Optimal exponent-pairs for the Bertalanffy-Pütter growth model

Katharina Renner-Martin Corresp., [1] , **Norbert Brunner** [1] , **Manfred Kühleitner** [1] , **Werner-Georg Nowak** [1] , **Klaus Scheicher** [1]

[1] Department of Integrative Biology and Biodiversity, Institute of Mathematics, Universität für Bodenkultur Wien, Vienna, Austria

Corresponding Author: Katharina Renner-Martin
Email address: kathi.renner-martin@gmx.de

The Bertalanffy-Pütter growth model describes mass *m* at age *t* by means of the differential equation $dm/dt = p \cdot m^a - q \cdot m^b$. The special case using the Bertalanffy exponent-pair a=2/3 and b=1 is most common (it corresponds to the von Bertalanffy growth function VBGF for length in fishery literature). For data fitting using general exponents, five model parameters need to be optimized, the pair $a<b$ of non-negative exponents, the non-negative constants p and q, and a positive initial value m_0 for the differential equation. For the case b=1 it is known that for most fish data any exponent a<1 could be used to model growth without affecting the fit to the data significantly (when the other parameters p, q, m_0 were optimized). Thereby, data fitting used the method of least squares, minimizing the sum of squared errors (*SSE*). It was conjectured that the optimization of both exponents would result in a significantly better fit of the optimal growth function to the data and thereby reduce *SSE*. This conjecture was tested for a data set for the mass-growth of Walleye (*Sander vitreus*), a fish from Lake Erie, USA. Compared to the Bertalanffy exponent-pair the optimal exponent-pair achieved a reduction of *SSE* by 10%. However, when the optimization of additional parameters was penalized, using the Akaike information criterion (*AIC*), then the optimal exponent-pair model had a higher (worse) *AIC*, when compared to the Bertalanffy exponent-pair. Thereby *SSE* and *AIC* are different ways to compare models. *SSE* is used, when predictive power is needed alone, and *AIC* is used, when simplicity of the model and explanatory power are needed.

Optimal exponent-pairs for the Bertalanffy-Pütter growth model

Katharina Renner-Martin, Norbert Brunner, Manfred Kühleitner, Werner-Georg Nowak, and Klaus Scheicher

Institute of Mathematics, Department of Integrative Biology and Biodiversity Research, University of Natural Resources and Life Sciences, Gregor Mendel Straße 33, A-1180 Vienna, Austria

Corresponding author:

Katharina Renner-Martin

E-mail: kathi.renner-martin@gmx.de

Statements

The authors declare no conflict of interests.

The first author was supported by a grant from the University of Natural Resources and Life Sciences, Vienna.

Optimal exponent-pairs for the Bertalanffy-Pütter growth model

Abstract. The Bertalanffy-Pütter growth model describes mass m at age t by means of the differential equation $dm/dt = p \cdot m^a - q \cdot m^b$. The special case using the Bertalanffy exponent-pair $a = 2/3$ and $b = 1$ is most common (it corresponds to the von Bertalanffy growth function VBGF for length in fishery literature). For data fitting using general exponents, five model parameters need to be optimized, the pair $a < b$ of non-negative exponents, the non-negative constants p and q, and a positive initial value m_0 for the differential equation. For the case $b = 1$ it is known that for most fish data any exponent $a < 1$ could be used to model growth without affecting the fit to the data significantly (when the other parameters p, q, m_0 were optimized). Thereby, data fitting used the method of least squares, minimizing the sum of squared errors (*SSE*). It was conjectured that the optimization of both exponents would result in a significantly better fit of the optimal growth function to the data and thereby reduce *SSE*. This conjecture was tested for a data set for the mass-growth of Walleye (*Sander vitreus*), a fish from Lake Erie, USA. Compared to the Bertalanffy exponent-pair the optimal exponent-pair achieved a reduction of *SSE* by 10%. However, when the optimization of additional parameters was penalized, using the Akaike information criterion (*AIC*), then the optimal exponent-pair model had a higher (worse) *AIC*, when compared to the Bertalanffy exponent-pair. Thereby *SSE* and *AIC* are different ways to compare models. *SSE* is used, when predictive power is needed alone, and *AIC* is used, when simplicity of the model and explanatory power are needed.

Subjects Computational Biology, Aquaculture (Fisheries and Fish Science), Mathematical Biology, Computational Science

Keywords Bertalanffy-Pütter differential equation, Akaike information criterion (*AIC*), Region of near-optimality

INTRODUCTION

Size-at-age (length or mass) is an important metric about animals (Google search: ca. 286,000 results), in particular for fisheries management (Ogle & Iserman, 2017). Consequently, various models for size-at-age have been proposed. This paper investigates a general class of growth models, defined from the Bertalanffy (1957) and Pütter (1920) differential equation (1):

$$(1)\ \frac{dm(t)}{dt} = p \cdot m(t)^a - q \cdot m(t)^b$$

Equation (1) describes body mass (weight) $m(t) > 0$ as a function of age t, using five model parameters: a, b, p, q, m_0. Thereby, $m_0 > 0$ is an initial value, i.e. $m(0) = m_0$. The exponent-pair $a < b$ ('metabolic scaling exponents') is assumed to be non-negative and also the constants p and q are non-negative. Several 'named models' are special instances of (1): To describe mass-at-age, Bertalanffy (1957) suggested the exponent-pair $a = 2/3$ and $b = 1$, West, Brown & Enquist (2001) proposed $a = 3/4$, $b = 1$, other authors considered $a = 1$, $b = 2$ (logistic growth of

Verhulst, 1838), Richards (1959) recommended $a = 1$ while retaining $b > 1$ as a free parameter, and the generalized Bertalanffy growth model assumes $b = 1$, using $a < 1$ as parameter. There are also models of type (1) for length-at-age, notably VBGF, the von Bertalanffy growth function with exponent-pair $a = 0$, $b = 1$ (bounded exponential growth) which is widely used in fishery literature (Google search for 'VBGF, fish': ca. 15,000 results). VBGF is equivalent to the model with the Bertalanffy exponent-pair ($a = 2/3$, $b = 1$) for mass-growth (Bertalanffy, 1957).

In the case of equal exponents, equation (1) is replaced by the differential equation (2). Its right hand side is the limit of the right hand side of (1), assuming b approaches a. As the special case $a = 1$ of equation (2) defines the Gompertz (1832) model, with respect to equation (2) the paper refers to the class of Gompertz models.

$$(2)\ \frac{dm(t)}{dt} = p \cdot m(t)^a - q \cdot \ln(m(t)) \cdot m(t)^a$$

In general, the solutions of (1) and (2) are non-elementary functions, namely hypergeometric functions and exponential integrals, respectively (Ohnishi, Yamakawa & Akamine, 2014; Marusic & Bajzer, 1993). The solutions of the more special 'named models' are elementary.

Parameter values for equations (1), (2) were obtained by identifying a growth function (i.e. a concrete solution of the differential equations) with the best fit to the data. Experience has shown that no single of the above-mentioned 'named models' was exactly correct for all species (c.f. Killen, Atkinson & Glazier, 2010 for fish; White, 2010 for mammals). Renner-Martin et al. (2018) explored the situation for the generalized Bertalanffy model ($b = 1$) and found that for most species of fish any exponent $0 \leq a < 1$ could be used to model growth without affecting the fit to the data significantly (when the other parameters p, q, m_0 were optimized). They explained this by data quality, as for wild-caught fish and also for wildlife data there is always the problem of 'haphazard' sampling, which may result in unreliable growth parameter estimates (Wilson et al., 2015).

The present paper asks, if the high variability for fish data is still observed in two dimensions, when both exponents (a, b) of (1) are optimized. In view of the computational complexity of optimizing the Bertalanffy-Pütter-model, the paper identifies optimal exponents for one fish data-set only.

Materials and methods

Data

The paper used 'FSAdata WalleyeErie2'from Ogle (2018) about Walleye (*Sander vitreus*) from Lake Erie, USA. A sub-sample (20,166 data-points) about male fish was retrieved. The data were insofar exceptional, as they informed about mass (in gram) and age (in years from otholits) of wild-caught fish, while most growth data for fish are length-at-age. Data were retrieved using MS Excel. Also a preliminary analysis was conducted in Excel (pivoting to identify average weights for the age classes). Figure 1 plots the data and the average weights.

There were few data about young fish (14 of age 0) and likewise few about older fish (22 with age 16-20 years), and none about fish with age 21-29 (maximal observed age reported in FishBase: Froese & Pauly, 2018). This may indicate gear bias (where small or large fish were not adequately sampled). In order to obtain more balanced class-sizes, smaller classes were merged; the outcome is Table 1, reporting of each class the average mass at the average age. Thus, 13 classes representing larger samples were evaluated instead of originally 20 age classes. The subsequent computations, i.e. search of optimal parameters for equations (1) and (2), used Mathematica 11.3.

At first it may appear troubling to take more than 20,000 data points and then aggregate them to merely 13 mass-at-age classes. However, for data fitting it was the distance between the model curve and the average of each class that mattered. The distances between the average and the other class data could not be improved by a growth model.

General approach to data fitting

As Shi et al. (2014) observed, already for the generalized Bertalanffy model (i.e. $b = 1$, a, p, q, m_0 are optimized) data fitting was impeded by numerical instability. Clearly, with more parameters to optimize the problem of convergence becomes more demanding and also powerful methods slow down. In order to avoid running into numerical instability by the use of too many parameters, the paper considered exponents lying on a grid. For each grid-point (exponent-pair a, b) model parameters (p, q, m_0) were identified that minimized the following function:

$$(3)\ SSE_{opt}(a,b) = \min_{m_0,p,q} (SSE) \text{ for growth functions with exponents } a, b$$

Thereby, the paper used the most common approach to data fitting, the method of least squares, which assesses the fit to the data by means of the sum of squared errors (*SSE*). However, even for simple models (meaning: certain values for the exponents are assumed and three parameters are optimized, e.g. p, q, m_0) literature reported that optimization failed to converge for certain data sets (Apostolidis & Stergiou, 2013). One of the reasons was the use of parametrizations that require bounded growth functions (e.g. Cailliet et al., 2006), whereas not all data may support bounded growth. Another reason was the observation that even for simple models the problem of data fitting may overtask straightforward optimization routines. In view of such difficulties with the convergence of optimization the paper did not add more complex model assumptions to (1) and (2), such as heteroscedastic growth that assumes a larger variance for a higher mass, or models that need additional parameters to distinguish different growth phases (Manabe et al., 2018). There are also various improvements of regression models, such as mixed-effect models to identify explanatory factors for growth (Strathe et al., 2010). However, such models require highly controlled experiments, whereas the present data are about wild-caught fish. Further, the purpose of optimization was the identification of a suitable growth curve for the considered species and not the identification of a growth curve that would minimize errors in relation to a given population. Therefore, no mass for class size were used for the computation of *SSE*. For the same reason, the observed variances of the age-classes were not used to assess the likelihood of each data point. (This means that the variance of the assumed normal distribution of errors was another implicit model parameter.) Further, optimization was not simplified by adding assumptions about parameter values, e.g. eliminating two parameters from optimization by using a literature value for the initial condition m_0 (rather than optimizing it) and using a literature value for the asymptotic mass (defined below). In this case $SSE_{opt}(a, b)$ could have been computed very fast from the optimization of one parameter, only, but at the cost of weakening the link to the data.

The use of grid-points helped to identify failures of optimization by a visual inspection (e.g. a grid-point with exceptionally high SSE_{opt}, when compared to neighboring grid-points). In order to do not miss the optimum, different approaches to data-fitting were used to identify and correct miscalculations. Thereby, computation time was an issue. For instance, commercially available software packages for fisheries management use powerful numerical methods to determine the model parameters even for the simple models (Mildenberger et al., 2017). These methods aim at optimizing one given model, where computing time is not an issue. The present paper aimed at optimizing a large number of models simultaneously in order to explore the function SSE_{opt}; i.e.

each grid point defined a model (defined from the exponent pair a, b) for which optimal parameters were identified. While for each grid-point SSE_{opt} could be obtained fast, optimizing over the whole grid was time consuming. For example, covering the region $0 \leq a \leq 1$, $a < b \leq 3$ by a grid with neighboring points at distance 0.01 would define 25,250 grid points. For this grid, assuming six optimizations per minute would require 70 hours computing time.

Optimization proceeded in three stages. First, SSE_{opt} was computed on a coarse grid (step-size 0.1) to sketch the shape of SSE_{opt} and locate a region of near-optimal exponents. This used methods of optimization that were fast, but not necessarily accurate. In the second stage, the computations were repeated with a finer grid (step-size 0.01) and using more accurate methods of optimization. These computations allowed to identify candidates for the optimum. In the final stage a search for the global optimum was performed, starting with these candidate points. The specific methods of optimization used in each step are explained below (c.f. the survey of Cedersund et al., 2015).

In order to speed up computations all approaches solved the differential equations (1) and (2) numerically (Leader, 2004). Using the analytic solutions of the differential equations (these are available in Mathematica) would make data fitting time consuming even for a given exponent pair. As the numerical methods used by Mathematica 11.3 work with high precision, this did not compromise the accuracy of optimization.

Starting values for data fitting

For most iterative methods of optimization, reasonable starting values for the parameters are needed to ensure convergence of optimization. For instance, the starting value for the initial value m_0 was the first data point of Table 1.

For the other parameters, practitioners use various rules of thumb for this purpose (Carvalho & Santoro, 2007), which utilize general considerations about the possible shape of the growth functions. For, the typical solutions of (1) and (2) are increasing, bounded and sigmoidal. (However, there are also non-sigmoidal solutions, e.g. $a = 0$, and unbounded solutions, e.g. $q = 0$ and $p > 0$.) Initially the rate of growth increases, until the inception point is reached. Subsequently it decreases to zero in the limit, when the asymptotic mass m_{max} is reached; there the right-hand side of (1) and (2), respectively, vanishes:

$$(4)\ m_{max} = \left(\frac{p}{q}\right)^{\frac{1}{b-a}} \text{ for } a < b\text{, eq. (1)}$$

To obtain a starting value q_0 for the parameter q, it was assumed that the asymptotic mass would exceed the maximal observed mass by 20%, i.e. the equation $1.2m_{max}$ was solved for q, referring to equation (4). This resulted in $q_0 = p_0/(1.2 \cdot m_{max})^{b-a}$, where $p_0$was the starting value for p.

In order to obtain a starting value for p, equation (1) was evaluated approximately at $t = 0$, using for the right-hand side the above-mentioned starting value for m_0 for m and q_0 for q. An approximate value for the derivative, $m'(0)$,was obtained from the derivative at $t = 0$ of the quadratic interpolation polynomial (Burden& Faires, 1993) through the first three points listed in Table 1. This polynomial was an approximation for the growth function in the neighborhood of $t = 0$. Solving (1) for $p = p_0$ resulted in the following equation:

$$(5)\ p_0 = \frac{m'(0) \cdot 1.2^b \cdot m_{max}^b}{1.2^b \cdot m_{max}^b \cdot m_0^a - 1.2^a \cdot m_{max}^a \cdot m_0^b}$$

These formulas defined starting values for m_0, p and q. The formulas were problematic for exponents close to the diagonal, as the function p_0 tends to infinity in the limit $a \to b$. Therefore, for exponents $b = a + 0.01$, in case that optimization using these starting values did not converge Simulated Annealing (see below) was used.

Preparatory screening

SSE_{opt} was computed for a coarse grid (distance 0.1 between adjacent points), using two general purpose methods for global optimization in parallel, simulated annealing and the Nelder-Mead amoeba method. Both methods are available for the Mathematica function NMinimize.

Simulated annealing was used, as it was expected to produce reasonable results. It used random numbers as starting values (using multiple starting values) and then altered them by random fluctuations, accepting parameters with lower values of *SSE*, but also accepting with a certain probability (that became lower in subsequent iteration steps) parameters with a higher *SSE* to escape from suboptimal local extrema (Vidal, 1993). In order to ensure replicability, the default random seed 0 was used. Therefore, if *SSE* was optimized repeatedly for the same grid-point, the outcome remained the same.

The amoeba method was used, because it is fast. Given the exponent-pair a, b, the method first evaluates four corners of a tetrahedron (simplex) in parameter space (dimensions m_0, p, q) and successively applies reflections (moving the point with highest *SSE* through the opposite side of the tetrahedron to a point with perhaps lower *SSE*) and shrinking (zooming in to a local minimum point).

In order to avoid obviously meaningless parameter values, constraints were added to ensure a biologically reasonable initial value $m_0 > 10$ and positive parameters $p > q$.

Semi-automated optimization

In order to employ also methods developed specifically for the least squares method, an alternative approach using the Mathematica function NonlinearModelFit was used. It implements the most common methods for nonlinear regression.

The optimization loop assumed a fixed value for a, whereas b proceeded from $b = a$ to $b = 2$ with step size 0.01. Further, for each exponent $a = n \cdot 0.01$ the hitherto obtained values of $SSE_{opt}(a, b)$ were plotted. If the plot showed a U-shape, then a minimum of SSE could be identified on the line $a = n \cdot 0.01$, $b > a$; otherwise (human intervention) more values of b were added to the loop until a U-shape could be recognized. (This approach assumed that then for still larger exponents b the fit could only become worse. This assumption was corroborated by the initial screening.)

The optimization started at $a = 0$, $b = 0.01$ with initial values for m_0, p and q explained above. For the subsequent computations, where a was kept fixed and b moved, the iterative optimization at the next b, namely at $b + 0.01$, started with the optimal parameters from the previous optimization (for b).

However, in order to ensure convergence (and an empirically meaningful outcome), SSE was minimized subject to certain constraints ($m_0 > 10$ and $q > 0$), whence many common methods from regression analysis (e.g. Levenberg-Marquardt algorithm) were not applicable. Instead, an interior point method was used. These methods (e.g. barrier methods initially developed in the 1960s) became popular in 1984, when an interior point method (Karmakar, 1984) solved linear optimization problems in polynomial time; Forsgen, Gill & Wright (2002) refer to the 'interior point revolution'. This setting was also advantageous for the present problem.

Custom-made simulated annealing

Based on this preparatory work, the $SSE_{opt}(a, b)$could be evaluated for almost all grid points. In order to improve the estimates of SSE at the best fitting grid points and to move from there to the optimal exponent-pair (no longer a grid-point), the authors used Simulated Annealing. However, rather than using the general purpose method employed by Mathematica, the authors developed a custom-made approach. The main difference was the use of a (sort of) geometric Brownian motion. (For each step, rather than adding a small random number to the parameters, they were multiplied by a random number, whence positive values were retained.) The optimization used a loop with 500,000 steps: It started with the parameter values obtained from the preparatory optimization steps.

RESULTS

Table 2 illustrates the results by means of the optimal parameters for three exponent-pairs, Bertalanffy, logistic, and the optimal pair (of Gompertz-type). As Figure 2 visualizes, all exponent-pair provided a reasonable fit. The optimization aimed at improvements of SSE_{opt} = 23,709 for the Bertalanffy-pair, which was obtained in the initial round of optimizations.

At first SSE_{opt} was evaluated at grid-points with exponent-pairs $0 \leq a \leq 1$ and $a < b \leq 1.5$, for growth functions (1) and with exponent-pairs $0 \leq a = b \leq 1$ of (2). The exponent-pairs were grid-points at distance 0.1 between successive grid-points. For each grid point the better of the outcomes from Simulated Annealing and from the amoeba method was used; $SSE_{opt}(0.7, 0.7)$ = 21,310 was optimal. However, the initial optimization became problematic for $b > 1.2$ and did not allow to decide, if optimization would require a search in the problematic region. Further, it could not be decided, if the optimum would be located on or above the diagonal.

The systematic search (semi-automated data fitting) was confined to equation (1). It used a fine grid (distance 0.01 between successive exponent-pairs), aiming at identifying for each exponent a with $0 \leq a \leq 1$ an exponent $b > a$ with minimal SSE. (It was sufficient to screen exponents $b \leq 2$.) The improved accuracy of this search was demonstrated for the Bertalanffy exponent-pair; lower $SSE_{opt}(0.67,1)$ = 23,534.6.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27152v1 | CC BY 4.0 Open Access | rec: 28 Aug 2018, publ: 28 Aug 2018

Figure 3 plots the outcome from the optimization at 14,282 grid points. The black dots indicate, for each exponent *a*, for which exponent *b* the value of *SSE* was minimal. Thereby, SSE_{opt}(0.67, 0.7) = 21,287.1 was the least observed optimized *SSE* for equation (1). This demonstrates that optimization showed the following pattern: For *a* = 0 the minimum *SSE* was reached close to *b* = 2. For the following values there was a distinct U-shape to be observed till *a* = 0.67. Finally, the optimum was attained close to the diagonal *a* = *b* (dots moving upwards), but the optimum value was increasing compared to the previous ones.

This pattern supported the hypothesis that the optimal *SSE* would be attained within the search region or at a diagonal point *a* = *b* on its boundary. However, the computations did not allow to decide, whether the global minimum of *SSE* was attained for *b* > *a*, i.e. for equation (1), or for *b* = *a*, i.e. equation (2). Further, optimization proceeded smoothly till *a* = 0.7, but for larger exponents optimization became increasingly more difficult and fewer results could be accepted. In particular, grid points near the diagonal were problematic.

These issues were tackled in the final step using a global optimization. It started with the near-optimal parameters found previously. For equation (1), starting from *a* = 0.68 and *b* = 0.69, the least SSE_{opt}(0.666703,0.705181) = 21,287.5 was achieved. However, for equation (2), i.e. on the diagonal *a* = *b*, a slightly better outcome SSE_{opt}(0.686028, 0.686028) = 21,286.4 was obtained (parameters in Table 2).

The custom-made method of Simulated Annealing of this paper improved insofar upon the same method as implemented by Mathematica (which was used in the initial step), as it was more accurate. Further, despite the high number of computing steps its performance was more reliable (no unexpected computer crashes).

Summarizing, during the three optimization steps the fit achieved by the Bertalanffy exponent-pair (a = 2/3, b = 1 with SSE_{opt} = 23,709) could be substantially improved. The first round of optimization identified a better exponent-pair (a = b = 0.7 with SSE_{opt} = 21,310). The second round of optimizations, using more accurate computations, found a still better exponent pair (a = 0.67, b = 0.7 with SSE_{opt} = 21,287.1). The final round of optimization converged to the minimal SSE_{opt} = 21,286.4 at a = b = 0.686028. Thus, by using different exponent-pairs and also by using more accurate optimization methods, SSE_{opt} could be reduced by 10% from the initial estimate using the Bertalanffy pair.

Discussion

The problem of the paper asks, if the 10% reduction of SSE_{opt} achieved by optimization the two exponents (in addition to the other parameters) was enough to reduce variability considerably. Thereby variability was defined with respect to the Akaike' (1974) information criterion *AIC* and the Akaike weight (Renner-Martin et al., 2018); higher variability meant an acceptable Akaike weight (2.5%or higher) for more models (i.e. more grid-point exponents). Specifically, the paper used an index AIC_c for small sample sizes (Burnham & Anderson, 2002; Motulsky & Christopoulos, 2003); for a discussion of alternative information measures c.f. Dziak et al. (2017). AIC_c was defined from the least sum of squared errors for the model, *SSE*(model), from the number N = 13 of data-points (size of Table 1 rather than the number of fish), and from the number K of optimized parameters:

$$(6)\ AIC_c(\text{model}) = N \ln\left(\frac{SSR(\text{model})}{N}\right) + 2 \cdot K + \frac{2 \cdot K \cdot (K+1)}{N-K-1}$$

$$(7)\ prob(\text{model}) = \frac{e^{-\Delta/2}}{1+e^{-\Delta/2}},\ \text{where } \Delta = AIC(\text{model}) - AIC(\text{best fitting model}) > 0$$

The Akaike weight *prob* compares a model with the best fitting model (least AIC_c): Its Akaike weight *prob*(model) is the probability that this model is true (assuming that one of the two models is true); the maximal Akaike weight is 50%. (This interpretation is based on the assumption of normally distributed errors. As the data were average values of large samples, this assumption was justified.)

Technically, the application of the above criteria required that several distinctions were made: First, the differential equations (1) and (2) that set the general framework for this study need to be distinguished from the different growth models that may or may not assume specific values for the exponent-pair. Thereby, each grid point defined a concrete model of type (1) with an assumed exponent-pair (a, b); e.g. logistic model with (a, b) = (1, 2). The (other) model parameters (m_0, p, q) were optimized (data fitting). However, the third round of optimization in addition sought for optimal exponents, referring to the general Bertalanffy-Pütter model and the general Gompertz model, respectively. Second, *SSE* and *AIC* are different ways to compare models, whereby *SSE* is used, when the focus is on the predictive power. The results pertain to the optimization of *SSE* alone. *AIC* is used, when both the simplicity of the model and its explanatory power are needed. Thereby, the *AIC* of models with assumed exponent-pairs was computed with $K = 4$ (as implicitly also *SSE* was optimized: shape parameter of the assumed normal distribution of the residuals). The *AIC* of the general Bertalanffy-Pütter model and the general Gompertz model was computed with $K = 6$ and $K = 5$, respectively, as also the exponents were optimized. Consequently, the best fitting model (least *SSE*) could have a higher (worse) *AIC* than other models. Third, for this paper the Akaike weights were interpreted in two ways. If the *AIC* was computed with the correct number of parameters, the Akaike weights were probabilities about the truth of a model. However, the paper used the Akaike weights also with an incorrect number of parameters, assuming $K = 4$ for all models; i.e. also the models with optimal exponents were treated as if these exponents were given in advance. For this application, the Akaike weight was merely a measure of the good fit (low *SSE*) that was comparable across different data-sets, but not a probability of truth.

Keeping these caveats in mind, the answer to the initial question about variability was negative, as shown by Figure 4 (all Akaike weights computed with $K = 4$): Amongst models defined by exponent pairs with $b = 1$, the comparison with the best-fitting model did affect the Akaike weights only slightly. For instance, for the Bertalanffy pair the Akaike weight was reduced from 36% (comparison with the optimal exponent a, assuming $b = 1$) to 34% (comparison with the best-fitting exponent-pair). For lower Akaike-weights the reduction was even smaller, whence the Akaike-weights could not be pushed below the 2.5% threshold.

Figure 5 illustrates, how the variability extended into two dimensions. The green area represents exponent-pairs, whose *AIC* was below the *AIC* of the best-fitting model. (Thereby, *AIC* for given exponent-pairs was computed with $K = 4$, while the *AIC* for best fitting Gompertz-type model was computed with $K = 5$, whence there was a penalty.) The red area represents additional exponent-pairs, whose fit was deemed as acceptable in the meaning above (Akaike weight of 2.5% or higher, using $K = 4$ also for the best fitting model).

The following examples illustrate these concepts. In Figure 2, the best fit was achieved by the optimal exponent-pair, followed by the Bertalanffy-pair, while logistic growth was worst. However, owing to the penalty in the definition of *AIC* for using more parameters, the Bertalanffy exponent-pair was in the green region of Figure 5. Therefore, when choosing between the Bertalanffy and the best fitting exponent-pair, the *AIC*-criterion would recommend to select the former one. By contrast, the logistic exponent-pair was outside the red or green regions of Figure 5, whence this fit was deemed as not acceptable. Summarizing, when comparing these exponent pairs, the Bertalanffy-pair would be selected; the logistic pair would be refuted due to its poor fit; and the optimal pair would be refuted, as the 10% reduction of *SSE* did not justify the optimization of an additional parameter.

Figure 6 indicates that equation (1) may indeed result in overfit due to the optimization of too many parameters. Using model (2) together with the optimal exponent, it plots the region of the 'other parameters' (m_0, p, q), where *SSE* was bounded by 10^7 (ca. 500 times the least *SSE*). Despite this large *SSE*, the region was extremely thin, suggesting some relation between the parameters.

Conclusion

The paper conducted a case study about the variability of the Bertalanffy-Pütter exponent-pairs (a, b) for fish. It was based on mass-at-age data of Walleye. For the case $b = 1$ it is known that for most fish-data any exponent $0 \leq a < 1$ could be used to model growth without affecting the fit to the data significantly (when the other parameters p, q, m_0 were optimized). In two dimension it was no longer true that any exponent-pair could provide an acceptable fit. For instance, the logistic growth function provided a reasonable fit to the data, if only a visual inspection was used, but in quantitative terms (Akaike weight), its fit was not acceptable in comparison to the optimal model. However, the paper showed that variability extended insofar into two dimensions, as it identified a large region of exponents with acceptable fit, including the Bertalanffy exponent-pair ($a = 2/3$, $b = 1$). Summarizing, the paper did not find a reason, why fishery management should deviate from its established practice to describe growth in term of the Bertalanffy models (VBGF for length, the Bertalanffy exponent-pair for mass).

However, a closer look at the structure of the set of optimal parameters indicated the potential for further research into the Bertalanffy-Pütter model, as for the best fitting parameters there seemed to exist additional relations suggesting that optimization might be further constrained by some functional relationship between the parameters (a, b, m_0, p, and q). Thus, the authors conjecture that a subclass of the Bertalanffy-Pütter model using fewer parameters may provide the same fit and therefore suffice for the modeling of growth. There remains the problem to find such a subclass that in addition is empirically meaningful.

REFERENCES

Akaike, H. (1974). A New Look at the Statistical Model Identification. IEEE Trans. Automatic Control, 19, 716-723.

Apostolidis, C.; Stergiou, K.I. (2014). Estimation of growth parameters from published data for several Mediterranean fishes. Journal of Applied Ichthyology, 30, 189-194.

Bertalanffy, L.v. (1957). Quantitative laws in metabolism and growth. Quarterly Reviews of Biology, 32, 217-231.

Burden, R.L.; Faires, J.D. (1993). Numerical Analysis, PWS Publishing Co. Boston, MA.

Burnham, K.P.; Anderson, D.R. (2002). Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach. Springer, Berlin.

Cailliet, G.M.; Smith, W.D.; Mollet, H.F.; Goldman, K.J. (2006). Age and growth studies of chondrichthyan fishes: The need for consistency in terminology, verification, validation, and growth function fitting. Environmental Biology of Fish, 77, 211-228.

Carvalho, E.P.d.; Santoro, K.R. (2007). On Starting Values for Parameters of Nonlinear Growth Functions with Application in Animal Science. Journal of Concrete and Applicable Mathematics, 5, 133-142.

Casella, G.; Berger, R.L. (2001). Statistical Inference. Thomson Learning, Stamford.

Cedersund, G.; Samuelsson, O.; Ball, G.; Tegner, J.; Gomez-Cabrero, D. (2015). Optimization in Biology. Parameter Estimation and the Associated OptimizationProblem. In Gomez-Cabrero, D.; Geris, L. (editors). Uncertainty in Biology, Springer Verlag, Heidelberg, Germany.

Dziak, J.J.; Coffman, D.L.; Lanza, S.T., Li, R. (2017). Sensitivity and specificity of information criteria. PeerJ Preprints, 5, published online DOI: 10.7287/peerj.preprints.1103v3.

Forsgen, A.; Gill, P.E.; Wright, M.H. (2002). Interior point methods for nonlinear optimization. SIAM Review, 44, 525-597

Froese, R.; Pauly, D. (2000). FishBase: concepts, design and data sources. ICLARM, Los Baños, Laguna, Philippines. 344 p.

Froese, R.; Pauly, D. (2018). FishBase data base: www.fishbase.org, last accessed in June 2018.

Gompertz, B. (1832). On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. Phil. Transactions Royal Society London, 123, 513-585.

Karmakar, N. (1984). A new polynomial-time algorithm for linear programming. Combinatorica, 4, 373-395.

Killen, S.S., Atkinson, D., Glazier, D.S. (2010). The intraspecific scaling of metabolic rate with body mass in fishes depends on lifestyle and temperature. Ecology Letters, 13, 184-193.

Knight, W. (1968). Asymptotic growth, an example of nonsense disguised as mathematics. Journal of the Fisheries Research Board of Canada, 25, 1303-1306.

Leader, J. J. (2004). Numerical Analysis and Scientific Computation. Addison-Wesley, Boston.

Manabe, A.; Yamakawa, T.; Ohnishi, S.; Akamine, T.; Narimatsu, Y.; Tanaka, H.; Funamoto, T.; Ueda, Y.; Yamamoto, T. (2018). Anovel growth function incorporating the effects of reproductive energy allocation. PLoS ONE, 13, published online, DOI: 10.1371/journal.pone.0199346.

Marusic, M.; Bajzer, Z. (1993). Generalized two-parameter equations of growth. Journal of Mathematical Analysis and Applications, 179, 446-462.

Mildenberger, T.K.; Taylor, M.H.; Wolff, M. (2017). TropFishR: anR-package for fisheries analysiswithlength-frequency data. Methods in Ecology and Evolution, doi: 10.1111/2041-210X.12791.

Motulsky, H.; Christopoulos, A. (2003). Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting. Oxford Univ. Press, Oxford.

Ogle, D. (2018). R for Fisheries Analysis, derekogle.com, last accessed in June 2018.

Ogle, D.; Iserman, D.A. (2017). Estimating Age at a Specified Length from the von Bertalanffy Growth Function. North American Journal of Fisheries Management, 37, 1176-1180.

Ohnishi, S.; Yamakawa, T; Akamine, T. (2014). On the analytical solution for the Pütter-Bertalanffy growth equation. J. Theoretical Biology, 343, 174-177.

Pütter, A. (1920). Studien über physiologische Ähnlichkeit. VI. Wachstumsähnlichkeiten. Pflüg. Archiv der Gesamten Physiologie, 180, 298-340.

Renner-Martin, K.; Brunner, N.; Kühleitner, M.; Nowak, W.G.; Scheicher, K. (2018). On the exponent in the Von Bertalanffy growth model. PeerJ, 6, published online, DOI: 10.7717/peerj.4205.

Ricard, F.H. (1975). Essai de selection sur la forme de la courbe de croisance chez Ie poulet. Annales Genet. Select. Anim., 7, 427-443.

Richards, F.J. (1959). A Flexible Growth Function for Empirical Use, Journal of Experimental Botany, 10, 290-300.

Shi, P.-J.; Ishikawa, T.; Sandhu, H.S.; Hui, C.; Chakraborty, A.; Jin, X.-S.; Tachihara, K.; Li, B.-L. (2014). On the 3/4-exponent van Bertalanffy equation for ontogenetic growth. Ecological Modelling, 276, 23-28.

Strathe, A.B.; Danfær, A.; Sørensen, H.; Kebreab, E. (2010). A multilevel nonlinear mixed-effects approach to model growth in pigs. Journal of Animal Science, 88, 638-649.

Verhulst, P.F. (1838). Notice sur la loi que la population suit dans son accroissement, Correspondence Mathematique et Physique (Ghent), 10, 113-121.

Vidal, R.V.V. (1993). Applied simulated annealing. Lecture notes in economics and mathematical systems. Springer-Verlag, Berlin.

West, G.B.; Brown, J.H.; Enquist, B.J. (2001). A general model for ontogenetic growth. Nature, 413, 628-631.

Wilson, K.L.; Matthias, B.G.; Barbour, A.B.; Ahrens, R.N.M.; Tuten, T.; Allen, M.S. (2015). Combining samples from multiple gears helps to avoid fishy growth curves. North American Journal of Fisheries Management, 35, 1121-1131.

White, C.R. (2010). Physiology: There is no single p. Nature, 464, 691-693.

Table and Figure Captions

Table 1. Average weight-at-age (rounded) for male Walleye, based on ca. 20,000 age-weight data points (rounded to one decimal for the ease of presentation; the computations of the paper used data rounded to three decimals).

Table 2. Optimal parameters for selected models.

Figure 1. Weight-at-age and average weight (red dots) of male Walleye from Lake Erie.

Figure 2. Comparison with the data of the growth curve using the Bertalanffy exponent-pair (red), the logistic exponent pair (blue) and of the best fitting growth curve (black); parameter values as in Table 2.

Figure 3. Contour plot of the optimal *SSE* on a grid of exponent-pairs with distance 0.01 between adjacent points and for each exponent a, plot of the exponent-pair with smallest *SSE* (black dots).

Figure 4. Plot of the Akaike weights for exponent-pairs with $b = 1$, using the least *AIC* amongst generalized Bertalanffy-models (red) and the least *AIC* amongst all considered models (blue); all *AIC*s using $K = 4$.

Figure 5. Plot of the grid points $a < b$ with *AIC* below *AIC* of the best fitting model (green; the *AIC* of the best fitting model was higher due to the penalty for an additional parameter) and with acceptable fit (red). The Bertalanffy and the logistic exponent-pairs are displayed in yellow.

Figure 6. Plot of part of the region of exponents m_0, p, q for model (2) with the optimal exponent $a = 0.686028$, where *SSE* does not exceed 10^7.

Table 1(on next page)

Average weight-at-age (rounded) for male Walleye, based on ca. 20,000 age-weight data points (rounded to one decimal for the ease of presentation; the computations of the paper used data rounded to three decimals)

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27152v1 | CC BY 4.0 Open Access | rec: 28 Aug 2018, publ: 28 Aug 2018

Table 1

Average weight-at-age (rounded) for male Walleye, based on ca. 20,000 age-weight data points (rounded to one decimal for the ease of presentation; the computations of the paper used data rounded to three decimals)

Age (years)	Weight (gram)	Class size	Comment
0	192.1	14	
1	423.7	4009	
2	761.8	5181	
3	1018.0	3870	
4	1221.6	2262	
5	1442.8	1519	
6	1644.5	1471	
7	1802.0	690	
8	1880.7	446	
9.5	1895.3	430	classes 9+10
11	1982.6	105	
12.4	2140.4	104	classes 12+13
15.3	2228.5	65	classes 14-20

Table 2 (on next page)

Optimal parameters for selected models

* 1st and 3rd refer to the initial and final rounds of optimization

Table 2

Optimal parameters for selected models

Model	Comment*	a	b	m_0	p	q	SSE
Bertalanffy	1st (a, b given)	2/3	1	203.8	11.2	0.86	23,709
logistic	1st (a, b given)	1	2	301.716	0.528051	0.000253611	72,283
optimal	3rd (a optimized)	0.686028	= a	175.67	21.3148	2.76054	21,286

* 1st and 3rd refer to the initial and final rounds of optimization

Figure 1

Weight-at-age and average weight (red dots) of male Walleye from Lake Erie

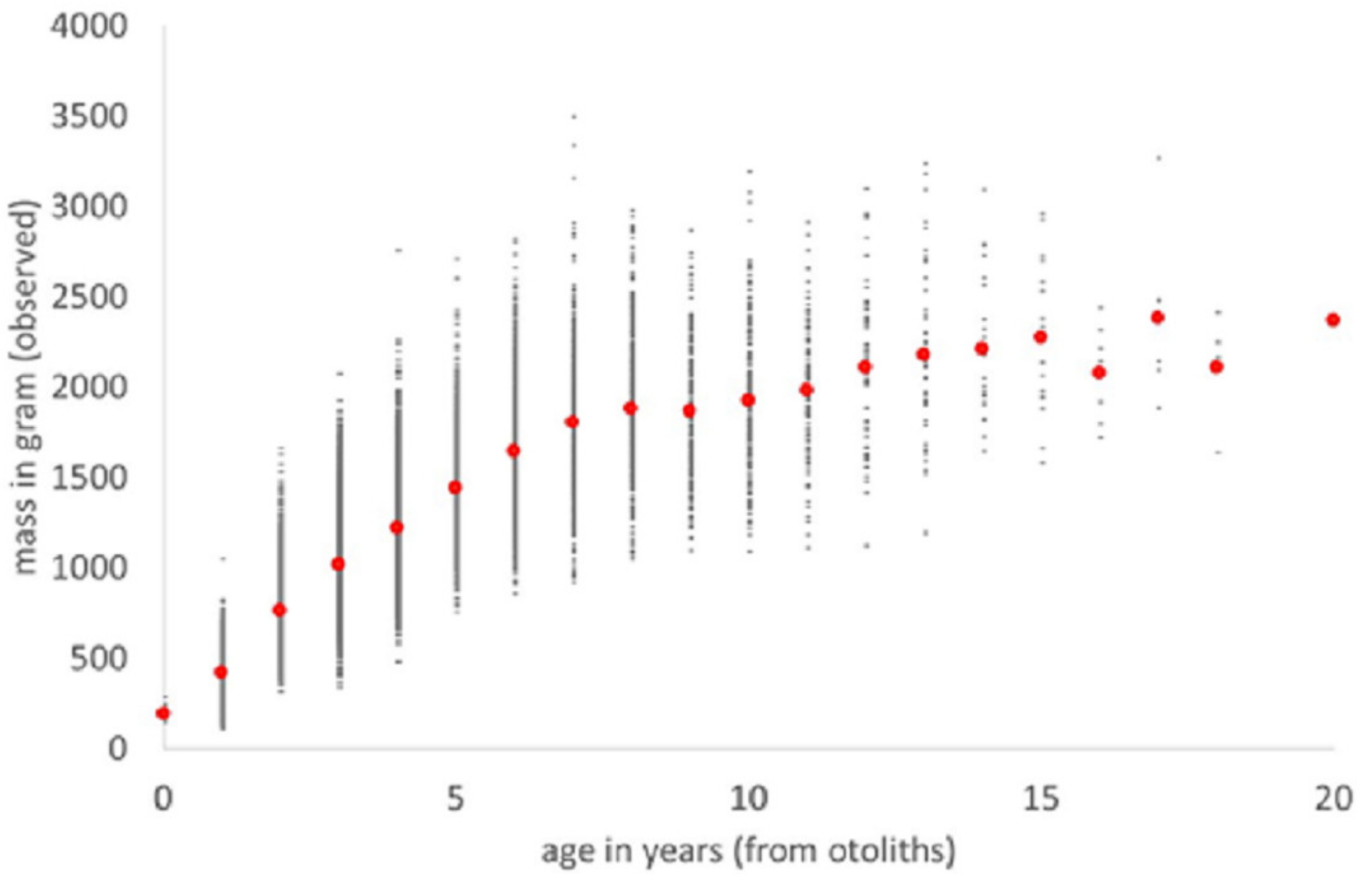
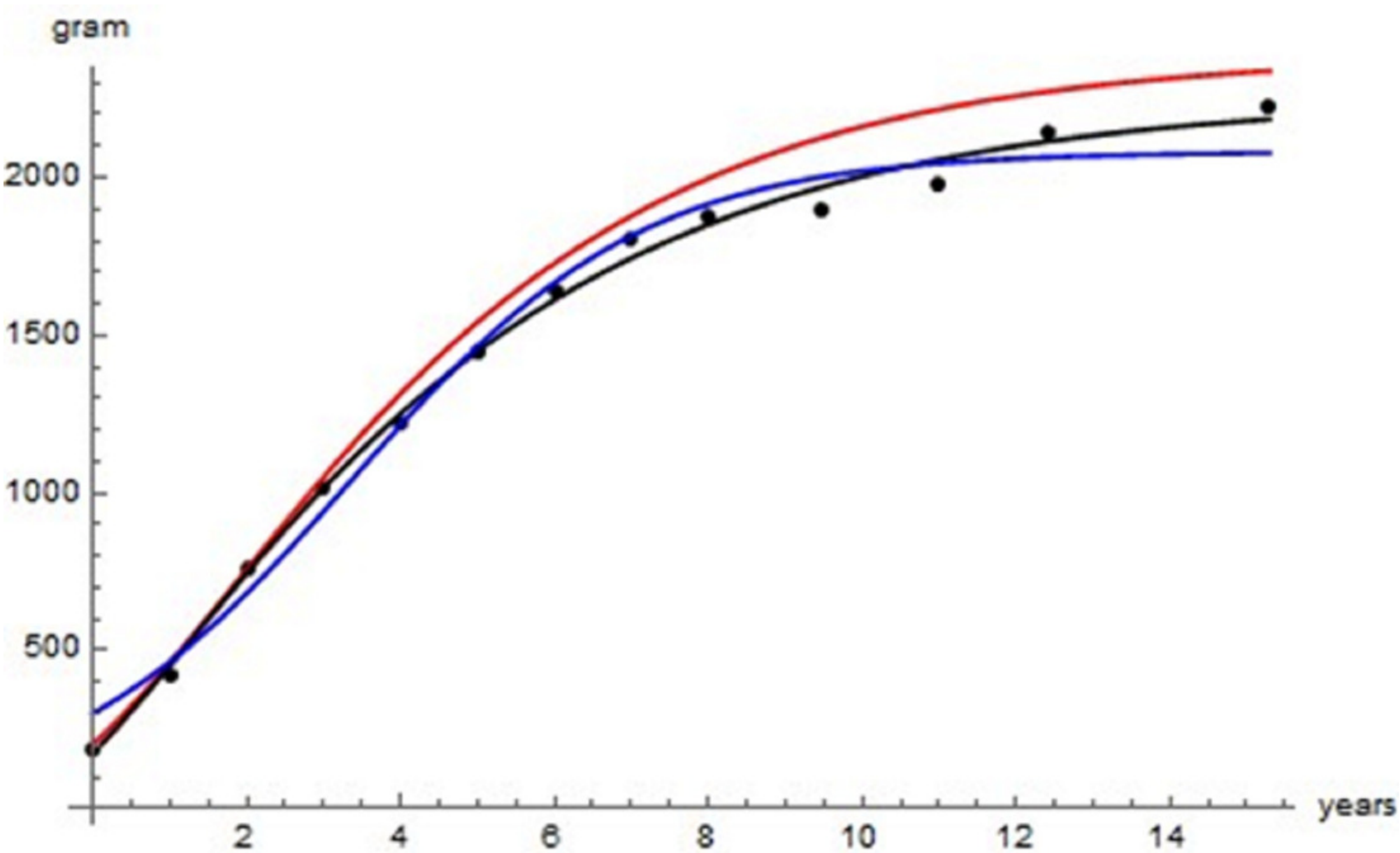


Figure 2

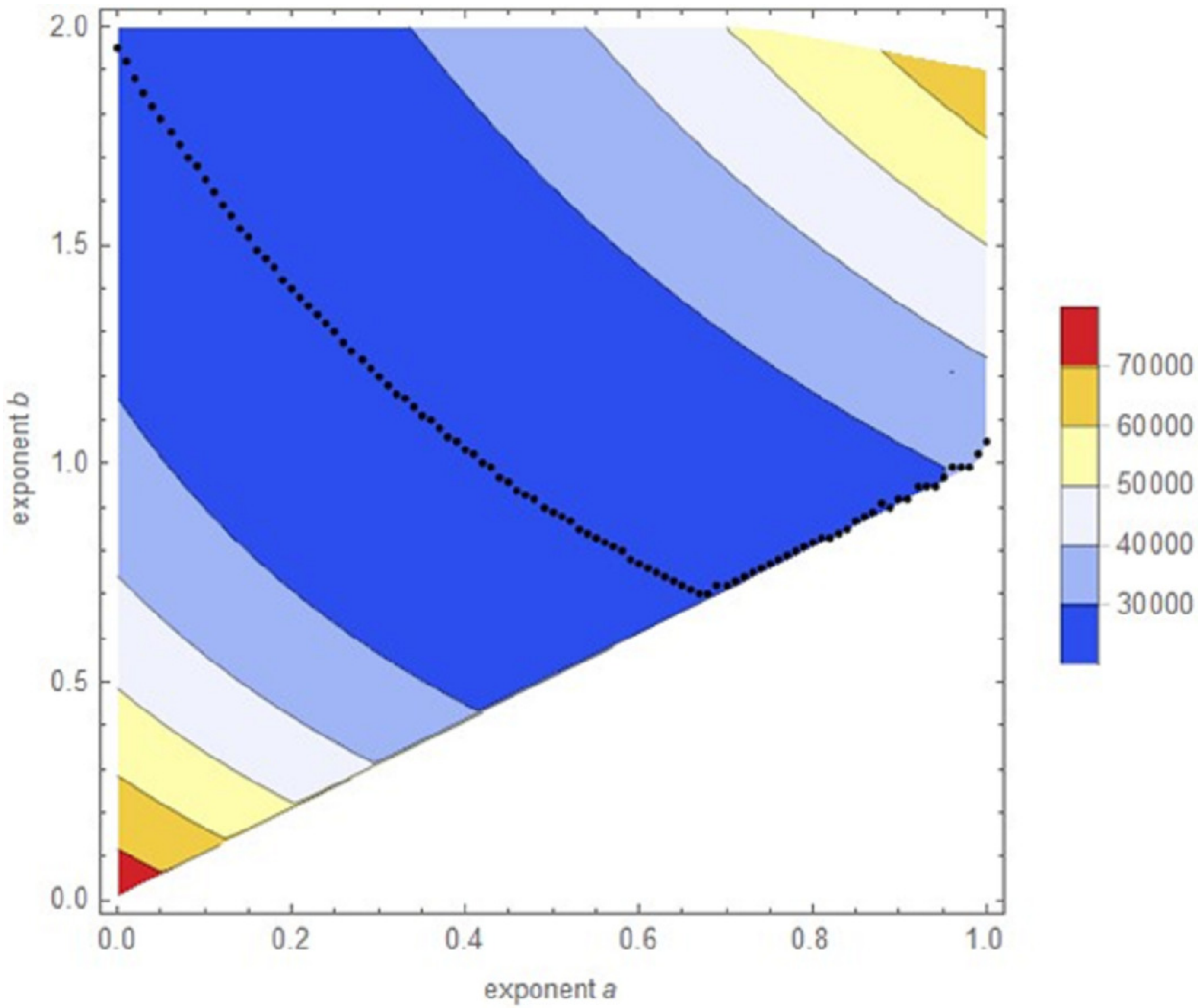
Comparison with the data of the growth curve using the Bertalanffy exponent-pair (red), the logistic exponent pair (blue) and of the best fitting growth curve (black); parameter values as in Table2.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27152v1 | CC BY 4.0 Open Access | rec: 28 Aug 2018, publ: 28 Aug 2018

Figure 3

Contour plot of the optimal SSE on a grid of exponent-pairs with distance 0.01 between adjacent points and for each exponent a, plot of the exponent-pair with smallest SSE (black dots).



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27152v1 | CC BY 4.0 Open Access | rec: 28 Aug 2018, publ: 28 Aug 2018

Figure 4

Plot of the Akaike weights for exponent-pairs with b = 1, using the least AIC amongst generalized Bertalanffy-models (red) and the least AIC amongst all considered models (blue); all AICs using K = 4

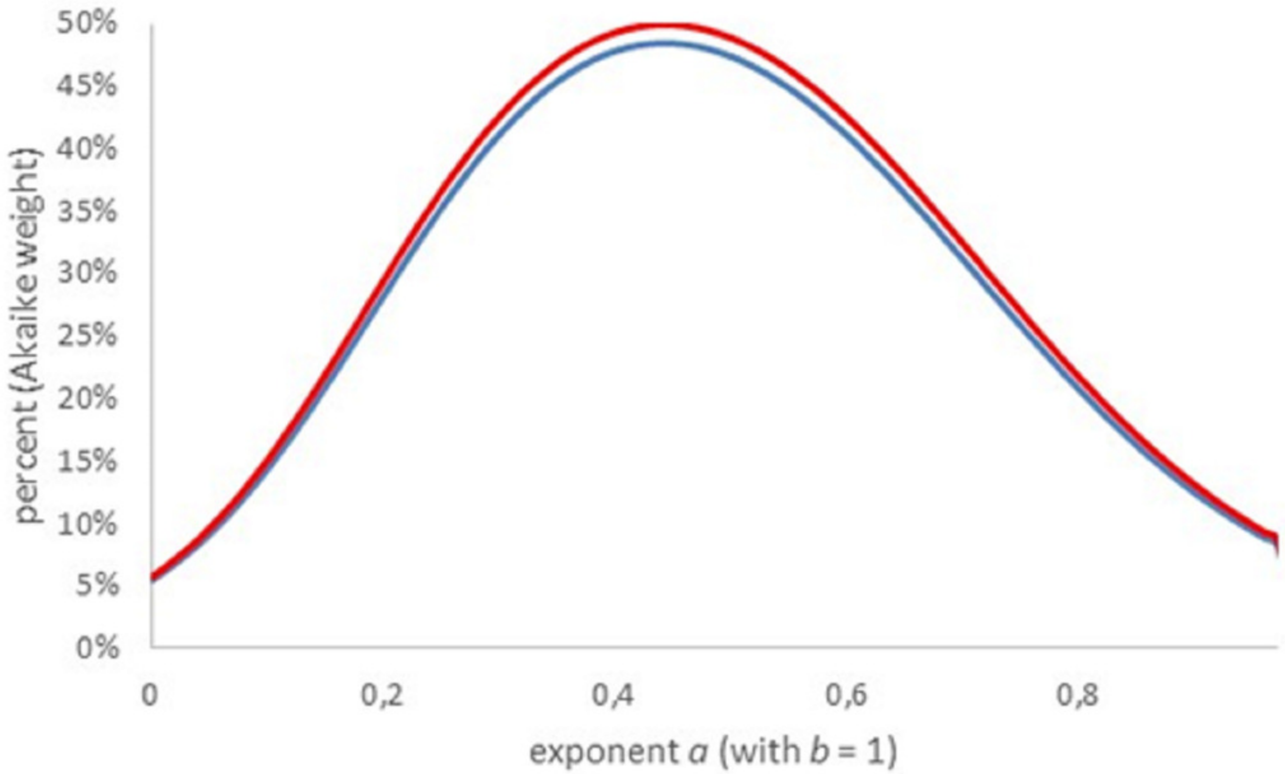


Figure 5

Plot of the grid points a < b with AIC below AIC of the best fitting model (green; the AIC of the best fitting model was higher due to the penalty for an additional parameter) and with acceptable fit (red). The Bertalanffy and the logistic exponent-pairs

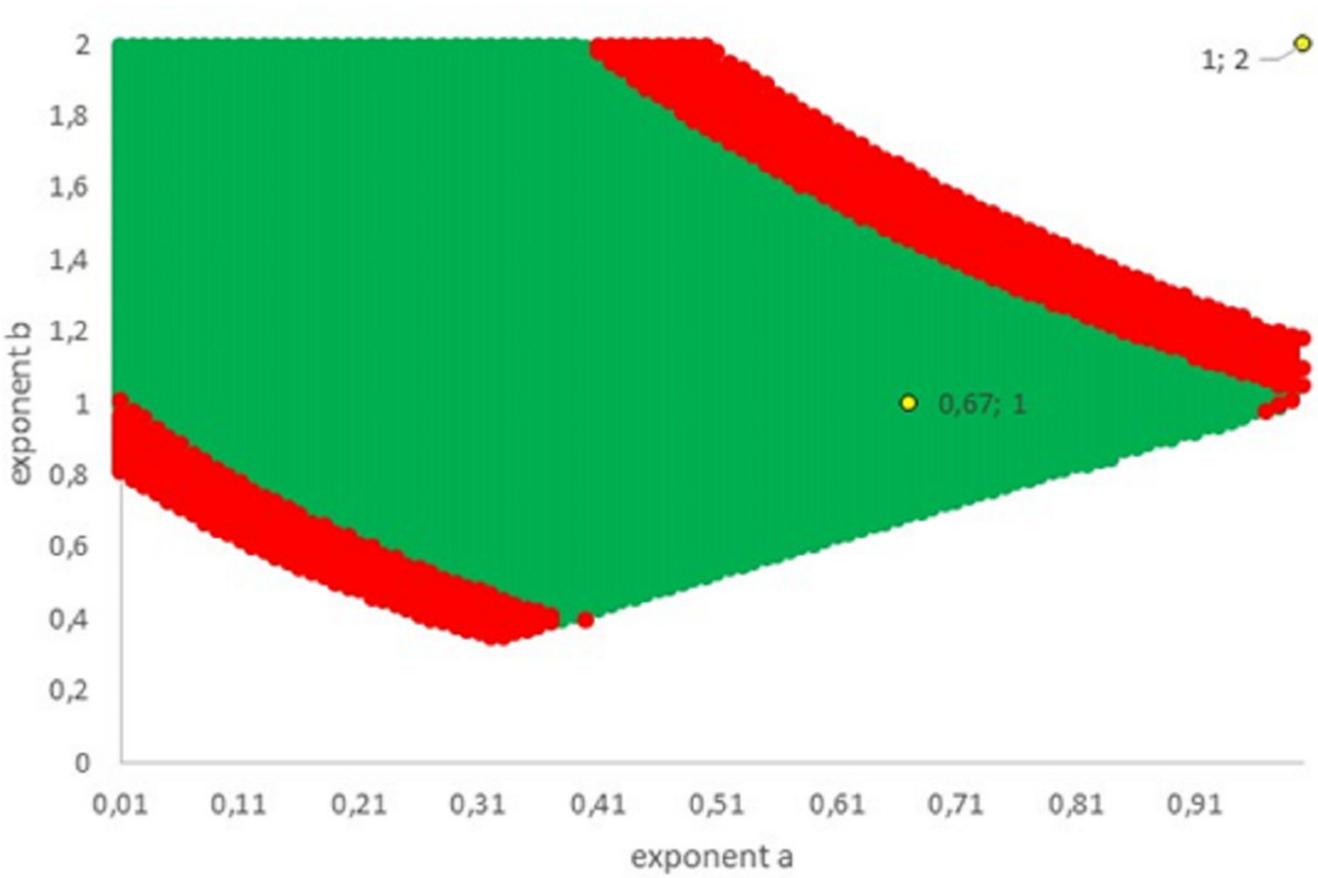
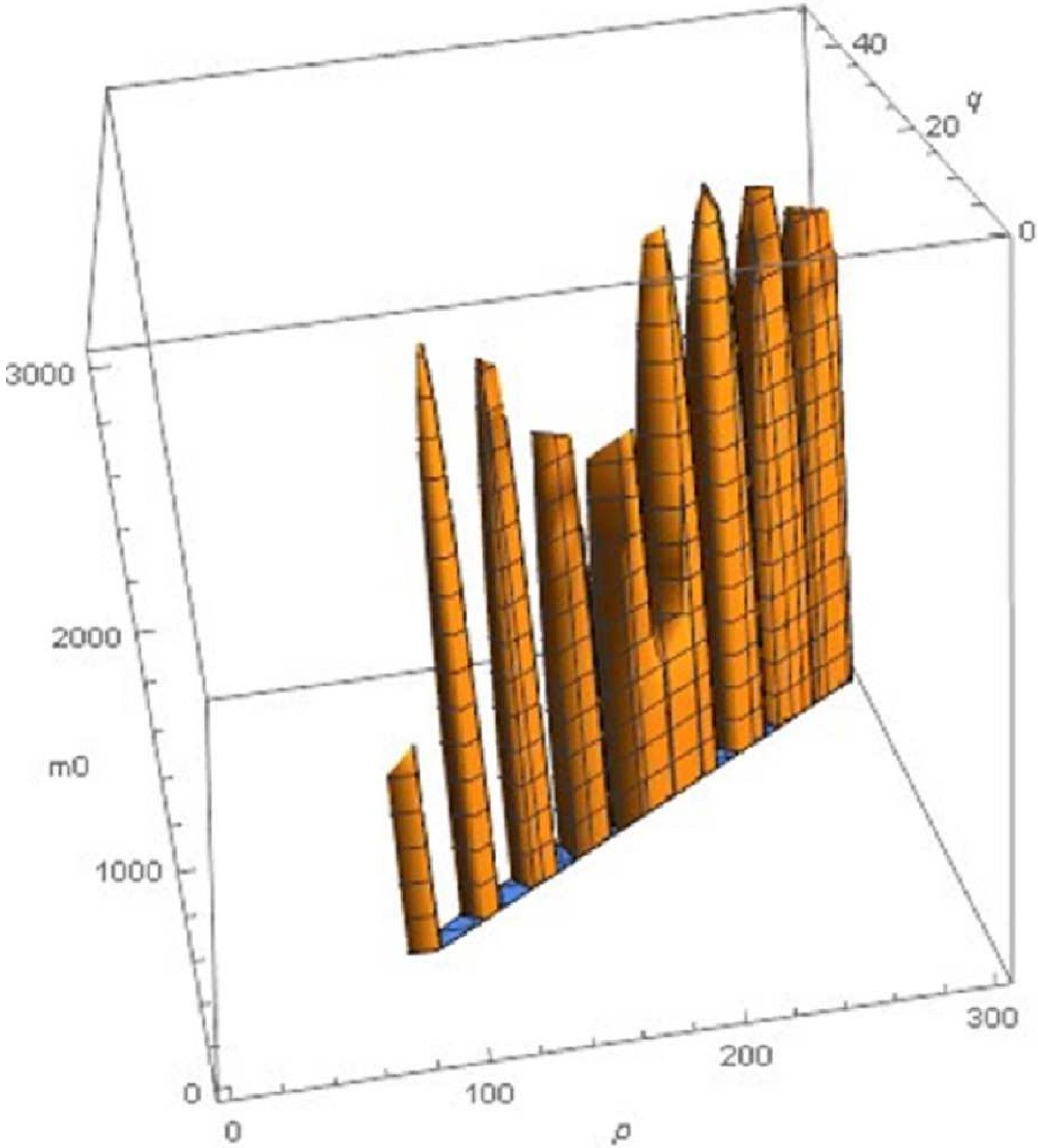


Figure 6

Plot of part of the region of exponents m_0 p, q for model (2) with the optimal exponent a = 0.686028, where SSE does not exceed 10^7.



40
q
20
0
3000
2000
m0
1000
0
0
100
200
300
ρ