Trade-offs between accuracy and interpretability in von Bertalanffy random-effects models of growth

SIMONE VINCENZI,[1,2,7] ALAIN J. CRIVELLI,[3] STEPHAN MUNCH,[4] HANS J. SKAUG,[5] AND MARC MANGEL[1,6]

[1]*Center for Stock Assessment Research, Department of Applied Mathematics and Statistics, University of California, Santa Cruz, California 95064 USA*
[2]*Dipartimento di Elettronica, Informazione e Bioingegneria Politecnico di Milano, Via Ponzio 34/5, I-20133 Milan, Italy*
[3]*Station Biologique de la Tour du Valat, Le Sambuc F-1320 France*
[4]*Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 110 Shaffer Road, Santa Cruz, California 95060 USA*
[5]*Department of Mathematics, University of Bergen, Box 7800, 5020 Bergen, Norway*
[6]*Department of Biology, University of Bergen, 5020 Bergen, Norway*

Abstract. Better understanding of variation in growth will always be an important problem in ecology. Individual variation in growth can arise from a variety of processes; for example, individuals within a population vary in their intrinsic metabolic rates and behavioral traits, which may influence their foraging dynamics and access to resources. However, when adopting a growth model, we face trade-offs between model complexity, biological interpretability of parameters, and goodness of fit. We explore how different formulations of the von Bertalanffy growth function (vBGF) with individual random effects and environmental predictors affect these trade-offs. In the vBGF, the growth of an organism results from a dynamic balance between anabolic and catabolic processes. We start from a formulation of the vBGF that models the anabolic coefficient (q) as a function of the catabolic coefficient (k), a coefficient related to the properties of the environment (γ) and a parameter that determines the relative importance of behavior and environment in determining growth (ψ). We treat the vBGF parameters as a function of individual random effects and environmental variables. We use simulations to show how different functional forms and individual or group variability in the growth function's parameters provide a very flexible description of growth trajectories. We then consider a case study of two fish populations of *Salmo marmoratus* and *Salmo trutta* to test the goodness of fit and predictive power of the models, along with the biological interpretability of vBGF's parameters when using different model formulations. The best models, according to AIC, included individual variability in both k and γ and cohort as predictor of growth trajectories, and are consistent with the hypothesis that habitat selection is more important than behavioral and metabolic traits in determining lifetime growth trajectories of the two fish species. Model predictions of individual growth trajectories were largely more accurate than predictions based on mean size-at-age of fish. Our method shares information across individuals, and thus, for both fish populations investigated, allows using a single measurement early in the life of individual fish or cohort to obtain accurate predictions of lifetime individual or cohort size-at-age.

Key words: brown trout (Salmo trutta); *longitudinal data; marble trout* (Salmo marmoratus); *model predictions; von Bertalanffy growth function; Slovenia.*

INTRODUCTION

Understanding individual growth will always be an important biological problem, as survival, sexual maturity, reproductive success, movement, and migration are commonly related to growth and body size (Peters 1983). Thus, variation in growth can have substantial consequences for ecological and evolutionary dynamics (Lomnicki 1988, Pelletier et al. 2007, Coulson et al. 2010).

Experimental and observational studies provide information on growth throughout an individual's lifetime or at specific life stages. However, a series of data points on size-at-age is difficult to interpret without reference to a model of growth, but nonlinear growth models allow us to condense the information contained in such a data series into a few parameters. In some growth models, those parameters are biologically interpretable, in the sense that they represent or summarize the most relevant biological processes and environmental factors determining variation in growth (West et al. 2004), while parameters of other growth models do not have a clear

mechanistic interpretation and are best considered as curve-fitting parameters (Appendix S1: Text S1).

Growth models have multiple applications in ecology and evolutionary biology. For instance, when managing human intervention in natural populations, we may be interested in:

- Understanding how growth rates and size-at-age vary in time and space depending on environmental conditions within and among populations of the same species (Vincenzi et al. 2014b).
- Inferring life-history strategies, that is, trade-offs between allocation of resources to competing physiological functions such as growth, maintenance, and reproduction throughout a lifetime (Roff 2007).
- Estimating heritability of growth and size-at-age (Carlson and Seamons 2008).

Another potential application of growth models (e.g., for fisheries management) is the prediction of lifetime growth trajectories of individuals or group of individuals (e.g., year-of-birth cohort, same-sex individuals) from observations at early life stages. Growth models have mostly been used for describing or interpreting population and individual processes, but have seldom been used for predictive purposes in ecology (Peters 1991), although an ample literature exists for health applications in humans (Berkey 1982, Radhakrishna Rao 1987, Shohoji et al. 1991, Roland et al. 2011).

For both basic and applied ecological goals, the choice of the growth model is often critical. Across growth models, we often face trade-offs between model complexity, biological interpretability of model parameters, ease of parameter estimation, and model accuracy (i.e., the combination of goodness of fit and predictive power). These trade-offs are commonly faced in other ecological contexts (Ludwig and Walters 1985, Adkison 2009). For instance, Ward et al. (2014) tested the predictive performance of short-term forecasting models of population abundance of varying complexity. They found that more complex parametric and nonparametric models often performed worse than simpler models, which simply treated the most recent observation as the forecast. In their case, the estimation of even a small number of parameters imposed a high cost while providing little benefit for short-term forecasting of species without obvious cyclic dynamics. However, when a signal of cyclic dynamics was clearly identifiable, more complex models were able to extract meaningful patterns from data and more accurately predict future abundances. Thus, the complexity of the best predictive model will be determined by the ecological situation.

A broad range of models describing the variation in size of organisms throughout their lifetime have been proposed (von Bertalanffy 1957, Lester et al. 2004, Kimura 2008, Quince et al. 2008, Kooijman 2009, Omori et al. 2009, Russo et al. 2009), with varying degrees of model complexity, biological interpretability of parameters, and data requirements for parameter estimation. For some growth models, parameters may or may not be biologically interpretable depending on model formulation. For instance, the parameters of the widely used von Bertalanffy growth function (vBGF; von Bertalanffy 1957) to model growth of organisms may be considered either curve-fitting parameters with no biological interpretation (i.e., providing just a phenomenological description of growth) or parameters that describe how anabolic and catabolic processes govern the growth of the organism (i.e., mechanistic description; see Mangel 2006). The classic vBGF has three parameters: asymptotic size, growth rate, and theoretical age at which size is equal to 0 (or size at age 0 in an equivalent formulation). In the original mechanistic formulation of von Bertalanffy, asymptotic size results from the relationship between environmental conditions and behavioral traits, and the growth coefficient is closely related to metabolic rates and behavioral traits (i.e., the same physiological processes affect both growth and asymptotic size). However, in the literature, asymptotic size and growth rate are commonly treated as independent parameters with no connection to physiological functions, thus becoming a phenomenological description of growth.

In the vast majority of applications of growth models, parameters are estimated at the population level, but interpreted as those of an average individual in the population. This approach fails to take into account the substantial variation in growth observed within populations, and severely limits the breadth and scope of the models (Sainsbury 1980, Siegfried and Sansó 2006, Vincenzi et al. 2014b). Individual variation in growth can arise from a variety of processes. For example, individuals within a population vary in their intrinsic metabolic rates and behavioral traits (e.g., aggressiveness or territoriality; Rosenfeld et al. 2014), which may have consequences for their foraging dynamics and access to resources. Realized growth is a combination of an individual's intrinsic growth potential, environmental conditions, intra- and interspecific competition, and stochastic events. For these reasons, the estimation of individual variation in growth is biologically and computationally difficult, and requires longitudinal data (Shelton and Mangel 2012). Random-effects models provide an intuitive framework for estimating heterogeneity of growth within and among populations, along with individual growth trajectories (Sainsbury 1980, Eveson et al. 2007, Sigourney et al. 2012).

Here, we explore how different formulations of the widely used vBGF with individual random effects offer different degrees of biological interpretability of model parameters, goodness of fit, and prediction of future growth trajectories or unobserved growth realizations. We start from the model that Snover et al. (2005) developed for management of coho salmon *Oncorhynchus kisutch*, which treats the anabolic factor in the vBGF as the product of the catabolic factor and a factor related to the properties of the environment, and show

how correlation among parameter estimates arises and how different functional forms and individual or group variability in the growth function's parameters provide a very flexible description of growth trajectories. However, flexibility comes at a cost, since it potentially reduces the biological interpretability of the parameters of the vBGF.

We use simulated data and test whether the same growth trajectories can be obtained using different formulations and parameter combinations of the vBGF. We investigate the correlation between parameter estimates, as the sign and strength of the correlation give insights on life-history strategies (Vincenzi et al. 2014b). We use one population of marble trout (*Salmo marmoratus*) and one population of brown trout (*Salmo trutta* L.) living in streams located in western Slovenia (Zakojska and Upper Volaja, respectively) as model systems for the fitting and application of the growth model.

Marble trout is a resident salmonid endemic in northern Italy and Slovenia that is at risk of extinction due to hybridization with brown trout (Vincenzi et al. 2008) and massive mortalities associated with catastrophic flood events, whose frequency is increasing (Vincenzi et al. 2014a). Brown trout was introduced in Upper Volaja in the 1920s, with no additional stocking of fish after the introduction. Growth patterns and size-at-age in salmonids contribute to determine survival (Woodson et al. 2013), sexual maturity, and reproductive success (Jonsson and Jonsson 2011), so that having a better understanding of growth has important implications for our understanding of the ecology of the two species, their population dynamics, the evolution of life-history traits, and for the effective application of conservation measures. For those populations, we test the goodness of fit of the models along with the empirical correlation between parameter estimates. Then, we test the ability of the models to predict future or unobserved size-at-age data. We finally discuss the biological interpretation of the vBGF's parameters, which are modeled as function of individual random effects and environmental predictors.

Material and methods

Growth model

We use an extension of the von Bertalanffy model (vBGF, von Bertalanffy 1957, Essington et al. 2001, Mangel 2006), which has been used to model the growth of organisms across a wide range of taxa (Kingsley 1979, Zullinger and Ricklefs 1984, Shine and Charnov 1992, Starck and Ricklefs 1998, Frisk et al. 2001, Lester et al. 2004, Tjørve and Tjørve 2010).

We start with a description of the standard vBGF and we then follow with a formulation of the vBGF that allows for a description of the growth process in terms of interaction between individual behavior and properties of the environment.

The standard von Bertalanffy growth function.—In the vBGF, the growth of an organism results from a dynamic balance between anabolic and catabolic processes (von Bertalanffy 1957), where anabolic processes are those leading to tissue growth, differentiation of cells, and increase in body size, and catabolic processes are those involving the breakdown of complex molecules and the release of energy. If $W(t)$ denotes mass at time t, the assumption of the vBGF is that anabolic factors are proportional to surface area, which scales as $W(t)^{2/3}$, and that catabolic factors are proportional to mass. If a and b denote these proportionality parameters, then the rate of change of mass is

$$\frac{dW}{dt} = aW(t)^{2/3} - bW(t). \qquad (1)$$

If we further assume that mass and length, $L(t)$, are related by $W(t) = \rho L(t)^3$, with ρ corresponding to mass per unit volume, then

$$\frac{dL}{dt} = q - kL \qquad (2)$$

where $q = a/3\rho$ and $k = b/3\rho$. In this parameterization, q is a coefficient of anabolism. The coefficient of catabolism, k, is commonly known as the von Bertalanffy growth coefficient and has the units of t^{-1}. The coefficient q, with unit size·t^{-1}, is proportional to the amount of resources available to an individual and will vary with environmental conditions and individual behavior. The asymptotic size (i.e., obtained by setting $q - kL = 0$ in Eq. 2) is $L_\infty = q/k$ and if $L(0) = L_0$ is size at age 0, we can readily solve the linear differential Eq. 2 by the method of the integrating factor. Two forms of the solution are

$$L(t) = L_\infty(1 - e^{-kt}) + L_0 e^{-kt} \qquad (3)$$

and

$$L(t) = L_\infty(1 - e^{-k(t-t_0)}) \qquad (4)$$

where t_0 is the hypothetical age at which length is equal to 0.

The vBGF model in Eq. 4 has three parameters: L_∞, k, and t_0 (in addition to the residual variance when parameter values are estimated), which are usually estimated at the population or group (e.g., cohort, sex) level. L_∞ is commonly treated as an independent parameter (i.e., not an explicit function of k, although a negative correlation between parameter estimates often emerges due to the presence of ridges in the likelihood surface, in particular when length-at-age data for older ages are relatively few or missing, see Vincenzi et al. 2014b) and the vBGF has often been used as a phenomenological and not mechanistic description of the growth process (although see Essington et al. [2001] and Temming and Herrmann [2009] for estimating consumption rates from vBGF parameter values). In this study, we will not make explicit reference to L_∞, as it does not provide any insights on the behavioral and physiological processes governing growth.
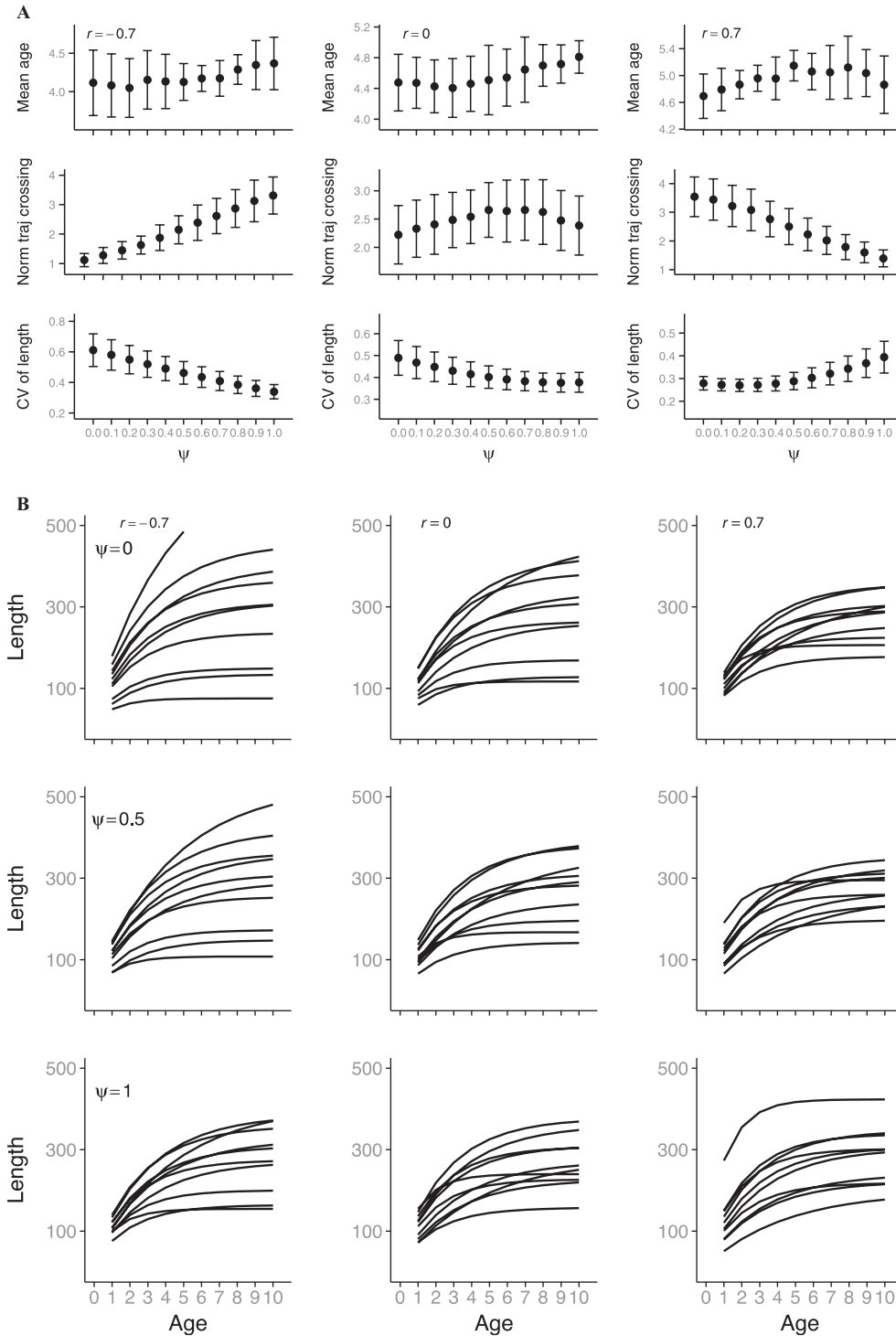
FIG. 1. (A) Mean age (yr) at growth trajectories crossing, number of trajectories crossing divided by the total number of individuals (Norm traj crossing), and coefficient of variation of length at age 10 (CV of length) for von Bertalanffy growth models as in Eqs. 7 and 8 with different values of ψ and Pearson's correlation between pairs of k_i and γ_i. Panels in the same column are for models with the same correlation between pairs of k_i and γ_i. For all models, individuals have the same asymptotic length $L_\infty = 300$ mm, $t_0 = -0.32$ y, $k_0 = 0.38$ y^{-1}, $\sigma_u = \sigma_v = 0.36$, $\gamma_0 = L_\infty (k_0^{1-\varphi})$. Vertical segments are standard deviations over 10 replicates with random drawing of individual random effects to simulate individual growth trajectories. (B) Ten growth trajectories simulated with the model in Eq. 8 and the same parameter estimates as in A. From left to right column, correlation between k_i and γ_i equal to -0.7, 0, 0.7. From top to bottom row, ψ equal to 0, 0.5, 1. See Table 1 for parameter definitions.

Model with individual variation in parameters of anabolism and catabolism.—There are biological reasons for k (the coefficient of catabolism) and q (the coefficient of anabolism) to be linked (Shelton and Mangel 2012, Shelton et al. 2013). Therefore, we turn to a model that combines individual and environmental variation and allows dependence between k and q along with individual variation, that is k and q are defined and estimated at the individual level. Since q is the coefficient of anabolism, it should be closely linked to bottom-up factors in the environment, such as food conditions (Mangel 2006). By letting q vary across individuals (q_i), we assume that "realized anabolism" may vary across individuals. The parameter k_i determines how metabolic rates scale with the size of individual i and thus relates to an individual's phenotypic capacity for growth. Snover et al. (2005, 2006) assume that k_i combines physiological and behavioral traits that determine individual activity and thus potentially affect the ability of an individual to obtain resources from the environment, although with a trade-off with energy expenditure. Thus, the "anabolic" conditions may be different for individuals and they may depend on k_i as well as some properties of the environment. Under these assumptions, we model q_i as a function of k_i, γ_i, and a parameter ψ constrained between 0 and 1 that determines the degree to which q_i depends on environmental (represented by γ_i) versus behavioral (represented by k_i) factors, and set

$$q_i = \gamma_i k_i^{\psi}. \tag{5}$$

With the formulation in Eq. 5, ψ is the result of the interaction of the environment and the foraging characteristics of the species, and may depend on patchiness of resources, fragmentation of the habitat, movement range of individuals. Units of γ depend on the value of ψ: when $\psi = 0$, γ has the units of q (size·t^{-1}); when $\psi = 1$, γ has the units of size, while units are fractal when $0 < \psi < 1$.

With Eq. 5, the expected length of individual i at age t is

$$L_i(t) = \gamma_i k_i^{(\psi-1)}(1 - e^{-k_i(t-t_0)}). \tag{6}$$

Eq. 6 is the formula describing length-at-age for individual i that we will use in this study.

1. The case with γ common and k varying among individuals.—The case with γ common and k varying among individuals has been investigated in Snover et al. (2005, 2006), and Shelton and Mangel (2012). In this case, parameters have a clear biological interpretation when (1) $\psi = 0$, (2) $\psi = 1$, and (3) ψ between 0 and 1.

When (1) $\psi = 0$, k_i has no effect on an individual's success at obtaining resources from the environment. Therefore, individuals with large k_i have lower realized growth increments, since higher activity comes at a higher energetic cost. When (2) $\psi = 1$, individuals with large k_i (i.e., aggressive and/or highly active individuals) have

greatly increased access to resources, and therefore experience faster length-specific growth. For (3) values of ψ between 0 and 1, the relative growth rate for each individuals with different k_i will change with an individual's length. Individuals with large k_i will grow faster at small lengths and individuals with small k_i will grow faster at large size; the length or age at which individuals experience equivalent growth and when growth trajectories cross depends on ψ.

2. The case with both γ and k varying among individuals.—We hypothesize that k_i and γ_i co-vary among individuals. This additional variability in γ increases the complexity of the model by increasing the number of parameters to be estimated, and allows for a greater flexibility of growth trajectories of individuals (Fig. 1). For instance, while with a common γ between 0 and 1 and only k varying among individuals all growth trajectories intersect at the same age, with both k and γ co-varying among individuals we obtain a distribution of ages at which growth trajectories intersect.

However, a biological interpretation of the two parameters k_i and γ_i, and in particular of ψ, becomes more challenging with respect to the case with only k varying among individuals. In fact, we will show that the biological interpretation of ψ depends on the sign and strength of the correlation between values of k_i and γ_i, and ψ may also be seen as a parameter giving additional flexibility to the vBGF, rather than describing the relative importance of environmental and behavioral factors in determining q_i.

As ψ is defined at the population level, but k and γ are allowed to vary among individuals, we describe potential growth trajectories by first fixing ψ (i.e., the relative importance of behavioral and environmental factors in determining anabolism) and explore how growth trajectories may change with different strength and sign of correlation between values of k_i and γ_i (Fig. 1).

When $\psi = 0$, the maintenance of size ranks through the lifetime of individuals and the mean age at which growth trajectories cross decreases going from a negative to a positive correlation between values of k_i and γ_i. In the limiting case of a correlation between values of k_i and γ_i equal to 1, growth trajectories never cross throughout the lifetime of individuals for any variance of k_i and γ_i (Appendix S1: Figs S1 and S2). When $\psi = 0$, aggressive/more active individuals (larger k_i) are always growing slower than less aggressive/active individuals. Thus, we may hypothesize that the more active individuals are either expending more energy than less active individuals without acquiring more resources, or are investing more energy on biological processes other than growth.

On the other hand, when $\psi = 1$ (i.e., asymptotic size is not an explicit function of k), the maintenance of size ranks (i.e., of size hierarchy) through an organism's lifetime and mean age at which growth trajectories cross

increases going from a negative to a positive correlation between values of k_i and γ_i (Fig. 1). In the limiting case of a correlation between k_i and γ_i equal to 1, growth trajectories never cross through organisms' lifetime (i.e., size ranks are always maintained throughout the lifetime of individuals), thus aggressive and/or highly active individuals are always able to obtain a higher share of resources (and invest them in growth). Thus, a similar pattern of growth at the population level emerges when $\psi = 0$ and values of k_i and γ_i are strongly negatively correlated or when $\psi = 1$ and values of k_i and γ_i are strongly positively correlated, but the biological and environmental processes leading to the emergence of similar growth trajectories are different. For values of ψ between 0 and 1, a rich variety of growth trajectories can be obtained depending on the correlation between values of k_i and γ_i as well as their variances (Fig. 1), although the biological interpretation of ψ is challenging with respect to the case of individual variability only for k, in particular with values of ψ far from either 0 or 1.

Parameter estimation and individual variation

Formulations of the standard vBGF with individually varying parameters (L_∞, k, t_0) have been proposed previously (Sainsbury 1980, Francis 1988, Wang and Thomas 1995, Laslett et al. 2002, Pilling et al. 2002, Eveson et al. 2007, Sigourney et al. 2012). Here, we present a novel formulation of the vBGF (as described in *Growth model*) specific for longitudinal data where γ, k, and t_0 are a function of shared predictor(s) and individual random effects.

We treat t_0 and ψ as population-level parameters (with no environmental predictors and no individual random effects), so that all individuals are assumed to share the same value. This improves the biological interpretation of the other parameters and helps with model fitting. Since k and γ must be non-negative, we use a log-link function to facilitate parameter estimation and convergence of the model-fitting procedure. For individual i in group j (e.g., sex, year-of-birth cohort) we thus set

$$\begin{cases} log(k_{ij}) = \log(k_0) + \alpha_j x_i + \sigma_u u_{ij} \\ log(\gamma_{ij}) = \log(\gamma_0) + \beta_j x_i + \sigma_v v_{ij} \\ \qquad\qquad t_0^{(ij)} = t_0 \end{cases} \tag{7}$$

where u_{ij} and v_{ij} are the standardized individual random effects, σ_u and σ_v are the standard deviations of the statistical distributions of the random effects (which we take to have prior distributions that are normal, while the posterior distribution is not guaranteed to be normal due to the nonlinearity of the likelihood function), k_0 and γ_0 are population-level parameters, α_j and β_j are group-level parameters, and x_i has value of 1 if individual i is in group j and 0 otherwise. The model with no predictors and no individual random effects contains only four parameters

(plus the residual variance), $\log(k_0)$, $\log(\gamma_0)$, t_0, ψ, that is, parameters are estimated at the whole population level with no individual variation in growth. In the following, we report and interpret parameter estimates of k_0 and γ_0 on their natural scale, as this allows to directly comparing their estimates to published values.

We use the Automatic Differentiation Model Builder (ADMB) software to estimate the parameters of the growth models (Vincenzi et al. 2014b). ADMB is an open source statistical software package for fitting nonlinear statistical models (Fournier et al. 2012, Bolker et al. 2013) that is becoming a standard tool for use in fisheries stock assessment and management. ADMB–RE (the random-effects module of ADMB) has the ability to fit generic random-effects models using an Empirical Bayes approach that implements the Laplace approximation (Skaug and Fournier 2006). Empirical Bayes (EB) refers to a tradition in statistics where the fixed effects and variance of a random-effects model are estimated by maximum likelihood, while estimates of random effects are based on Bayes formula. Although traditionally random effects are predicted and fixed effects are estimated, we refer in this paper to estimates of k_i and γ_i. Model fitting in ADMB-RE automatically stops when the maximum gradient (i.e., the larger of the partial derivatives of the likelihood function with respect to model parameters) is $<10^{-4}$ (appropriate with log-transformed model parameters).

The length of individual i in group j at age t is

$$L_{ij}(t) = \gamma_{ij} k_{ij}^{\psi-1} (1 - e^{-k_{ij}(t-t_0)}) + \varepsilon_{ij} \tag{8}$$

where ε_{ij} is normally distributed with mean 0 and variance $\sigma_\varepsilon^{2^{ij}}$ (estimated in the model-fitting procedure).

For simplicity, we do not explicitly introduce process stochasticity, so that the likelihood is (Hilborn and Mangel 1997)

$$\prod_{j=1}^{J} \prod_{i=1}^{n_j} \prod_{l=1}^{m_{ij}} \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp\left(-\frac{\left(L_{ijl} - L(t_{ijl}; \gamma_{ij}, k_{ij}, t_0^{(ij)}, \psi)\right)^2}{2\sigma_\varepsilon^2} \right) \tag{9}$$

where n_j is the number of individuals in group j, J is the number of groups, m_{ij} is the number of observations from individual i of group j, l is an index that run over these observations. Further, the observed length measurements for individual i in group j are denoted by L_{ijl}, while t_{ijl} is the age of the individual when the l-th measurement is made. In the following, we will simply use k_i and γ_i for the individual-level parameters.

Note that Eq. 9 is only the likelihood for the observation part of the model. To obtain the likelihood that is used for parameter estimation it is necessary to include the contributions from the random effects, and to integrate the joint likelihood with respect to the random effects (Vincenzi et al. 2014b).

TABLE 1. Model parameters, model assumptions, and data-generating parameter values or their empirical estimates/relationship.

Parameter	Description
L_∞	asymptotic length reached in the limit of infinite time
k	coefficient of catabolism or vBGF growth coefficient
t_0	age at which length is 0
q	coefficient of anabolism
γ	parameter describing the environmental contribution to anabolism
ψ	parameter bounded between 0 and 1 describing the interaction between the environment and the foraging characteristics of the species
k_0 and γ_0	population-level parameters in the linear models for $\log(k)$ and $\log(\gamma)$
α and β	group-level parameters in the linear models for $\log(k)$ and $\log(\gamma)$
u and v	standardized individual random effects in the linear models for $\log(k)$ and $\log(\gamma)$
σ_u and σ_v	standard deviations of the statistical distributions of the random effects in the linear models for $\log(k)$ and $\log(\gamma)$
Model assumptions	
$L_\infty = q/k$	asymptotic size emerges from the relationship between the coefficients of anabolism and catabolism
$q = \gamma k^\psi$	the coefficient of anabolism q depends on environmental (represented by γ) versus behavioral (represented by k) factors, whose respective importance is modulated by the value of the parameter ψ
Estimated/data-generating parameter values and relationship between parameter values	
ψ_s	data-generating ψ
r_s	data-generating Pearson's correlation between individual-level pairs of k and γ
ψ_f	fixed value of ψ when fitting length-at-age data
r_f	empirical correlation between estimated pairs of k and γ

We give a description of model parameters, model assumptions, and imposed parameter values or their empirical estimates/relationship in Table 1.

Case study

We use as model systems for fitting and application of the growth model one population of marble trout living in Zakojska stream and one population of brown trout living in Upper Volaja stream in the western region of Slovenia (Vincenzi et al. 2012, Fig. 2). The population of Zakojska was established in 1996 by stocking age-1 fish that were the progeny of parents from a relict, genetically pure marble trout population (Crivelli et al. 2000). Fish hatched in Zakojska for the first time in 1998 and the 1998 cohort is the first included in the analysis. Upper Volaja was sampled for the first time in 2006 and the oldest cohort to be included in the analysis was born in year 2000. The two populations were sampled annually in June. Fish were collected by electrofishing and measured for length and mass to the nearest millimeter and gram, respectively. If fish were caught for the first time (or if the tag had been lost) and they were longer than 110 mm, they were tagged with Carlin tags (Carlin 1955) and age was determined by reading scales. Males and females in both marble and brown trout are morphologically indistinguishable at the time of sampling. The probability of capture at time t of a fish alive at time t was higher than 80% (Vincenzi et al. 2008). Marble trout females reach sexual maturity when longer than 200 mm, usually at age 3 or older, while age at first reproduction for brown trout in Upper Volaja occurs at age 2 or older. The maximum observed age for fish born in the streams was 9 and 10 yr

in Zakojska and Upper Volaja, respectively. The last sampling occasion included in the data set was June 2013. In Upper Volaja, the last cohort included was the one born in 2011. Due to a flood that almost completely wiped out the population in 2007 (Vincenzi et al. 2012), there were no fish born in Zakojska in 2008–2010. Also in Zakojska, the last cohort included was the one born in 2011. Density of fish age-1 and older (number/m²) was (mean ± standard deviation [SD]) 0.05 ± 0.04 in Zakojska from 1998 to 2013 and 0.05 ± 0.05 in Upper Volaja from 2006 to 2013. In total, 1141 unique fish were included in the Zakojska data set and 1649 in the Upper Volaja data set.

Statistical analysis

Simulated data.—As this is the first time the model in Eq. 8 is proposed, we started by studying the behavior of the model using simulated data. First, we tested whether the same growth trajectories could be described using the parameter and growth functions in Eqs. 7 and 8 with different values of ψ from 0 to 1 with a step of 0.1. To do so, we first simulated 400 (potentially) 10-yr long unique growth trajectories with a true (i.e., data-generating) ψ (ψ_s) in Eq. 8 and for different scenarios with positive, negative, or no correlation (r_s) between the 400 pairs of k_i and γ_i. Specifically, we imposed a correlation structure between normal distributions of individual random effects for k_i and γ_i (u_i and v_i, respectively), we randomly drew 400 (u_i, v_i) pairs from the joint probability distribution of random effects, and then obtained 400 ($\log(k_i)$, $\log(\gamma_i)$) pairs following Eq. 7. To simulate a realistic empirical case, we used a mortality rate M and excluded one observation, on average, per individual.
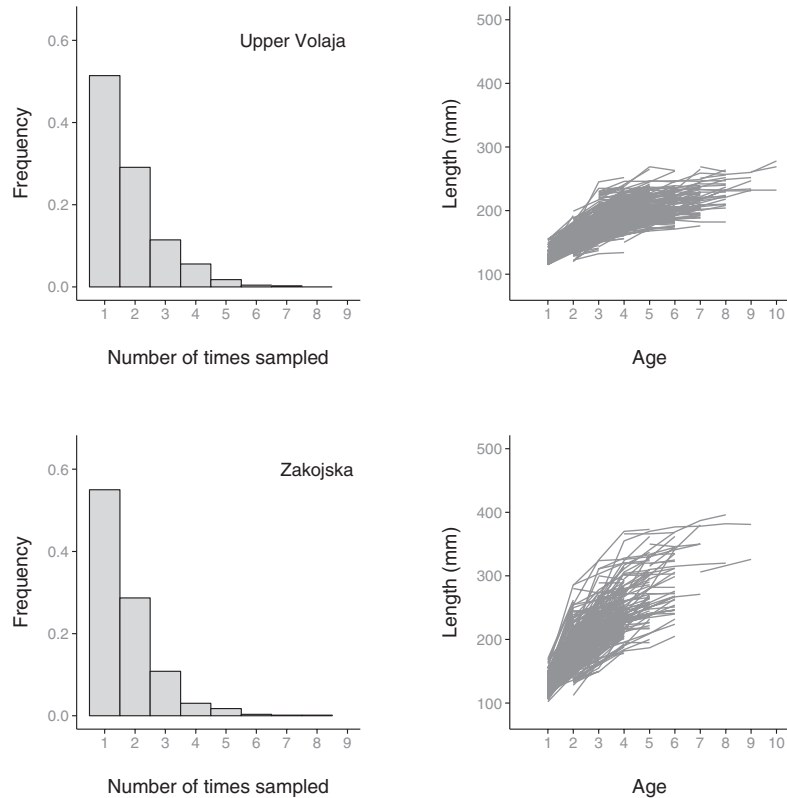
Fig. 2.   Frequency of sampling events per individual and empirical growth trajectories for the populations of Upper Volaja (brown trout, *Salmo trutta*) and Zakojska (marble trout, *Salmo marmoratus*) streams, western Slovenia.

We did not introduce group parameters to simulate individual growth trajectories (i.e., $\alpha_j = \beta_j = 0$ in Eq. 7). Then, we fitted the model to the simulated growth trajectories by fixing the value of ψ (ψ_f) in Eq. 8. Due to the random sampling of u_i and v_i from the joint probability distribution, we ran 10 random replicates for each combination of ψ_s, ψ_f, and r_s, and recorded the convergence of the model-fitting procedure as determined by the maximum gradient in ADMB-RE, average over replicates of mean absolute error (MAE; i.e., mean absolute difference between simulated length-at-age and length-at-age predicted by the fitted model) and mean correlation between the 400 estimated pairs of k_i and γ_i (r_f) across replicates ($\bar{r}_f$) that successfully converged. We did not compare data-generating parameter values and parameter estimates as they can be meaningfully compared only when $\psi_s = \psi_f$. We also recorded how many times over the 10 random replicates for each combination of ψ_s, ψ_f, and r_s the model fitting procedure failed to converge. Convergence failure means that ADMB-RE was not able to obtain a sufficiently small likelihood gradient, with the (default) criterion being 10^{-4} for all parameters. The particular reason for lack of convergence may vary across simulation replicates, and although it is not feasible to investigate each case in detail, it is good practice to keep track of the number of cases in which convergence was not achieved.

Selection of the best growth model for Zakojska and Upper Volaja and prediction of unobserved data.—We checked the maximum gradient component to ensure that a satisfactory convergence was reached. Except for the case of nonlinear regression explained in the next paragraph, each model we tested included individual random effects as in Eq. 7. Following Vincenzi et al. (2014b), we introduced year-of-birth cohort as fixed categorical effects to test whether its inclusion as predictor included model fitting to data for either population (α_j and β_j in Eq. 7). The model may not be always not statistically identifiable, in that ψ can only sometimes be estimated (Shelton and Mangel 2012). Thus, we fitted separately models with or without cohort as predictors of γ_i, k_i, or both with ψ from 0 to 1 with a step of 0.1. We used the Akaike Information Criterion (AIC; Akaike 1974, Burnham and Anderson 2002) to select the best model. We then investigated correlation between the estimates of k_i and γ_i. We tested whether the inclusion of individual random effects for both k and γ (thus increasing model complexity) increased model accuracy with respect to models that include individual random effects only for k (i.e., models in Shelton and Mangel 2012, Shelton et al. 2013). For the latter model, in Eq. 8, we thus fixed $\sigma_v = 0$.

We tested whether vBGF models with random effects for both k and γ with different values of ψ predicted

substantially different mean cohort-specific growth trajectories. In addition, we tested whether fitting nonlinear least-squares regression with no random effects on cohort-specific data (using the nls function in R; R Development Core Team 2011) leads to substantially different mean cohort-specific growth trajectories with respect to random-effect models.

We tested the predictive ability of (1) the best overall vBGF model with individual variation for both k and γ (where ψ_b is the value of ψ for the best model), as well as the models with (2) $\psi = 0$ and (3) $\psi = 1$ for both populations. We also tested the predictive ability of the best overall vBGF model with variation only for k. For each population, we: (1) randomly sampled one-third of fish that have been sampled more than three times throughout their lifetime (validation sample); (2) deleted from the data set all observations except the first one from each individual fish in the validation sample; (3) estimated the parameters of the vBGF for each individual including those in the validation sample; and (4) predicted the missing observations.

We compared the predictions of the vBGF to the predictions given by the mean length-at-age of the cohort of the fish. We used MEA and R^2 with respect to the 1:1 line observed data vs. predicted data as measures of predictive ability. The predictive abilities of the vBGF models were tested using the same 10 random validation samples for each population.

Results

Simulated data

Model fitting with simulated data showed that when growth trajectories had a negative correlation r_s between k_i and γ_i, the average correlation between k_i and γ_i across replicates $\bar{r}_f$ tended to remain negative in the area below the 1:1 line in the $\psi_s - \psi_f$ plane and around zero or positive above the line (Fig. 3a). When growth trajectories were simulated starting from a positive r_s, $\bar{r}_f$ tended to remain close to 0 or slightly negative in the fitted models below the 1:1 line and mostly positive above the line (Fig. 3c). When r_s was equal to 0, the empirical correlation between estimated k_i and γ_i in the fitted models tended to be around 0 for the majority of combinations of ψ_s and ψ_f (Fig. 3b). Similar results were obtained when using different variances of the individual random effects (Appendix S1: Figs S3 and S4).

The probability of convergence of the model fitting procedure varied across combinations of ψ_s, ψ_f, and r_s. Although a clear pattern of probability of convergence did not emerge, the model-fitting algorithm converged for most of the $\psi_s - \psi_f$ combinations and replicates (Fig. 3d–f). The average of MAE across replicates was smaller than 2 mm (thus an almost perfect fit) in more than 90% of the combinations of ψ_s and ψ_f.

Case study

Observed trajectories showed higher individual variation in growth and length at age in the marble trout population of Zakojska than in the brown trout population of Upper Volaja (Fig. 2).

For the vBGF models without cohort as a predictor for either k_i and γ_i, the correlation between k_i and γ_i was function of ψ (Fig. 4), and tended to shift from a negative to a positive correlation with increasing values of ψ for both populations. In this case, for the population of Upper Volaja the best model according to AIC had $\psi = 0.6$ (AIC = 23855.4), while the model with $\psi = 0$ had lower AIC than the model with $\psi = 1$ (23951.4 vs. 24059.2). For the population of Zakojska, the best model according to AIC had $\psi = 0.3$ (AIC = 17387.8), while the model with $\psi = 0$ had lower AIC than the model with $\psi = 1$ (17395.9 vs. 17445.1). The joint distribution of ψ and sign and strength of the correlation between k_i and γ_i (Fig. 4) suggested maintenance of size ranks throughout fish lifetime for both Zakojska and Upper Volaja trout populations, with growth trajectories crossing on average after sexual maturity (Fig. 1). Every model predicted the observed data to high accuracy (Zakojska, range of MAE = 7.1–8.5 mm, range of R^2 = 0.98–0.98; Upper Volaja, MAE = 3.7–5.0 mm, R^2 = 0.97–0.98). Assuming a lifespan of 10 yr (i.e., predicting 10 yr of length-at-age for each fish), growth trajectories predicted using estimated parameters for models with different value of ψ (and without cohort as predictor of either parameter) had similar mean age at crossing of growth trajectories and CV of length at age 10, but substantially different number of trajectories crossing throughout the lifetime of fish (Fig. 5).

For marble trout, the best model using AIC as model-selection criterion had cohort as predictor both in k_i and γ_i and $\psi = 0.3$ (Table 2). For brown trout, the best model had cohort as predictor for only γ_i and $\psi = 0.6$. For both Zakojska and Upper Volaja populations, the models with individual random effects only in k_i performed far worse than the models with individual random effects for both parameters (Table 2).

Cohort-specific models for marble trout and brown trout with cohort as predictor both in k_i and γ_i provided essentially the same mean trajectories when ψ was equal to 1 and when ψ was the one giving the smallest AIC (i.e., $\psi = 0.3$ for Zakojska and $\psi = 0.5$ for Upper Volaja; Fig. 6). Cohort-specific vBGF models with no random effects fitted with standard nonlinear least-squares regression predicted substantially greater length-at-age than random-effects models for the marble trout population of Zakojska, while they provided the same mean cohort-specific growth trajectories as the random-effects models for the brown trout population of Upper Volaja (Fig. 6).

Prediction of unobserved length-at-age.—In the population of Upper Volaja and Zakojska, 132 and 63 fish were sampled more than three times during their lifetime, respectively. The vBGF model with both k and γ function
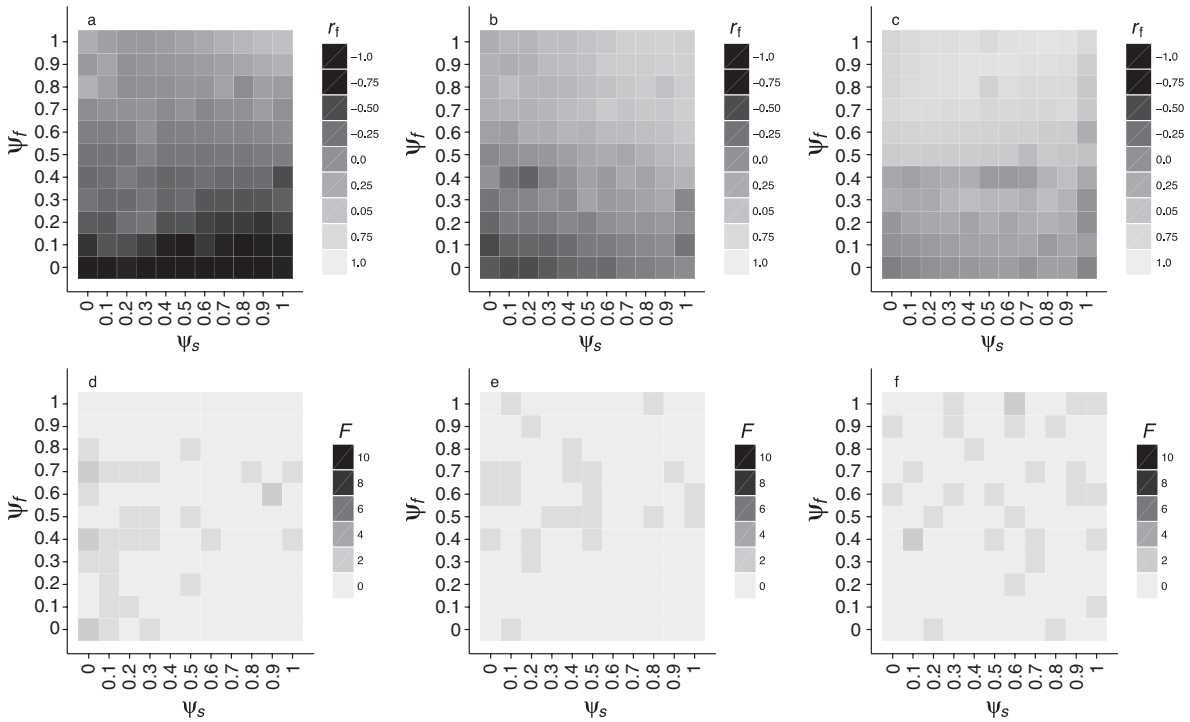
FIG. 3. (a–c) Average Pearson's correlation $\bar{r}_f$ between k_i and γ_i across replicates that successfully converged when fitting von Bertalanffy growth models as in Eqs. 7 and 8 with $\psi = \psi_f$ on growth trajectories simulated with $\psi = \psi_s$, $L_\infty = 300$ mm, $t_0 = -0.32$ y, $k_0 = 0.38$ y^{-1}, $\sigma_u = \sigma_v = 0.37$, $\gamma_0 = L_\infty(k_0^{1-\varphi})$, mortality rate $M = 0.8$, and correlation r_s between k_i and γ_i equal to (a, d) −0.9, (b, e) 0, (c, f) 0.9. (d–f) Number of replicates F that did not converge for every combination of ψ_s and ψ_f out of the 10 replicates. Plots for $\sigma_u = \sigma_v = 0.60$ and $\sigma_u = \sigma_v = 0.14$ are provided in Appendix S1 (Figs. S3 and S4).

of cohort, individual random effects, and $\psi = 1$ provided consistently better prediction of the missing observations than models with $\psi = 0$, ψ giving the best AIC value for models with both k and γ function of cohort and than prediction based on mean length-at-age of the respective cohort (Table 3; Figs 7 and 8). The best model with individual variation only in k provided substantially worse predictions than the best model with individual variation for both k and γ.

DISCUSSION

Our formulation of the von Bertalanffy growth function balances biological details of the growth process and model fitting, and thus provides a flexible and powerful framework for estimating and understanding the role of abiotic and biotic factors in determining organisms' growth. This unification is achieved by an ecological, rather than purely statistical, focus that considers growth in terms of the behavior–environment interaction. Adding complexity in the form of individual variability in both mechanistic parameters (k and γ) of our formulation of the von Bertalanffy growth function increases model accuracy with respect to the model including individual variability only in k, which is the parameter summarizing physiological and behavioral traits that determine individual activity. We now discuss the results of our simulations, parameter

estimation, and model selection using two fish populations as a case study, and their implications for our understanding of the determinants of variation and for management and conservation.
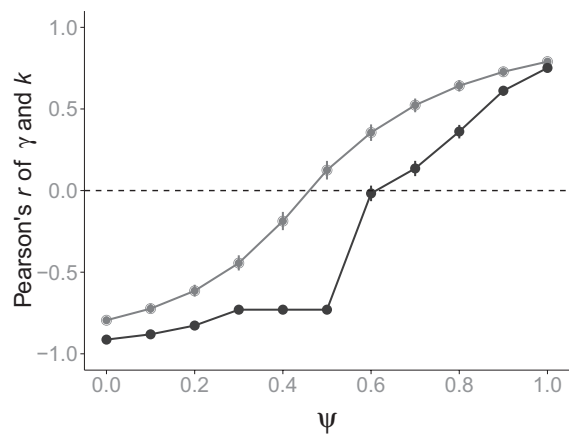

FIG. 4. Correlation (Pearson's r) between estimates of k_i and γ_i for different values of ψ for the von Bertalanffy growth model with no predictors other than individual random effects for either k and γ for Zakojska (gray) and Upper Volaja (black). Vertical segments are 95% confidence intervals of r. The best model with no predictors other than individual random effects according to AIC for Upper Volaja and Zakojska had $\psi = 0.6$ and $\psi = 0.4$, respectively.
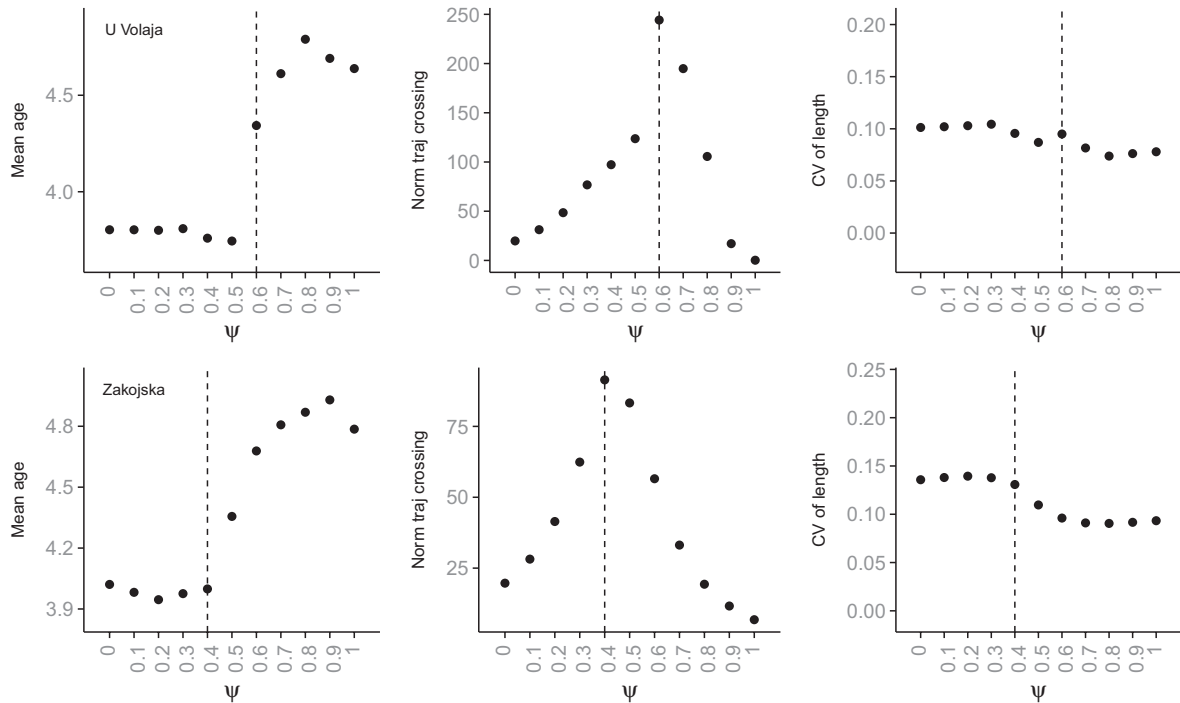
Fig. 5. Mean age at crossing of growth trajectories, total number of trajectories crossing divided by the total number of individuals in the population (1649 for Upper Volaja and 1147 for Zakojska), and coefficient of variation of length at age 10 for von Bertalanffy growth models (with no predictors except individual random effects for either model parameter) with model parameters estimated for values of ψ from 0 to 1 with step 0.1. All growth trajectories of unique individuals were predicted for a theoretical lifespan of 10 yr according to the estimated model parameters at the individual level. Vertical dashed lines identify the best model according to Akaike information criteria (AIC) for models with no predictors for either parameter.

Relationship between ψ and the correlation between model parameters

Our simulation showed that models with different values of parameter describing the interaction between the environment and the foraging characteristics of the species (ψ) and variability in both the parameter of catabolism (k) and the parameter describing the environmental contribution to anabolism (γ) are in general able to describe very similar growth trajectories. A clear pattern of probability of convergence of the model fitting procedure did not emerge from our simulations, but model fitting was successful in the vast majority of cases. This flexibility has to be ascribed to the many degrees of

TABLE 2. The 10 best von Bertalanffy growth models according to AIC for the marble trout (*Salmo marmoratus*) population of Zakojska and the brown trout (*Salmo trutta*) population of Upper Volaja, western Slovenia.

	Zakojska				Upper Volaja		
Model	ψ	AIC	npar	Model	ψ	AIC	npar
$k, γ$	0.3	17105	29	γ	0.6	23269	19
$k, γ$	0.4	17106	29	$k, γ$	0.5	23277	31
$k, γ$	0.2	17112	29	$k, γ$	0.6	23283	31
$k, γ$	0.1	17123	29	γ	0.5	23303	19
$k, γ$	0.5	17127	29	$k, γ$	0.3	23328	31
$k, γ$	0	17132	29	$k, γ$	0.2	23338	31
γ	0.4	17137	20	$k, γ$	0.4	23346	31
γ	0.3	17144	20	γ	0.4	23348	19
$k, γ$	0.6	17153	29	$k, γ$	0	23350	31
γ	0.2	17159	20	$k, γ$	0.7	23363	31
$k, γ$	0.4	17831	27	$k, γ$	0.6	24977	27

Notes: Parameters included in the model column are those that are a function of cohort; npar is the number of model parameters; AIC = Akaike Information Criterion. The last row reports the best models when individual random effects are included only for k.
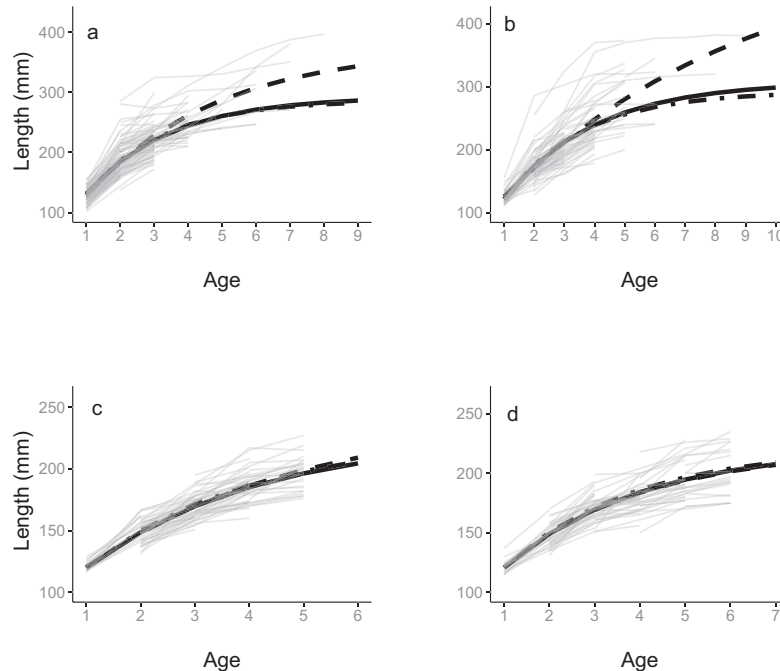
FIG. 6.   Cohort-specific growth trajectories for the trout populations of Zakojska (a, cohort 1999; b, cohort 2001) and Upper Volaja (c, cohort 2008, d, cohort 2007). Dashed line: prediction of model with no random effects fitted on cohort data with non-linear least square regression. Dash-dotted line: cohort-specific model with cohort as predictor for both k_i and γ_i with $\psi = 0.3$ and $\psi = 0.5$ (best models) for Zakojska and Upper Volaja, respectively. Solid line: cohort-specific model with cohort as predictor for both k_i and γ_i with $\psi = 1$.

freedom of our formulation of the von Bertalanffy growth function with individual random effects. Furthermore, when simulating growth trajectories with negative or positive correlation between pairs of k_i and γ_i, the sign of the correlation tended to remain negative and positive, respectively, when fitting models with other values of ψ. This pattern emerged only in the case of simulated data, since a clear change of sign of the correlation between pairs of k_i and γ_i was found when fitting the growth models to empirical data. This has to be ascribed to some unrealistic growth trajectories that are obtained when keeping the same variance for individual random effects for each value of ψ used to generate the growth trajectories (Appendix S1: Fig. S5).

Case study

Model selection, parameter estimates, and trade-off between accuracy and interpretability in growth models.—
All models predicted the observed data very well, although there were small differences in performance among models for either population. However, when predicting growth trajectories using the estimated model parameters for models with different values of ψ, the best model (for either population) among those with no predictor for either model parameter was the one predicting the highest number of crossing growth trajectories. The mean absolute error was very low in

each model, thus realized growth trajectories were almost perfectly predicted by each model. It follows that the differences in predicted growth trajectories should be mostly ascribed to differences in prediction of growth trajectories for fish that have been sampled one or a few times early in life, that is, the best model predicted that size ranks for the growth trajectories that were not realized due to early mortality were less maintained (i.e., more trajectories crossing) with respect to the other models.

Accuracy describes the ability of a model to explain observed data and make correct predictions, while interpretability concerns to what degree the model allows for understanding processes. Often a trade-off exists between accuracy and interpretability; more complex models are usually opaque, while more interpretable models often do not provide the same accuracy or predictive power of more complex models (Breiman 2001). McCullagh and Nelder (1989) wrote: "Data will often point with almost equal emphasis on several possible models, and it is important that the statistician recognize and accept this". However, different models may give different insights on the relation between the predictors and response variables (length-at-age), and how to determine which model most accurately reflects the data remains a challenge. One way is to use model selection procedures that trade off goodness-of-fit (the likelihood) and model complexity (number of parameters) to select for the best model (Burnham and

TABLE 3. Mean ± standard deviation (SD) of R^2 and mean absolute error (MAE, mm) of predictions of validation data for 10 random validation samples as provided by (1) the vBGF model with variation for both for both k and γ (vBGF [k_i, γ_i]) including cohort as predictor for both k and γ with $\psi = 0$, 1, and ψ of the best model according to AIC ($\psi = 0.5$ for Upper Volaja and $\psi = 0.3$ for Zakojska), (2) the vBGF model with variation for only k (vBGF [k_i]) including cohort as predictor for both k and γ with ψ of the best model according to AIC ($\psi = 0.6$ for Upper Volaja and $\psi = 0.4$ for Zakojska).

| Population | vBGF (k_i, γ_i) | | | | | | vBGF (k_i) | | | |
| | $\psi = 1$ | | $\psi = 0$ | | Best ψ | | Best ψ | | Cohort | |
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
Zakojska	0.61 ± 0.22	32.0 ± 13.0	0.54 ± 0.22	35.1 ± 12.3	0.52 ± 0.16	36.3 ± 9.5	0.44 ± 0.07	38.2 ± 5.46	0.52 ± 0.09	36 ± 5.1
Upper Volaja	0.57 ± 0.21	15 ± 3.4	0.56 ± 0.2	15.2 ± 3.3	0.55 ± 0.21	15.2 ± 4.2	0.47 ± 0.09	17.2 ± 1.23	0.51 ± 0.2	16 ± 2.9

Note: We also report mean ± SD of R^2 and MAE of predictions with mean length-at-age of the respective cohorts (cohort).

Anderson 2002, Johnson and Omland 2004). In our work, the AIC analysis showed for either population that models with individual random effects for both parameters performed substantially better than models with individual random effects only for k. Thus, increasing the complexity of the model by allowing individual variation in both parameters increased the accuracy of the growth models.

The best model for the marble trout population of Zakojska included cohort as a categorical predictor for both γ and k, while for the brown trout population of Upper Idrijca, the best model included cohort as predictor of k. That means that parameter values as well as the resulting predicted growth trajectories of fish seem to be more similar to those of fish in the same cohort than to those of the population as a whole.

The prediction of mean cohort-specific growth trajectories using models with or without random effects showed different results for the populations of Upper Volaja and Zakojska. For the brown trout population of Upper Volaja, cohort-specific models with no random effects (i.e., estimated using the nls function in R) and random-effects models with cohort as predictor of both k_i and γ_i with either $\psi = 1$ or ψ of the best model provided essentially the same prediction of mean cohort-specific growth trajectories. On the other hand, for the marble trout population of Zakojska the random-effects models provided essentially the same predictions of mean cohort-specific growth trajectories, while the cohort-specific models with no random effects tended to predict substantially higher length-at-age for fish older than 4 years old. This occurred because there was higher variation in length-at-age in Zakojska than in Upper Volaja and some big fish tended to have a longer lifespan in Zakojska, thus growth trajectories tended to be "pulled up" by the big, older individuals. This result supports the use growth models with individual random effects, in particular when there is substantial variability in both growth rates and size-at-age of individuals living in the same population. However, in both populations the random-effects models provided essentially the same predictions of mean cohort-specific growth trajectories.

Biological interpretation of the selected growth models and parameter estimates.—Across taxa, climatic vagaries during the first stages of life have the potential to influence the mean growth trajectories of cohorts, as well as other life histories. Strong empirical evidence of early induced effects on later growth rate, life-history traits, and behavior of organisms is quite recent (Danchin and Wagner 2010, Salvanes et al. 2013, Ait Youcef et al. 2015). Jonsson and Jonsson (2014) recently discussed how conditions fish encounter early in their life cycle could leave lasting effects on morphology, growth rate, life history, and behavioral traits. Vincenzi et al. (2014a,b) found that other processes may be potentially responsible for variability of mean growth trajectories of cohorts, such as high variance in reproductive success combined with either high heritability of growth or heterogeneity in site profitability accompanied by limited movement. High heritability of growth (Carlson
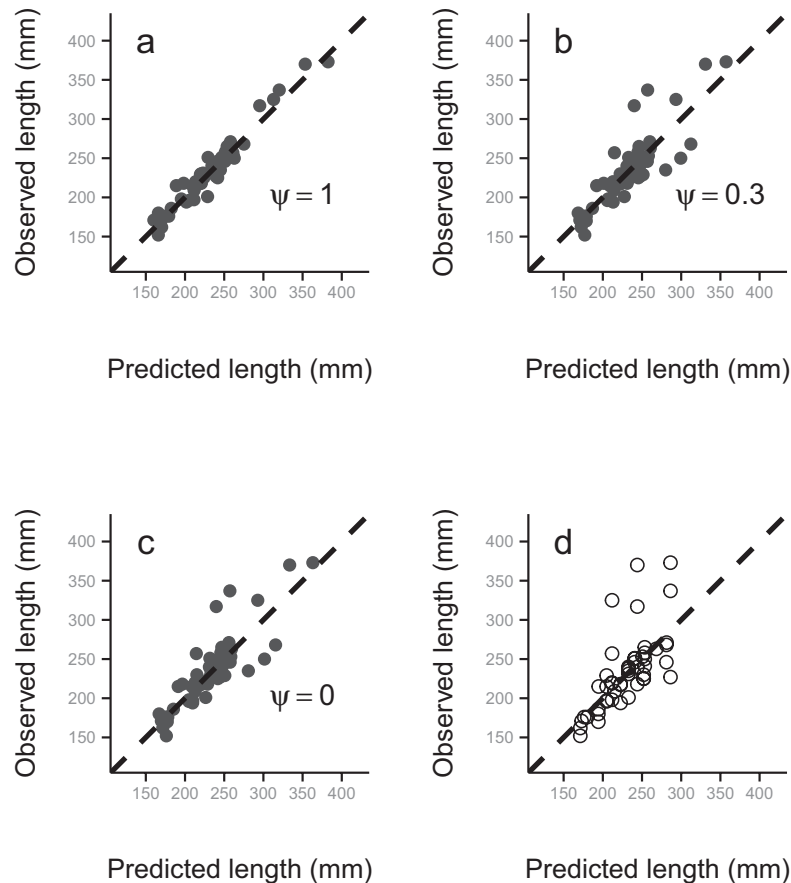
Fig. 7. Example of prediction of validation data for the marble trout population of Zakojska with the von Bertalanffy growth model with cohort as predictor of k_i and γ_i and (a) $\psi = 1$, (b, best model) $\psi = 0.3$, and (c) $\psi = 0$. (d) Prediction of validation data using mean length-at-age for the cohort of the individual whose growth is predicted.

and Seamons 2008), maternal decisions on the timing and location of spawning (Letcher et al. 2011), and dominance established early in life (Gilmour et al. 2005) are all processes that may in combination or by themselves explain the maintenance of size ranks throughout fish lifetime.

Trade-offs between growth and survival have been found across species and taxa (Pauly 1980) as well as at the individual level within populations at the early life stages (Biro and Post 2008, Woodson et al. 2013). Given the similarity in growth of fish in the same cohort, we may hypothesize cohort effects also in survival, either in the direction of higher mortality for faster-growing cohorts due to trade-offs between growth and survival, or of higher survival for faster-growing cohorts when faster growth is a signal of higher quality of individuals.

The biological interpretation of model parameters is easier when only the individual random effects (and not cohort) are included as predictors. In this case, for both populations the model with $\psi = 0$ performed substantially better than the model with $\psi = 1$. This result, along with the strong negative empirical correlation between estimates of k_i and γ_i when $\psi = 0$, suggests that size ranks are largely maintained throughout marble and brown trout lifetime,

crossing of growth trajectories mostly occurs after sexual maturity, and that more aggressive/active fish are on average growing slower than those less aggressive/active.

One hypothesis is that both trout populations live in an environment in which resource acquisition depends less on intrinsic behavioral traits and more on habitat, and thus more active individuals are expending more energy than less active individuals without acquiring more resources. Support for this hypothesis comes from the mean bigger size-at-age found in both Upper Volaja and Zakojska for trout living in the uppermost part of the streams, as more food (in particular, invertebrate drift) is available there.

As for growth trajectories crossing mostly after sexual maturity, one potential explanation is sex-specific energetic investment in reproduction, with females allocating more energy to reproduction than growth with respect to males.

Predicting unobserved data.—The variation in growth and size that characterizes organisms can almost always be modeled retrospectively. However, the limited number of attempts at predicting missing size observations or unobserved growth trajectories may also depend on the
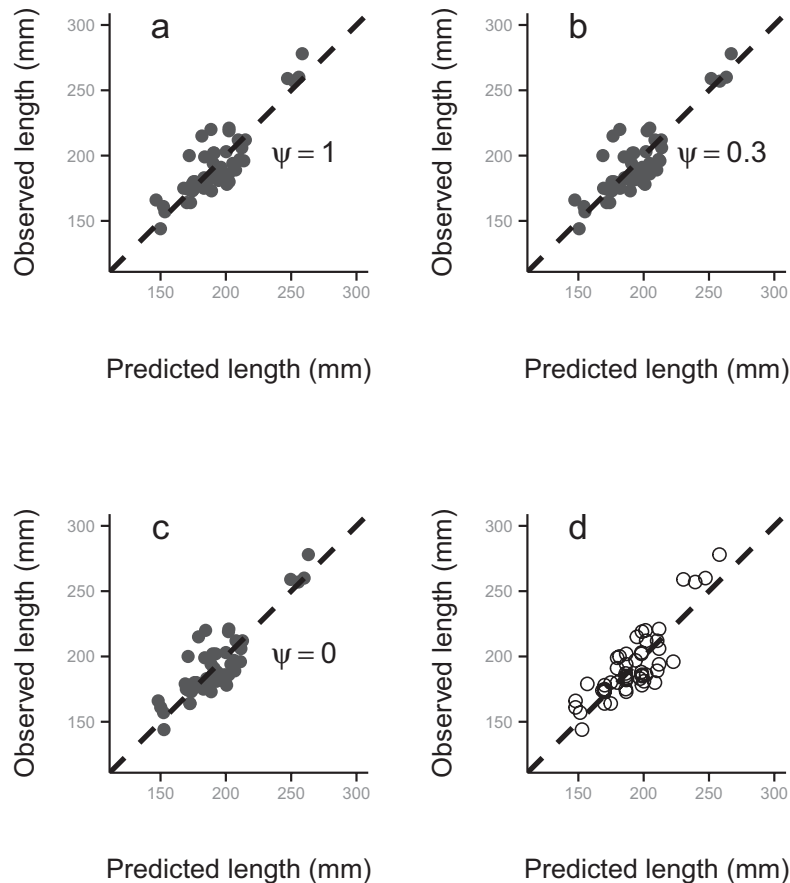
FIG. 8. Example of prediction of validation data for the brown trout population of Upper Volaja with the von Bertalanffy growth model with cohort as predictor of k_i and γ_i and (a) $\psi = 1$, (b, best model) $\psi = 0.3$, (c) $\psi = 0$. (d) Prediction of validation data using mean length-at-age for the cohort of the individual whose growth is predicted.

intrinsic unpredictability of some growth curves, for which it may be impossible to accurately predict later portions of the growth trajectory when only a few observations early in life are available (e.g., ocean growth of anadromous salmonids when only a few observations relative to the freshwater phase are available; Norton et al. 1976). The vBGF models with cohort as predictor of both k_i and γ_i, and $\psi = 0$, 1, or ψ of the best overall model (ψ_b) provided good predictions of unobserved growth trajectories for both the marble and brown trout populations, and except for one case ($\psi = \psi_b$ for Zakojska) the predictions were consistently better than predictions of the best model with individual variability only for k and of predictions based on the mean length-at-age of the fish cohort. However, for neither the marble trout population of Zakojska nor for the brown population of Upper Idrijca did the best model selected according to AIC provide the best prediction of unobserved growth trajectories. Although the best model did not formally overfit, the additional flexibility provided by a value of ψ not equal to 0 or 1 did not translate in more accurate predictions of unobserved growth trajectories.

Conclusions and implications for management

The purpose of a scientific investigation should drive model formulation and the type and amount of data collected. Random-effects models and powerful software and routines allow the fitting of complex models, but often complexity comes at the cost of interpretability of model parameters. Our work shows that adding additional complexity to the von Bertalanffy growth function (e.g., cohort as predictor of vBGF's parameters, variability in both k and γ) may offer substantial advantages in terms of understanding of the determinants of growth patterns and predicting or estimating the future or unobserved size-at-age of individuals. When using the model formulation that we propose in this study and for ease of interpretation of model parameters, we recommend limiting model selection to models with $\psi = 0$ (vBGF as formulated by von Bertalanffy, in which asymptotic size is an explicit function of the growth coefficient) or $\psi = 1$ (vBGF as commonly fitted, in which asymptotic size is independent of the growth coefficient). In those two cases, model selection may give clearer insights on processes leading to individual and group variation in growth while providing accurate predictions of

unobserved or future size-at-age data and growth trajectories. Further insights on the processes leading to variation in growth would come from combining parameter estimation and model selection with estimates of metabolic rates, patchiness of resources, movement of fish, and costs of reproduction. Further investigation on these trade-offs are needed using other growth models and other species.

By furthering our understanding of variation in life-history processes that depend on, or correlate with, growth processes, our modeling approach has relevant implications for more applied contexts. For instance, our results support the hypothesis that both trout populations live in an environment in which resource acquisition depends more on habitat selection than on intrinsic behavioral traits (although, especially early in life, intrinsic difference in behavioral traits contributes to habitat selection). The hypothesis is also supported by the consistently bigger size-at-age of fish occupying the uppermost part of the western Slovenian streams in which other marble trout populations live (where a larger portion of stream drift is available since no fish are present upstream) than of those fish living further downstream (Vincenzi et al. 2010, 2014b, 2015). Trout are typically stationary feeders that hold relatively fixed positions from which they make short forays to feed; according to our model-selection results, habitat choice or chance (such as being born more upstream, especially when natural barriers reduce or impair upstream movement) are critical for growth and fitness of the individual. Riverscapes are highly spatially heterogeneous and the effects of habitat type and quality on individual fitness may be strongest and best explained at the micro-habitat spatial scale (Fausch 1984). The importance of habitat selection may thus suggest the use of spatially explicit models for studying the population dynamics of the two species, as well as for predicting the evolution of growth and other life-history traits (Ayllón et al. 2016).

Estimates of growth are also fundamental to any assessment of population demographics and population dynamics for management. For instance, age-structured stock assessment methods are based on size-at-age that is often derived from parameters of the von Bertalanffy growth model for that species (Katsanevakis and Maravelias 2008). We have shown that for the salmonid populations that we used as a model system, our model allows one to use a single measurement early in the life of individual fish (or, equivalently, a set of measurements from a cohort) to obtain accurate predictions of lifetime individual or cohort size-at-age.

Literature Cited

Adkison, M. D. 2009. Drawbacks of complex models in frequentist and Bayesian approaches to natural-resource management. Ecological Applications 19:198–205.

Ait Youcef, W., Y. Lambert, and C. Audet. 2015. Variations in length and growth of Greenland Halibut juveniles in relation to environmental conditions. Fisheries Research 167:38–47.

Akaike, H. A. I.. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control AC-19:716–723.

Ayllón, D., S. F. Railsback, S. Vincenzi, J. Groeneveld, A. Almodóvar, and V. Grimm. 2016. InSTREAM-Gen: Modelling eco-evolutionary dynamics of trout populations under anthropogenic environmental change. Ecological Modelling 326:36–53.

Berkey, C. S. 1982. Bayesian approach for a nonlinear growth model. Biometrics 38:953–61.

von Bertalanffy, L. 1957. Quantitative laws in metabolism and growth. The Quarterly Review of Biology 32:217–231.

Biro, P. A., and J. R. Post. 2008. Rapid depletion of genotypes with fast growth and bold personality traits from harvested fish populations. Proceedings of the National Academy of Sciences of the United States of America 105:2919–2922.

Bolker, B. M., et al. 2013. Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. Methods in Ecology and Evolution 4:501–512.

Breiman, L. 2001. Statistical modeling: the two cultures. Statistical Science 16:199–231.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer Verlag, New York, New York, USA.

Carlin, B. 1955. Tagging of salmon smolts in the River Lagan. Report of the Institute of Freshwater Research, Drottningholm 36:57–74.

Carlson, S. M., and T. R. Seamons. 2008. A review of quantitative genetic components of fitness in salmonids: implications for adaptation to future change. Evolutionary Applications 1:222–238.

Coulson, T., S. Tuljapurkar, and D. Z. Childs. 2010. Using evolutionary demography to link life history theory, quantitative genetics and population ecology. Journal of Animal Ecology 79:1226–40.

Crivelli, A. J., G. Poizat, P. Berrebi, D. Jesensek, and J.-F. Rubin. 2000. Conservation biology applied to fish: the example of a project for rehabilitating the marble trout (*Salmo marmoratus*) in Slovenia. Cybium 24:211–230.

Danchin, É., and R. H. Wagner. 2010. Inclusive heritability: combining genetic and non-genetic information to study animal behavior and culture. Oikos 119:210–218.

Essington, T. E., J. F. Kitchell, and C. J. Walters. 2001. The von Bertalanffy growth function, bioenergetics, and the consumption rates of fish. Canadian Journal of Fisheries and Aquatic Sciences 58:2129–2138.

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

Fausch, K. D. 1984. Profitable stream positions for salmonids: relating specific growth rate to net energy gain. Canadian Journal of Zoology 62:441–451.

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software 27:233–249.

Francis, R. I. C. C. 1988. Maximum likelihood estimation of growth and growth variability from tagging data. New Zealand Journal of Marine and Freshwater Research 22:42–51.

Frisk, M. G., T. J. Miller, and M. J. Fogarty. 2001. Estimation and analysis of biological parameters in elasmobranch fishes: a comparative life history study. Canadian Journal of Fisheries and Aquatic Sciences 58:969–981.

Gilmour, K. M., J. D. Dibattista, and J. B. Thomas. 2005. Physiological causes and consequences of social status in salmonid fish. Integrative and comparative biology 45:263–73.

Hilborn, R., and M. Mangel. 1997. The ecological detective. Princeton University Press, Princeton, New Jersey.

Johnson, J. B., and K. S. Omland. 2004. Model selection in ecology and evolution. Trends in Ecology & Evolution 19:101–8.

Jonsson, B., and N. Jonsson. 2011. Ecology of Atlantic salmon and brown trout: habitat as a template for life histories. Springer.

Jonsson, B., and N. Jonsson. 2014. Early environment influences later performance in fishes. Journal of Fish Biology 85:151–188.

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

Kimura, D. K. 2008. Extending the von Bertalanffy growth model using explanatory variables. Canadian Journal of Fisheries and Aquatic Sciences 65:1879–1891.

Kingsley, M. C. S. 1979. Fitting the von Bertalanffy growth equation to polar bear age–weight data. Canadian Journal of Zoology 57:1020–1025.

Kooijman, B. 2009. Dynamic energy budget theory for metabolic organisation. Cambridge University Press, New York.

Laslett, G. M., J. P. Eveson, and T. Polacheck. 2002. A flexible maximum likelihood approach for fitting growth curves to tag – recapture data. Canadian Journal of Fisheries and Aquatic Sciences 59:976–986.

Lester, N. P., B. J. Shuter, and P. A. Abrams. 2004. Interpreting the von Bertalanffy model of somatic growth in fishes: the cost of reproduction. Proceedings of the Royal Society B: Biological Sciences 271:1625–1631.

Letcher, B. H., J. A. Coombs, and K. H. Nislow. 2011. Maintenance of phenotypic variation: repeatability, heritability and size-dependent processes in a wild brook trout population. Evolutionary Applications 4:602–615.

Lomnicki, A. 1988. Population ecology of individuals. Princeton University Press, Princeton, New Jersey.

Ludwig, D., and C. J. Walters. 1985. Are age-structured models appropriate for catch-effort data? Canadian Journal of Fisheries and Aquatic Sciences 42:1066–1072.

Mangel, M. 2006. The theoretical biologist's toolbox. Cambridge University Press, New York.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman & Hall/CRC, London, UK.

Norton, L., R. Simon, H. D. Brereton, and A. E. Bogden. 1976. Predicting the course of Gompertzian growth. Nature 264:542–545.

Omori, K., A. Sogabe, H. Hamaoka, and H. Ohnishi. 2009. Storage and the regulation of body mass in animals: a general growth equation based on an energy balance model. Ecological Modelling 220:2618–2623.

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. ICES Journal of Marine Science 39:175–192.

Pelletier, F., T. Clutton-Brock, J. Pemberton, S. Tuljapurkar, and T. Coulson. 2007. The evolutionary demography of ecological change: linking trait variation and population growth. Science 315:1571–4.

Peters, R. H. 1983. The ecological implications of body size. Cambridge studies in ecology. Cambridge University Press, New York.

Peters, R. H. 1991. A critique for ecology. Cambridge University Press, New York.

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

Quince, C., B. J. Shuter, P. A. Abrams, and N. P. Lester. 2008. Biphasic growth in fish II: empirical assessment. Journal of Theoretical Biology 254:207–214.

R Development Core Team. 2011. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Radhakrishna Rao, C. 1987. Prediction of future observations in growth curve models. Statistical Science 2:434–471.

Roff, D. 2007. Life-history evolution. Sinauer Associates Inc., Sunderland, Massachusetts, USA.

Roland, C. H., N. Cameron, and L. Molinari. 2011. Methods in human growth research. Cambridge University Press, Cambridge, UK.

Rosenfeld, J., T. Van Leeuwen, J. Richards, and D. Allen. 2014. Relationship between growth and standard metabolic rate: measurement artefacts and implications for habitat use and life-history adaptation in salmonids. Journal of Animal Ecology 84:4–20.

Russo, T., P. Baldi, A. Parisi, G. Magnifico, S. Mariani, and S. Cataudella. 2009. Lévy processes and stochastic von Bertalanffy models of growth, with application to fish population analysis. Journal of Theoretical Biology 258:521–9.

Sainsbury, K. J. 1980. Effect of individual variability on the von Bertalanffy growth equation. Canadian Journal of Fisheries and Aquatic Sciences 37:241–247.

Salvanes, A. G. V., O. Moberg, L. O. E. Ebbesson, T. O. Nilsen, K. H. Jensen, and V. A. Braithwaite. 2013. Environmental enrichment promotes neural plasticity and cognitive ability in fish. Proceedings of the Royal Society B: Biological Sciences 280:20131331.

Shelton, A. O., and M. Mangel. 2012. Estimating von Bertalanffy parameters with individual and environmental variations in growth. Journal of Biological Dynamics 6(Suppl 2):3–30.

Shelton, A. O., W. H. Satterthwaite, M. P. Beakes, S. B. Munch, S. M. Sogard, and M. Mangel. 2013. Separating intrinsic and environmental contributions to growth and their population consequences. The American Naturalist 181:799–814.

Shine, R., and E. L. Charnov. 1992. Patterns of survival, growth, and maturation in snakes and lizards. The American Naturalist 139:1257.

Shohoji, T., K. Kanefuji, T. Sumiya, T. A. O. Qin, I. Arts, R. February, E. Bayes, I. S. M. Cooperative, and S. Mathematics. 1991. A prediction of individual growth of height according to an Empirical Bayesian approach. Annals of the Institute of Statistical Mathematics 43:607–619.

Siegfried, K. I., and B. Sansó. 2006. Two Bayesian methods for estimating parameters of the von Bertalanffy growth equation. Environmental Biology of Fishes 77:301–308.

Sigourney, D. B., S. B. Munch, and B. H. Letcher. 2012. Combining a Bayesian nonparametric method with a hierarchical framework to estimate individual and temporal variation in growth. Ecological Modelling 247:125–134.

Skaug, H. J., and D. A. Fournier. 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. Computational Statistics & Data Analysis 51:699–709.

Snover, M. L., G. M. Watters, and M. Mangel. 2005. Interacting effects of behavior and oceanography on growth in salmonids with examples for coho salmon (*Oncorhynchus kisutch*). Canadian Journal of Fisheries and Aquatic Sciences 62:1219–1230.

Snover, M. L., G. M. Watters, and M. Mangel. 2006. Top-down and bottom-up control of life-history strategies in coho salmon (*Oncorhynchus kisutch*). The American Naturalist 167:140–157.

Starck, J. M., and R. E. Ricklefs. 1998. Avian growth and development, J. M. Starck, and R. E. Ricklefs, editors. Oxford University Press, Oxford, UK.

Temming, A., and J. P. Herrmann. 2009. A generic model to estimate food consumption: linking von Bertalanffy's growth model with Beverton and Holt's and Ivlev's concepts of net conversion efficiency. Canadian Journal of Fisheries and Aquatic Sciences 66:683–700.

Tjørve, K. M. C., and E. Tjørve. 2010. Shapes and functions of bird-growth models: how to characterise chick postnatal growth. Zoology 113:326–33.

Vincenzi, S., A. J. Crivelli, D. Jesensek, J.-F. Rubin, G. Poizat, and G. A. De Leo. 2008. Potential factors controlling the population viability of newly introduced endangered marble trout populations. Biological Conservation 141:198–210.

Vincenzi, S., A. J. Crivelli, D. Jesensek, and G. A. De Leo. 2010. Detection of density-dependent growth at two spatial scales in marble trout (*Salmo marmoratus*) populations. Ecology of Freshwater Fish 19:338–347.

Vincenzi, S., A. J. Crivelli, D. Jesensek, and G. A. De Leo. 2012. Translocation of stream-dwelling salmonids in headwaters: insights from a 15-year reintroduction experience. Reviews in Fish Biology and Fisheries 22:437–455.

Vincenzi, S., A. J. Crivelli, W. H. Satterthwaite, and M. Mangel. 2014a. Eco-evolutionary dynamics induced by massive mortality events. Journal of Fish Biology 85:8–30.

Vincenzi, S., M. Mangel, A. J. Crivelli, S. Munch, and H. J. Skaug. 2014b. Determining individual variation in growth and its implication for life-history and population processes using the Empirical Bayes method. PLoS Computational Biology 10:e1003828.

Vincenzi, S., M. Mangel, D. Jesensek, J. C. Garza, and A. J. Crivelli. 2015. Within and among-population variation in vital rates and population dynamics in a variable environment. bioRxiv. http://dx.doi.org/10.1101/028662

Wang, Y., and M. R. Thomas. 1995. Accounting for individual variability in the von Bertalanffy growth model. Canadian Journal of Fisheries and Aquatic Sciences 1375:1368–1375.

Ward, E. J., E. E. Holmes, J. T. Thorson, and B. Collen. 2014. Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. Oikos 123:652–661.

West, G. B., J. H. Brown, and B. J. Enquist. 2004. Growth models based on first principles or phenomenology? Functional Ecology 18:188–196.

Woodson, L. E., B. K. Wells, P. K. Weber, R. B. MacFarlane, G. E. Whitman, and R. C. Johnson. 2013. Size, growth, and origin-dependent mortality of juvenile Chinook salmon Oncorhynchus tshawytscha during early ocean residence. Marine Ecology Progress Series 487:163–175.

Zullinger, E. M., and R. E. Ricklefs. 1984. Fitting sigmoidal equations to mammalian growth curves. Journal of Mammalogy 65:607–636.

Supporting Information

Data Availability