



Age composition and growth without age data: a likelihood-based model

Diego Corrêa Alves^{a,b,*}, Lilian Paula Vasconcelos^b, Angelo Antonio Agostinho^{b,c}

^a Departamento de Estatística, Universidade Estadual de Maringá, Avenida Colombo, 5790, 87020-900 Maringá, PR, Brazil

^b Programa de Pós-Graduação em Ecologia de Ambientes Aquáticos Continentais, Departamento de Biologia, Universidade Estadual de Maringá, Avenida Colombo, 5790, 87020-900 Maringá, PR, Brazil

^c Núcleo de Pesquisas em Limnologia, Ictiologia e Aquicultura (Nupélia), Universidade Estadual de Maringá, Avenida Colombo, 5790, 87020-900 Maringá, PR, Brazil

ARTICLE INFO

Handled by A.E. Punt

Keywords:

Fish growth
Fisheries
Population dynamics
Modelling
Stock assessment
von Bertalanffy

ABSTRACT

We proposed models capable of jointly estimating age composition and somatic growth parameters (L_{∞} and K) from length-frequency data without the need to obtain age data. The proposed approach consists of a linear regression in which both the regression coefficients (age composition) and the predictor variables (size distribution at each age) are unknown. The predictor variables correspond to theoretical simulated values from a growth curve, whose parameters are jointly estimated with the regression coefficients using a robust global optimization algorithm, differential evolution, which uses stochastic procedures with parallel methods of direct search. The proposed models were assessed using a simulation study with two sets of virtual fish populations, representing two different growth curves. The parameter estimates of the age composition were equally precise and accurate among models in which the growth parameters were estimated or known *a priori*. Furthermore, the estimates obtained by the models that also estimated the growth parameters were unbiased and accurate. The estimates of growth parameters are an alternative for cases in which the relationship between length and age is unknown, outdated or limited. The models presented in this study can be applied to various groups of organisms other than fish.

1. Introduction

The age composition of a population is an important input for population dynamics models and fishery stock assessments, making it possible to estimate recruitment and mortality (Shepherd, 1984; Magnusson and Hilborn, 2007; Maunder and Piner, 2015). However, obtaining age data is still difficult. One option could be aging individuals from a random sample that represents the population. This aging is performed based on the number of rings formed consistently in hard body parts, such as otoliths, scales or other bone parts. However, the time required to obtain and prepare these structures, the difficulty in reading these rings and the lack of knowledge of the time and periodicity of their formation in many species make this process very expensive and often imprecise and inaccurate (Campana, 2001; Chang and Maunder, 2012). Thus, it is practically unfeasible to obtain age composition data only by aging individuals, at least for most fish species, as well as for other animals or plants.

As an alternative, mathematical and statistical methods have been developed to obtain the age composition of a population based on length-frequency data, which are faster, easier and cheaper to obtain. These methods are based on the association between length and age in

individuals obtained from samples of the study population. In general, it is possible to evaluate this relationship using two approaches; the first expresses the age distribution of each length class, while the second assumes, in an inverse manner, that the length distribution is influenced by age (Kimura and Chikuni, 1987; Hoenig et al., 1994).

The first approach is based on the ideas developed by Fridriksson (1934) and uses an age-length key (ALK), also known as a classical or forward key (Kimura and Chikuni, 1987; Hoenig et al., 1994). The ALK can be represented by a matrix, with j lines (the number of size classes) and i columns (the number of age groups). Each row of this matrix contains the age distribution of a size class; that is, each pair j, i of this matrix represents the conditional probability of an individual being of age i , given its size class j , $P(\text{age} = i \mid \text{size-class} = j)$. With this method, the rows should add to one. To estimate these conditional probabilities comprising the age-length key, a stratified sample by size-class is required, with cross-classification of age and length. Then, based on these stratified samples, it is possible to estimate the conditional probabilities and to classify the age groups into a second length-composition sample. Although this approach restricts the number of individuals to be aged, it traditionally requires age-length stratified subsamples for each population and period corresponding to length-composition samples. This is

* Corresponding author at: Departamento de Estatística, Universidade Estadual de Maringá, Avenida Colombo, 5790, 87020-900 Maringá, PR, Brazil.
E-mail address: dcalves@uem.br (D.C. Alves).

because the ALK uses the information of the age composition of the population and the period in which the stratified sample was obtained, generating a source of bias when applied to other populations or to the same population during a different period. However, this approach has been improved to minimize the bias generated by using “lagged” ALKs (Kimura and Chikuni, 1987; Hoening et al., 1994).

The second approach arises from the ideas of Clark (1981), who aimed to remove the ALK bias. This approach proposed standardizing the ALK through the inversion of the conditionals, resulting in the length distributions for each age, representing somatic growth. This association can also be expressed by a matrix, usually referred to as the inverse-ALK, transition matrix (P), or length-distribution matrix, hereinafter referred to as the transition matrix (Clark, 1981; Bartoo and Parker, 1983; Kimura and Chikuni, 1987; Parrack and Cummings, 2003). Each column of the transition matrix contains the probability distribution of an individual belonging to size class j , due to its age i (P (size-class = j | age = i)). Thus, the columns of the transition matrix should add to one. Finally, Bartoo and Parker (1983) tried to eliminate the need for the stratified sample required for ALK and inverse-ALK construction, suggesting the theoretical-transition matrix could be estimated using a growth curve. The latter procedure added a stochastic component, in this case the transition matrix, which also eliminated the bias of the traditional deterministic method based on the growth curve, cutting or slicing (Bartoo and Parker, 1983).

Both approaches depend to some extent on age samples, either to compose ALK or to estimate length distribution by age (growth), although the growth-curve approach requires aging fewer individuals (Bartoo and Parker, 1983). This method assumes that the transition matrix is fixed and known, or that the somatic growth curve of individuals of the population does not change over time, at least between the periods in which the transition matrix and the length-frequency data were obtained (Clark, 1981; Bartoo and Parker, 1983). However, in many situations, the growth pattern may change, hindering the application of these methods.

Thus, a statistical model based on likelihood is proposed to estimate both age composition and somatic growth parameters from a single length-frequency sample, but that be representative of the population and without sample bias. The proposed model is a continuation of the ideas raised by Bartoo and Parker (1983). However, it does not require the use of age samples or *a priori* knowledge of the growth parameters responsible for generating the transition matrix. Additionally, it eliminates the assumption that the transition matrix is fixed and known.

2. Methods

First, two sets of population parameters were defined, representing two populations with different life-history strategies. These parameters were used to generate 500 stochastic datasets per scenario through an operational model (Fig. 1). The operational model determines the biological mechanism for data generation and was based on an age-structured population dynamics model. The length-frequency data generated by the operational model were used to fit four statistical models that were then evaluated. Two of these models estimated only the age composition parameters, while the other two models estimated both the parameters of the age composition and growth (Fig. 1). The four models were based on the traditional von Bertalanffy equation (VBGE), although any growth model could have been adopted.

Two sets of scenarios representing different life-history strategies were selected. The first scenario was characterized by a growth curve with high overlap of the length's distributions by age groups, while the second scenario was represented by a growth curve with less overlap among age groups. Thus, the peacock bass *Cichla kelberi* and goliath catfish *Brachyplatystoma rousseauxii* species formed the basis for the first and second scenarios, respectively. *C. kelberi* is a species with sedentary habits, displaying parental care and a faster life cycle than *B. rousseauxii*. The goliath catfish does not display parental care and performs

the longest freshwater fish migration in the world, migrating hundreds or even thousands of kilometers to complete its life cycle (Barthem et al., 2017). Fishes with highly developed parental care, such as the peacock bass, have slow young-of-year and adult growth (Winemiller and Rose, 1992), which leads to a greater overlap in the length distribution. On the other hand, seasonal species (*sensu* Winemiller, 1989), like *B. rousseauxii*, were selected to synchronize large clutches in short periods in which the growth and survival of young were favorable (Winemiller and Rose, 1992). This contributes to reduce the overlaps in length distribution.

2.1. Operational model

The operational model was structured by age and year (Fig. 1), with 10 age groups, where $N_{i,y}$ is the number of individuals in the age group i in the middle of year y ($N_{i,y}$, $i = \{0,1,2,\dots,9\}$, $y = \{1,2,3,\dots,12\}$). The number of individuals in age group i and year y equals the product of the number of individuals in the age group $i-1$ and year $y-1$ and the negative of the exponential rate Z , defined as the instantaneous rate of total mortality, as follows:

$$N_{i,y} = N_{i-1,y-1} e^{-Z} \quad (1)$$

The number of individuals in the first age group in the middle of year y was defined as the recruitment of year y ($N_{0,y}$) and assumed to be a random variable with a uniform probability distribution in the interval $(10^3, 10^6)$, according to Parrack and Cummings (2003). The population parameters of the age composition (a_i , with i varying from 0 to 9) corresponded to the number of individuals in each age group in the last year, $y = 12$, $N_{i,y=12}$. The length (s) of each individual was randomly generated through a Gaussian (normal) probability distribution, with the expected length defined as a function of age (t), using the von Bertalanffy equation with a standard deviation equal to 3 (Bartoo and Parker, 1983).

$$E(s|t) = L_{\infty} (1 - e^{-K(t-t_0)}) \quad (2)$$

where L_{∞} , K and t_0 are parameters of the von Bertalanffy equation, which represent the asymptotic length, for which growth is zero, growth coefficient and theoretical age when the length is zero, respectively. The values of these growth parameters, as well as the Z , used as input in the operational model, were based on the study of Gomiero et al. (2010), for *C. kelberi* ($L_{\infty} = 57.7$; $K = 0.53$; $t_0 = -0.39$; $Z = 0.66$), and of Córdoba et al. (2013), for *B. rousseauxii* ($L_{\infty} = 153.3$; $K = 0.22$; $t_0 = -0.49$; $Z = 1.14$). Thus, the index of overlap, obtained according to Hart and Reynolds (2002), were 0.556 and -0.004 for *C. kelberi* and *B. rousseauxii*, respectively. This index reflects the proportion in which the 95% zone for the age group is overlapped by the 95% zone for the next age group, being that the modes disappear with values higher than 0.25 (Hart and Reynolds, 2002), as is the cases involving *C. kelberi*.

Finally, 500 data sets were generated for each scenario, resulting in a dataset consisting of 500 length frequencies, with 1 cm length-class, for each species, whose true parameters of age composition, von Bertalanffy equation (VBGE), and total mortality were known (Fig. 1). Therefore, this database was used to fit the four estimation models, as described below.

2.2. Estimation models

The estimation models are based on an age-transition matrix for lengths (P matrix). Traditionally, this matrix is assumed to be fixed and known and can be obtained with theoretical values through a somatic growth model with known parameters or values estimated by an inverse age-length key. In this study, the P matrix was composed of theoretical values, derived from the von Bertalanffy growth function, computed as described by Parrack and Cummings (2003) as follows:

Download English Version:

<https://daneshyari.com/en/article/8885425>

Download Persian Version:

<https://daneshyari.com/article/8885425>

[Daneshyari.com](https://daneshyari.com)