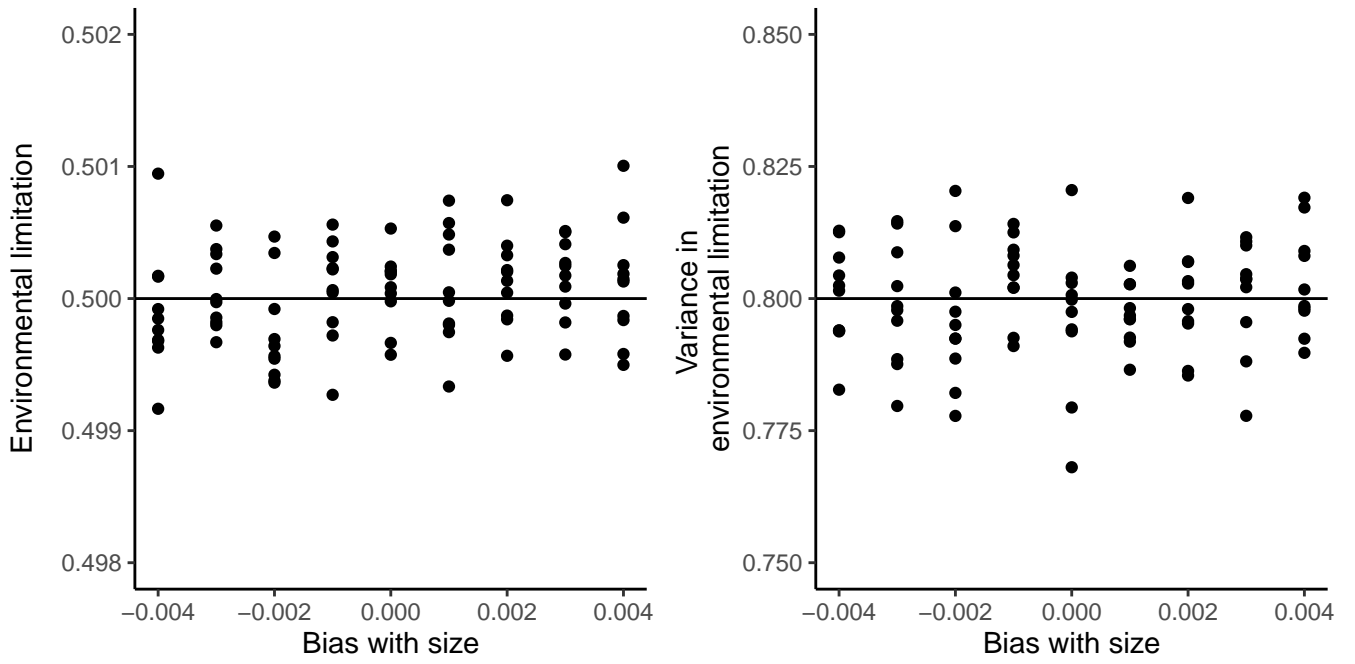


Figure A2: Average of the estimated mean and variance of the environmental limitation when the data contains a bias with length. A positive bias with length indicates that larger individuals contribute more to the samples, while a negative bias with length indicates that smaller individuals contribute more to the samples. The black line indicates the actual values of the mean and variance in environmental limitation in the simulation.

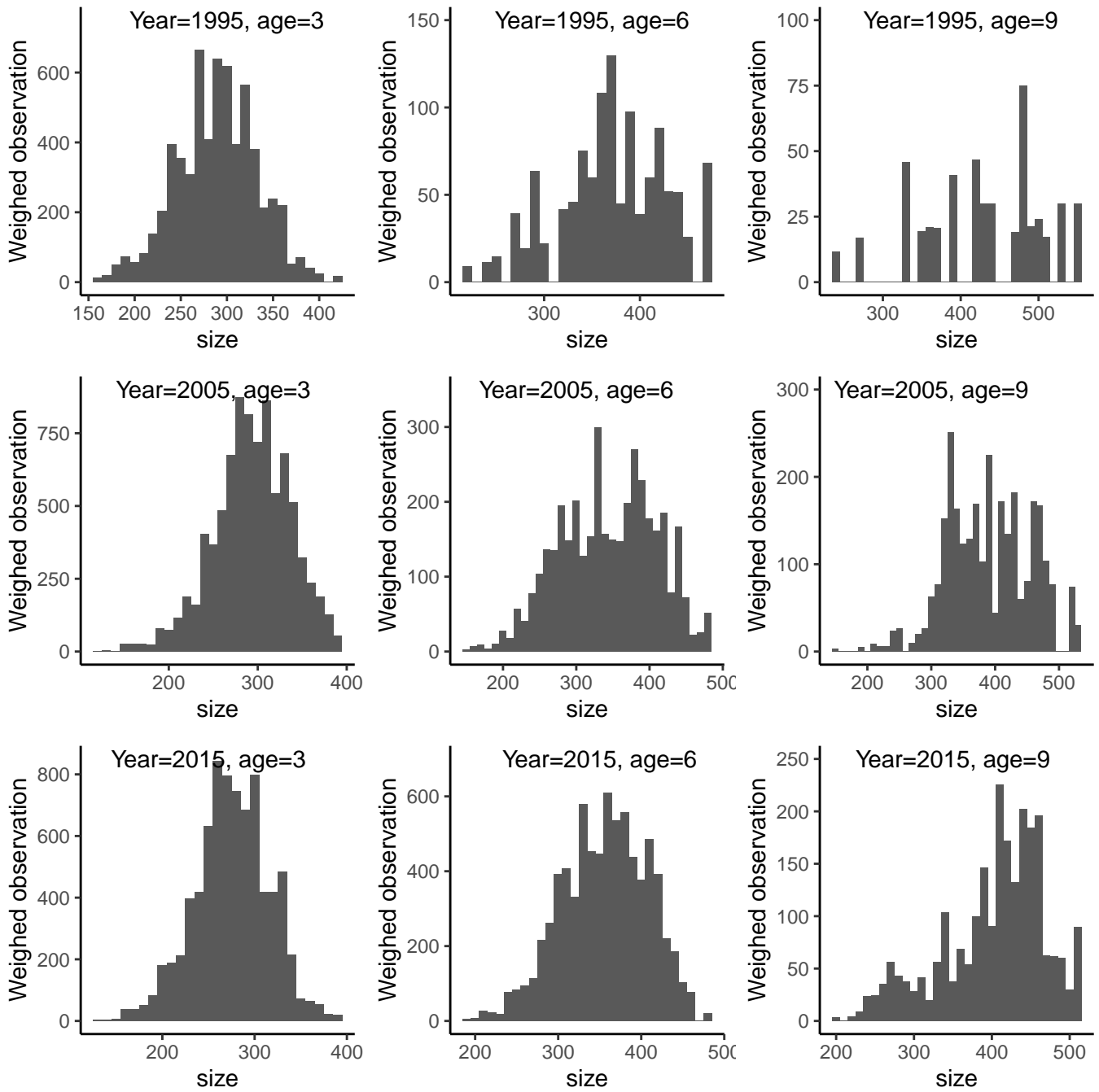


Figure A3: Weighted distributions of length at age at a given year. The number of observed individuals with a specific length is weighted by the inverse of the catch per unit effort of that length in the given year. The length at age seems to approximately follow a normal distribution.