

Using nonlinear hierarchical models for analyzing annulus-based size-at-age data

G. Bruce Schaalje, Jared L. Shaw, and Mark C. Belk

Abstract: Size-at-age data for fish (derived from otoliths or other structures) are valuable but statistically messy. The data are typically serially correlated and unbalanced, with both time-independent and time-varying covariates. Appropriate growth models are typically nonlinear, with an unknown functional form. We recommend the use of nonlinear hierarchical models for the analysis of such data. We illustrate the use of these methods by applying the recently introduced SAS procedure NLMIXED (SAS Institute Inc., Cary, N.C.) to otolith-based estimated standard lengths of Utah chub (*Gila atraria*) collected in four locations with predators and four locations without predators.

Résumé : Les données de taille en fonction de l'âge chez les poissons (calculées à partir des otolithes ou d'autres structures) sont utiles, mais difficiles à traiter statistiquement. Ces données sont typiquement déséquilibrées et en corrélation séquentielle et elles possèdent des covariables dépendantes et indépendantes du temps. Les modèles de croissance qui leur conviennent sont généralement non linéaires et leur forme fonctionnelle est inconnue. Nous recommandons l'utilisation de modèles hiérarchiques non linéaires pour ces données. Nous en illustrons l'usage en appliquant la procédure récemment introduite de NLMIXED de SAS (SAS Institute Inc., Cary, N.C.) à des longueurs standard estimées à partir des otolithes chez des menés de l'Utah (*Gila atraria*) provenant de quatre sites sans prédateurs et de quatre sites avec prédateurs.

[Traduit par la Rédaction]

Introduction

Many studies of ecology and evolution of fish populations use size measurements derived from scale and otolith annuli of captured fish (Osenberg et al. 1988; Jennings and Philipp 1992; Johnson and Belk 1999). Yearly estimates of body size-at-age (typically standard or total length) are back-calculated from annulus measurements (Tesch 1968; Campana 1990). A great deal of size-at-age information can potentially be obtained from relatively few fish, and the method has proven to be a powerful tool for characterizing growth patterns, comparing these patterns among populations, sexes, etc., and relating these patterns to environmental variables.

The main problem with analysis of annulus-based growth estimates is that the data are messy in a statistical sense. Because the captured fish are invariably of different ages, the series of size measurements are of different lengths. Growth patterns of fish, or other animals, are usually best described by nonlinear rather than linear functions of age (Chen et al. 1992). The back-calculated sizes within fish are serially correlated. An additional complication is that some explanatory variables or covariates apply to the fish (independent of their age, e.g., sex) and some vary over time within fish (e.g., age). Hence, data from annulus-based growth studies can be described as unbalanced, nonlinear

repeated measures data, with both time-independent and time-varying covariates.

The statistical analysis of such data is further complicated by the fact that there is no consensus about which nonlinear function of age is appropriate for modeling fish growth (Chen et al. 1992). The choice of an appropriate nonlinear function often has to be made anew for each study because the functional form is dependent on such things as species, location, age-classes, and weight classes of the fish.

Several statistical approaches have been used in the analysis of annulus-based growth data. For example, to analyze differences among fish groups, one-way analysis of variance (ANOVA) and the one-way nonparametric Kruskal–Wallis procedure have been applied to size data separately for each age-class (Johnson et al. 1995; Belk 1998). Repeated measures ANOVA, treating age as a categorical variable, has been applied to complete series of size measurements for the fish (Jennings and Philipp 1992). Another approach has involved fitting a nonlinear, polynomial, or segmented linear model to the series of measurements for each fish and then comparing the parameter estimates among groups of fish using ANOVA or multivariate ANOVA (Johnson and Belk 1999). In a complicated analysis, Osenberg et al. (1988) computed growth rates as differences between annulus-based sizes for each size-class of each fish and then carried out factor analysis using the matrix of pairwise correlations of

Received 13 February 2002. Accepted 21 August 2002. Published on the NRC Research Press Web site at <http://cjfas.nrc.ca> on 23 September 2002.
J16761

G.B. Schaalje and J.L. Shaw. Department of Statistics, Brigham Young University, Provo, UT 84602, U.S.A.
M.C. Belk. Department of Zoology, Brigham Young University, Provo, UT 84602, U.S.A.

¹Corresponding author (e-mail: shaalje@byu.edu).

growth rates between size-classes. They examined the factor loadings for the various age-classes and interpreted them in terms of ecological factors.

Although valid, such approaches are incomplete, inefficient, and non-optimal. Some of these methods exclude data for certain fish or certain age-classes in order to deal with the unbalanced nature of the data. Some of them do not allow the use of nonlinear functions. Some ignore the serial correlations of the measurements. Potvin et al. (1990) emphasized the importance of using appropriate statistical tools for repeated measures data in ecological studies.

Another troubling issue in the statistical analysis of annulus-based growth data is that it is common for authors to apply several different statistical procedures to the same data set to answer different questions (Osenberg et al. 1988; Johnson and Belk 1999). Although not wrong, this practice can be confusing and inefficient. A more straightforward and productive general statistical paradigm would be first to select an appropriate model for the data using model fitting, model checking, and model comparison procedures and then to use the selected model to answer scientific questions by estimating appropriate functions of parameters of the selected model or by using the model to make predictions (McCullagh and Nelder 1989).

Fortunately, the last few years have seen great advances in statistical methodology for applying this general approach to exactly the type of data generated in annulus-based growth studies of fish. The methodology is associated with nonlinear mixed or hierarchical models. Reasonably accessible monographs have been written about the selection and analysis of nonlinear hierarchical models (Davidian and Giltinan 1995; Vonesh and Chinchilli 1996), and software has recently become available in major statistical packages to facilitate the use of these methods (SAS Institute Inc. 1999; Pinheiro and Bates 2000). McRoberts et al. (1998) recently recommended the use of nonlinear hierarchical models to investigate size-at-age relationships of animals and used these methods to analyze size-at-age data for black bears.

The methodology of nonlinear hierarchical models has apparently not been previously used for annulus-based fish growth data. The purpose of this paper is to describe the methodology and to apply it to data from a previously published fish growth study (Johnson and Belk 1999). In addition to demonstrating the methodology, a new nonlinear function for modeling fish growth is selected and new insights into Utah chub (*Gila atraria*) ecology are obtained.

Materials and methods

Nonlinear hierarchical models

Nonlinear hierarchical models are generalizations of both simple nonlinear statistical models and linear hierarchical models (Davidian and Giltinan 1995). The expected value of the response variable is modeled as a nonlinear function of known predictor variables, unknown fixed parameters, and unknown random quantities. The models are described as hierarchical because they accommodate data with a nested structure. For example, fish growth data consist of different numbers of repeated measurements for each of several individuals. Covariates and random effects are usually associated with all levels of the nested structure. The covariance

structure of the random effects is specified as part of the model.

Sheiner et al. (1972) first recognized the need to account for variation within and among subjects in nonlinear models for pharmacokinetic data and proposed early linearized versions of nonlinear mixed models. Beal and Sheiner (1982, 1988) and Sheiner and Beal (1985) refined and extended the linearization methods. Racine-Poon (1985) used the expectation-maximization (EM) algorithm to estimate parameters of nonlinear mixed models. Lindstrom and Bates (1990) proposed a general nonlinear mixed model for longitudinal data. They used a variation of the linearization method together with Newton-Raphson iteration to estimate the mean and variance parameters. Vonesh and Carter (1992) considered nonlinear models for repeated measures with a variety of covariance structures and used a four-step generalized least squares method of estimation. Davidian and Giltinan (1995) reviewed, synthesized, and extended nonlinear hierarchical models and associated methods.

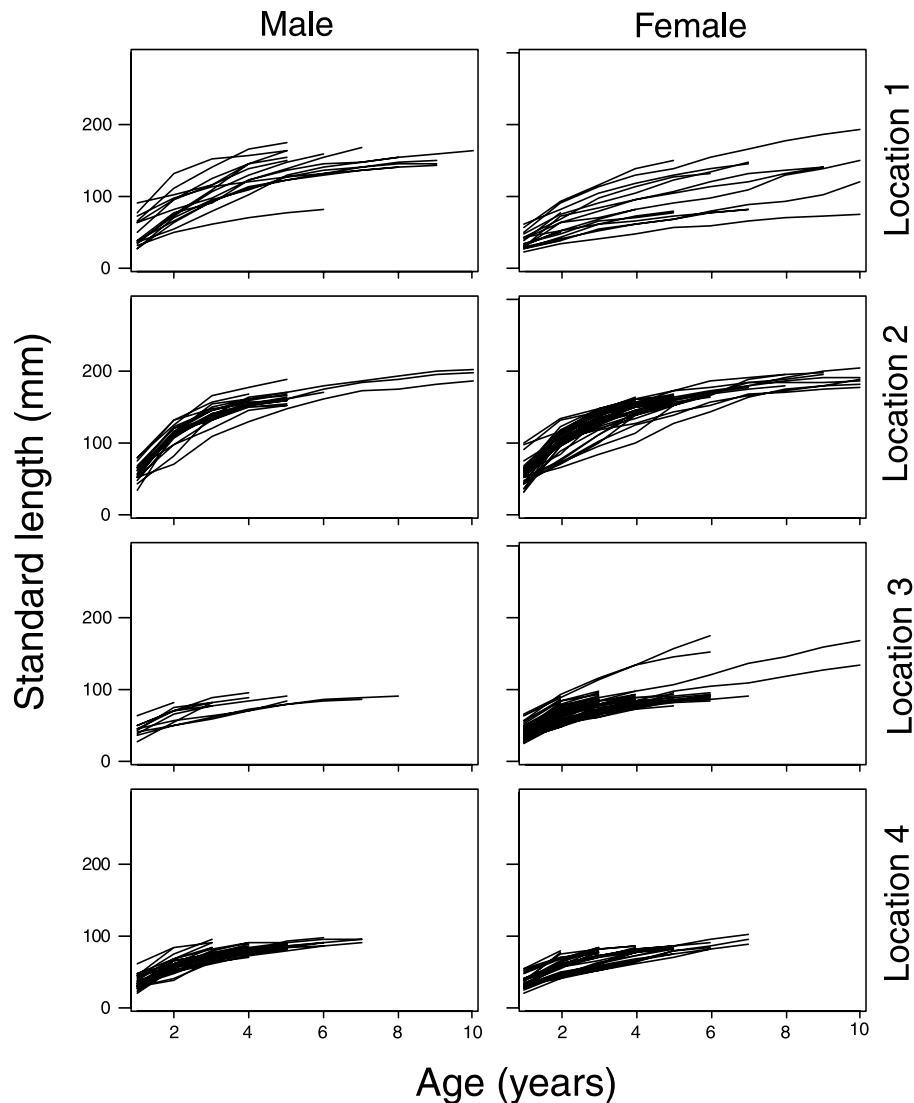
Nonlinear hierarchical models have received attention from researchers in the last decade because of the availability of high-speed computers and computationally intensive algorithms to find maximum likelihood estimates and their approximate covariance matrix. The likelihood function associated with a nonlinear hierarchical model can generally only be expressed as a product of several integrals involving conditional distributions, and therefore the iterative algorithm for maximizing the likelihood must include a numerical method for approximating these difficult integrals at each step. For low-dimensional integrals, Gaussian quadrature can be used. Pinheiro and Bates (1995) compared several different computational methods of finding maximum likelihood estimates for nonlinear mixed models (SAS Institute Inc. 1999).

Given the parameter estimates with their asymptotic covariance matrix, the delta method (Lehmann 1997) can be used to compute confidence intervals, prediction intervals, and hypothesis tests for nonlinear functions of the parameters.

A number of computer programs are currently available to carry out the extensive computations associated with nonlinear mixed-effect models. These include the NLME function of S-Plus (Pinheiro and Bates 2000), the SAS macro NLINMIX (Littell et al. 1996), and the SAS procedure NLMIXED (SAS Institute Inc. 1999). The NLMIXED procedure of SAS, the most recently developed of these programs, fits two-level nonlinear mixed effects models that are analogous to two-level linear random-coefficients models. It is unique in several ways. NLMIXED allows the conditional distribution of the response variable to follow the normal, binomial, or Poisson distribution. Several iterative optimization techniques are included in the SAS procedure as well as several integral approximation methods for evaluating the likelihood at each iteration. The ESTIMATE statement of NLMIXED can be used to estimate arbitrary linear or nonlinear functions of the fixed parameters along with their standard errors. The procedure also computes empirical Bayes estimates of random effects in the model.

To facilitate the selection of an appropriate nonlinear function on which to base the nonlinear hierarchical model, NLMIXED computes the Akaike information criterion (AIC; Akaike 1974) and the Bayesian information criterion (BIC;

Fig. 1. Otolith-based standard lengths of male and female Utah chub (*Gila atraria*) at four locations without predators. Each line represents a fish.



Schwarz 1978) for each fitted model. These criteria use the maximized likelihood penalized by the number of parameters in the model to indicate how well a model fits the data. Various candidate nonlinear functions can be used in developing nonlinear hierarchical models, and it makes sense to select the function producing the best value of the AIC or BIC (Littell et al. 1996).

Utah chub data

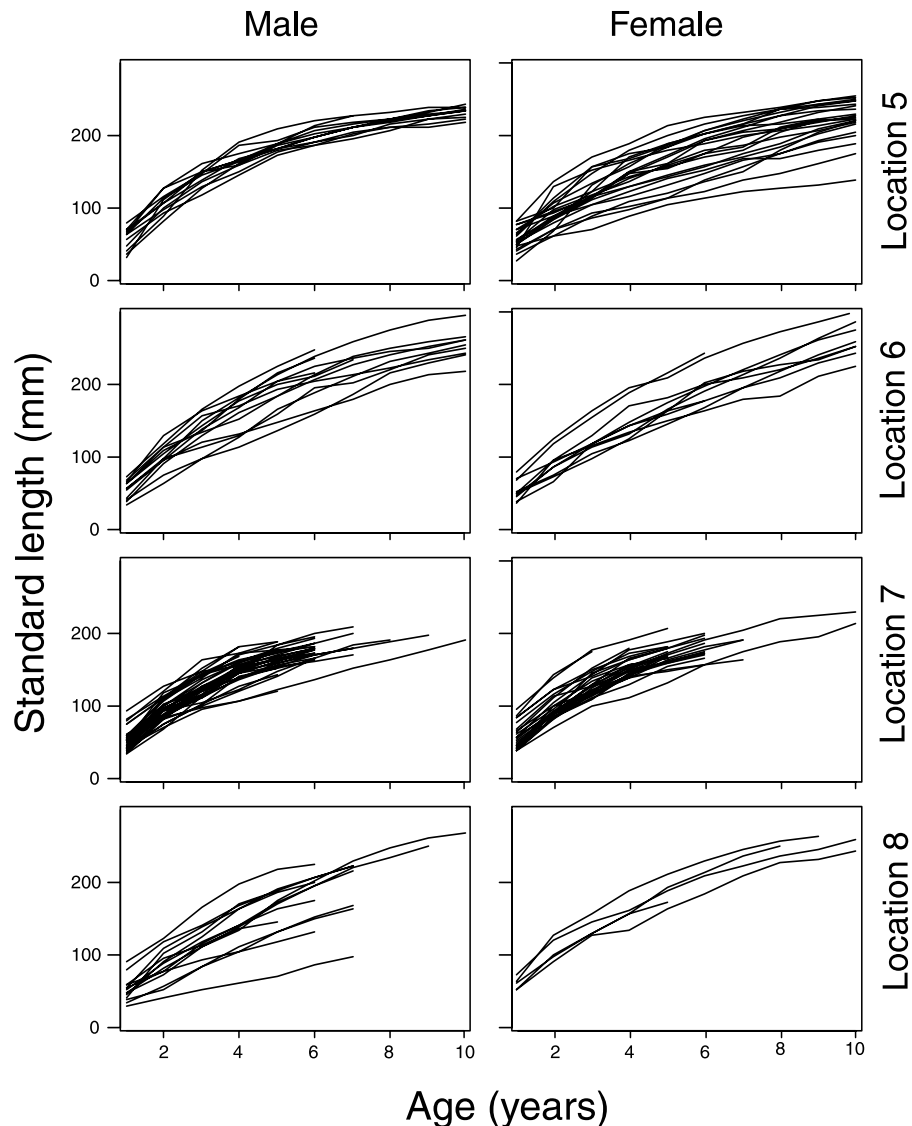
Details of the fish collection and otolith measurement methods were given by Johnson and Belk (1999). Briefly, Utah chub were netted in eight locations in Wyoming and Utah, U.S.A. Sample sizes for adult Utah chub from the eight locations ranged from 19 to 67, and fish ages ranged from 1 to 29 years, but annulus measurements were made for the first 10 years only. Four of the locations were free of chub predators; predators were present in the other four locations. Back-calculated standard lengths for each age were generated using a modified Fraser–Lee formula (after Campana 1990). Profile plots for otolith-based standard

lengths of the sampled fish by sex and location show that the growth curves all closely follow a nonsigmoid asymptotic shape, but with considerable variation in slopes and asymptotes among fish (Figs. 1, 2). The purpose of the study was to compare predator and nonpredator locations with regards to such life history traits as age at maturity, standard length at maturity, juvenile growth rate, and asymptotic standard length. Using three different statistical approaches, Johnson and Belk (1999) concluded that Utah chub in predator locations grew faster, reached a larger adult size, and matured at later ages and at larger sizes than did fish in nonpredator locations. This was consistently true for all locations except one nonpredator location (location 2), where the average growth rate was high compared with the other nonpredator locations. Johnson and Belk (1999) made no distinctions between sexes.

Statistical analysis of the Utah chub data

Our analysis followed the general statistical paradigm of model selection and checking followed by use of the se-

Fig. 2. Otolith-based standard lengths of male and female Utah chub (*Gila atraria*) at four locations with predators. Each line represents a fish.



lected model to answer scientific questions. For most nonlinear hierarchical modeling, the appropriate nonlinear function upon which the statistical model should be based is unknown and so the model selection stage involves two steps: selection of an appropriate nonlinear function and then selection of an appropriate statistical model based on the selected nonlinear function. We used the SAS procedure NLMIXED for all calculations associated with nonlinear hierarchical models.

Nonlinear function selection

An appropriate nonlinear function of age upon which to base the statistical model for Utah chub growth was selected by proposing a number of candidate growth functions with approximately the desired asymptotic shape, fitting a nonlinear hierarchical model based on each of these candidate functions to the Utah chub data, and calculating AIC and BIC for each model. To focus AIC and BIC on selection of the nonlinear function rather than on other aspects of the statistical model, nonlinear hierarchical models in this stage

should include all or most of the potential statistical complexity. Also, the candidate nonlinear functions to be compared should be of comparable complexity.

As candidate nonlinear functions for use in developing a statistical model for the Utah chub data, four three-parameter functions discussed by Ratkowsky (1990) were used, all of which exhibit the general nonsigmoidal increasing shape displayed by the chub sizes (Figs. 1, 2) in at least part of their range:

- (1) $\alpha - \beta\gamma^x$
- (2) $\alpha/(1 + \beta x^{-\gamma})$
- (3) $\alpha x^{\beta x^{-\gamma}}$
- (4) $\alpha(1 - \beta^{-x})^\gamma$

In all of these functions, x is age and the parameters α , β , and γ are positive real numbers. Function 1 is the three-parameter von Bertalanffy growth function (von Bertalanffy

Table 1. Akaike information criterion (AIC) and Bayesian information criterion (BIC) values for the four candidate nonlinear functions describing growth of Utah chub.

Function	Information criterion	
	AIC	BIC
1	1369.3	1299.5
2	1459.9	1389.9
3	1501.1	1431.0
4	1434.1	1364.1

Note: Larger values indicate better fit.

Table 2. Likelihood ratio tests comparing six simpler models to the full statistical model.

Model	χ^2 value	df	<i>p</i> value
Additive	128	21	<0.0001
Constant α	213	15	<0.0001
Constant β	55	15	<0.0001
Constant γ	251	15	<0.0001
Nonrandom α	1352	2	<0.0001
Nonrandom β	705	2	<0.0001

1957), which has been commonly used historically for modeling growth of animals (McRoberts et al. 1998). However, there is no theoretical reason to believe that the von Bertalanffy growth function would be best or even appropriate for size-at-age relationships for fish (Chen et al. 1992).

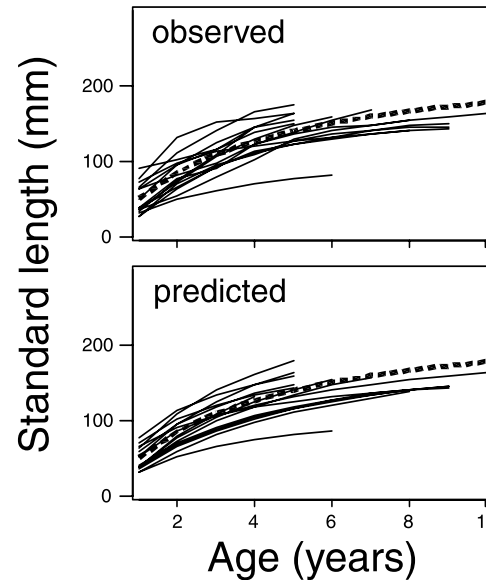
Relatively complex two-level nonlinear hierarchical statistical models were formulated based on each of these candidate functions. As an example, the statistical model based on function 3 was

$$y_{ijkl} = (\alpha_{ij} + a_{ijk})x_{ijkl}^{(\beta_{ij} + b_{ijk})x_{ijkl}^{-\gamma_{ij}}} + \varepsilon_{ijkl}$$

where y_{ijkl} is the l th standard length and x_{ijkl} the l th age of the k th Utah chub of sex i in location j . It was assumed that $a_{ijk} \sim n(0, \sigma_a^2)$, $b_{ijk} \sim n(0, \sigma_b^2)$, $\text{cov}(a_{ijk}, b_{ijk}) = \sigma_{ab}$, and $\varepsilon_{ijkl} \sim n(0, \sigma_\varepsilon^2)$. This model thus included different fixed parameters α_{ij} , β_{ij} , and γ_{ij} for each sex–location combination, normally distributed and correlated random effects a_{ijk} and b_{ijk} for each fish within each sex–location combination, and a normally distributed random error associated with each standard length estimate for each fish. All of the models included age as a time-varying covariate, and sex and location as time-independent covariates. Hierarchical models based on the other candidate functions were formulated analogously.

Starting values for the parameter estimates were needed in order to fit these models and obtain maximum likelihood estimates of the parameters using NLMIXED. To find reasonable starting values, MAPLE V was used to plot the candidate functions for ranges of values of the parameters. Proc NLIN of SAS was then used to fit models based on the nonlinear functions to the data for each chub having at least four measurements. Means and variances of these parameter estimates were calculated and used as starting values for

Fig. 3. Otolith-based standard lengths as well as model-based predicted standard lengths of male Utah chub (*Gila atraria*) at location 1. Each line represents a fish. The growth curve based on the estimated parameters for males in location 1 (broken line) is superimposed on the plots.



estimation of the fixed parameters and variance–covariance parameters of the random effects.

Statistical model selection

Once the nonlinear function of age was selected, NLMIXED was used to choose among several versions of the nonlinear hierarchical model of varying levels of statistical complexity. The full statistical model, also used in the preceding section, included fixed parameters α_{ij} , β_{ij} , and γ_{ij} for all sex–location combinations for a total of 48 fixed parameters. This model also involved the four variance–covariance parameters σ_a^2 , σ_b^2 , σ_{ab} , and σ_ε^2 .

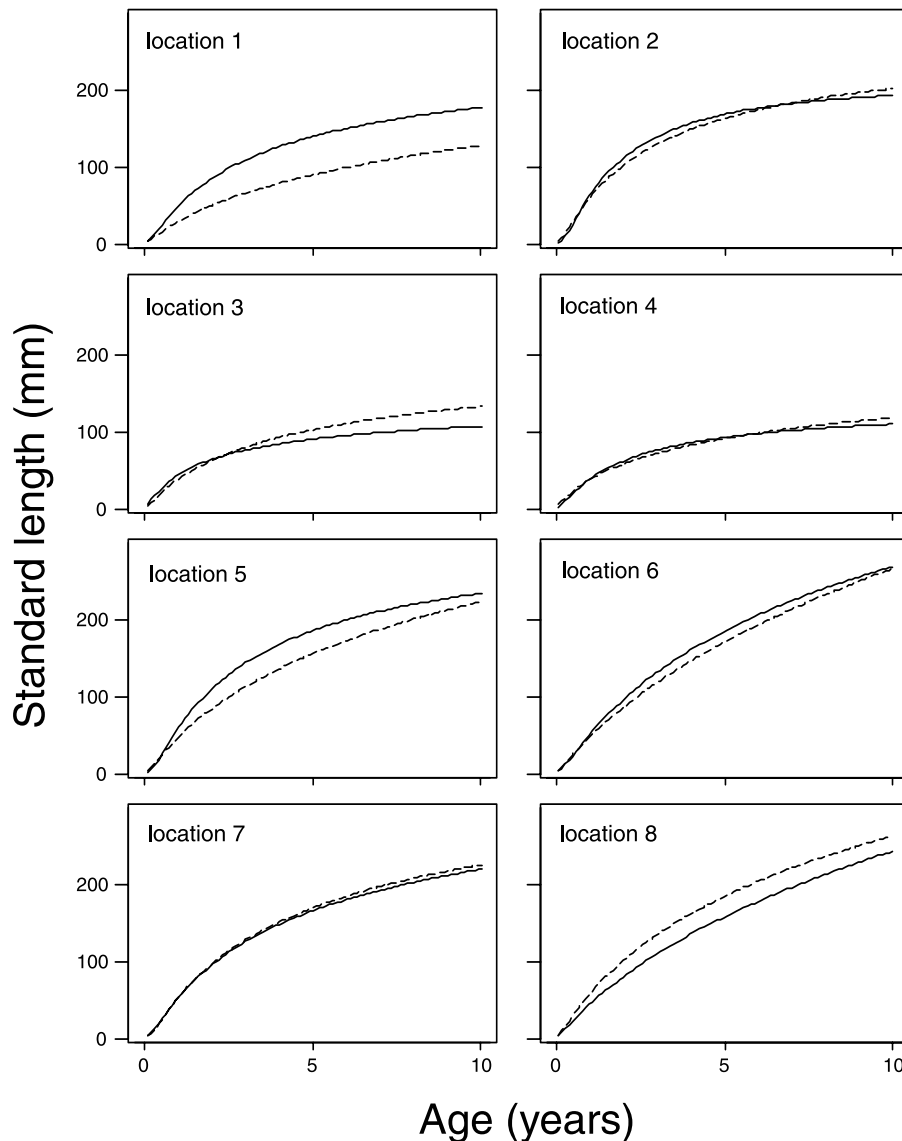
Six reduced statistical models were considered. The first of these, the additive model, had the same variance–covariance structure as the full model and thus involved the same four variance–covariance parameters. However, the parameters α_{ij} , β_{ij} , and γ_{ij} were constrained so that the effects of sex and location were additive. For example, α_{ij} was modeled as $\alpha + \theta_i + \phi_j$, and similar models were set up for β_{ij} and γ_{ij} . This model involved nine fixed parameters corresponding to each of α , β , and γ for a total of 27 fixed parameters.

The constant α model involved a single parameter α for all sex–location combinations but different parameters β_{ij} and γ_{ij} for each sex–location combination (33 fixed parameters, four variance–covariance parameters). The constant β and constant γ models were analogous to the constant α model.

The nonrandom α model had no random effects a_{ijk} (48 fixed parameters, two variance–covariance parameters). The nonrandom β model had no random effects b_{ijk} (48 fixed parameters, two variance–covariance parameters).

The goal in selection of a statistical model was to choose the most parsimonious well-fitting model. To choose between the full model and any of the reduced versions, likelihood ratio tests were carried out by calculating the difference between -2 times the log likelihood for each model

Fig. 4. Growth curves based on estimated parameters for Utah chub (*Gila atraria*) by location and sex: male, solid line; female, broken line.



and the full model. These test statistics are approximate chi-square random variables under the null hypothesis (i.e., if the reduced model is correct), with degrees of freedom calculated as the difference in the number of parameters for the two models (Lehmann 1997). The test is easily computed because NLMIXED automatically calculates the log likelihood for each model.

Using the selected model

After selecting and fitting an appropriate nonlinear statistical model, the model was used to carry out growth rate comparisons. For the particular nonlinear function selected for the Utah chub data, only one of the parameters had a natural biological interpretation. However, desired scientific questions could still be investigated using carefully chosen functions of the parameters. As an example, we considered the size at age 4 to be a meaningful indicator of the growth rate for each sex–location combination. This is a nonlinear function of the three parameters α_{ij} , β_{ij} , and γ_{ij} found by substituting $x = 4$ years, as well as sex and location information,

into the selected nonlinear hierarchical model. This function was estimated by replacing the parameters in the function by their estimates, so that for function 3 the estimate of size at age 4 for a particular sex–location combination would be

$$\hat{\alpha}_{ij}(4^{\hat{\beta}_{ij}(4^{-\hat{\gamma}_{ij}})})$$

The approximate standard error of this estimate was obtained using the delta method (Lehman 1997), which involves a first-order Taylor series approximation evaluated at the estimated asymptotic variance–covariance matrix of the parameter estimates. These estimation and delta method calculations are carried out automatically by invoking the ESTIMATE statement of the NLMIXED procedure.

Using exactly the same general methods, more complicated functions of the parameters were also estimated along with their standard errors. Male–female differences in the average size at age 4 were estimated for each of the locations. For function 3, the average male–female difference in size at age 4 for location j would be estimated as

Table 3. Estimates (with standard errors, SE) of parameters of the statistical model based on function 3, as well as estimated standard length at age 4 (SE) for both sexes at every location for Utah chub (*Gila atraria*).

Sex	Predators	Location	$\hat{\alpha}_{ij}$	$\hat{\beta}_{ij}$	$\hat{\gamma}_{ij}$	Size at age 4 (mm)
M	No	1	49.1 (2.5)	0.923 (0.051)	0.217 (0.020)	126.8 (3.8)
		2	61.3 (2.6)	1.078 (0.047)	0.334 (0.014)	157.0 (3.1)
		3	43.6 (4.0)	0.634 (0.102)	0.210 (0.074)	84.2 (4.1)
		4	37.0 (2.2)	0.866 (0.067)	0.263 (0.034)	85.1 (2.9)
	Yes	5	59.0 (3.3)	1.075 (0.058)	0.252 (0.011)	168.8 (4.3)
		6	49.1 (3.0)	1.062 (0.047)	0.156 (0.012)	160.7 (5.3)
		7	53.0 (1.8)	0.981 (0.036)	0.199 (0.013)	148.7 (2.5)
		8	42.8 (2.7)	0.951 (0.045)	0.099 (0.018)	135.0 (5.3)
F	No	1	28.6 (2.3)	0.865 (0.052)	0.123 (0.025)	78.5 (5.1)
		2	58.2 (1.7)	0.949 (0.032)	0.243 (0.011)	148.8 (2.1)
		3	38.0 (1.6)	0.823 (0.040)	0.179 (0.022)	92.5 (2.2)
		4	36.6 (2.2)	0.715 (0.059)	0.147 (0.037)	82.2 (2.7)
	Yes	5	46.5 (2.3)	0.947 (0.034)	0.141 (0.010)	136.7 (4.0)
		6	45.7 (3.5)	0.954 (0.052)	0.094 (0.015)	145.8 (6.2)
		7	52.8 (2.1)	1.009 (0.039)	0.202 (0.014)	151.9 (3.1)
		8	55.9 (5.5)	0.934 (0.083)	0.141 (0.022)	162.1 (7.3)

Note: M, male; F, female.

Table 4. Estimated average differences (male minus female) in standard length at age 4 (in mm) with standard errors (SE) of Utah chub (*Gila atraria*) at each location.

Location	Predators	Difference	SE	<i>t</i>	<i>p</i> value
1	No	48.2	6.3	7.66	<0.0001
2		8.2	3.7	2.18	0.0296
3		-8.3	4.7	-1.78	0.0751
4		2.9	4.0	0.73	0.4639
5	Yes	32.1	5.9	5.46	<0.0001
6		14.9	8.1	1.84	0.0660
7		-3.1	4.0	-0.79	0.4311
8		-27.1	9.0	-3.00	0.0029

$$\hat{\alpha}_{1j}(4^{\hat{\beta}_{1j}(4^{-\hat{\gamma}_{1j}})}) - \hat{\alpha}_{2j}(4^{\hat{\beta}_{2j}(4^{-\hat{\gamma}_{2j}})})$$

Similarly, average predator–nonpredator location differences were estimated overall and for each sex. Because locations were treated here as fixed effects, our inferences about the difference between predator and nonpredator locations applies to these particular locations, not to the general populations of such locations.

Other possibly interesting indicators of growth patterns are the age at which a specified size is achieved, the derivative of the size-at-age function at a specified size or age, the size difference between two specified ages, and the age at which the derivative of the size-at-age function is maximized. If these functions of the parameters of the model exist, they can similarly be estimated and compared statistically.

Results

Nonlinear function selection

Difficulties were encountered when we attempted to fit the hierarchical model based on the von Bertalanffy growth

function to the data. Regardless of starting values, the procedure would not converge. Through a process of trial and error, including fitting the model to each group separately, it was discovered that if the males and females for location 3 and the males for location 4 were left out, the hierarchical model based on the von Bertalanffy growth function could be fit. Hence, the decision was made to eliminate locations 3 and 4 when fitting all of the nonlinear models to the data so the AIC and BIC values would be comparable among the nonlinear models. The AIC and BIC values both indicated that function 3 should be selected as the nonlinear function upon which to base the statistical model (Table 1). Interestingly, the model based on the von Bertalanffy growth function had the worst values of AIC and BIC.

Statistical model selection

Likelihood ratio tests based on the complete data set (including locations 3 and 4) indicated that the full statistical model based on function 3 (with different values α_{ij} , β_{ij} , and γ_{ij} for all sex–location combinations) was significantly better than the additive, constant α , constant β , constant γ , nonrandom α , and nonrandom β models (Table 2). Estimates and standard errors of the 48 fixed parameters of the nonlinear hierarchical model were computed by NLMIXED (Table 3), as were estimates of the variance–covariance parameters σ_a^2 , σ_b^2 , σ_{ab} , and σ_ϵ^2 . These latter estimates were 150.0, 0.024, -0.017, and 40.0, respectively. The predicted growth curves agreed extremely well with the observed data (Fig. 3), and the estimated growth curves for all sex–location groups exhibited interesting differences (Fig. 4).

Growth rate comparisons

Using the full model, the average size at age 4 was estimated for each sex–location group (Table 3). Male–female differences in the average size at age 4 varied among locations (Table 4). In three of the locations, two without predators and one with predators, males were significantly larger than females (using *p* value ≤ 0.05). Females were signifi-

cantly larger at age 4 than males in one predator location. In the remaining four locations, the male–female differences in size at age 4 were not significant.

In general, fish at age 4 were significantly larger in locations with predators than in locations with no predators. Male chubs were, on average, 11.8 mm larger in predator locations than in nonpredator locations. The standard error (SE) of this difference was 5.2 mm. Female chubs were estimated to be 25.0 mm (SE = 5.8 mm) larger in predator than in nonpredator locations. Location 2 was the only nonpredator location in which the estimated size at age 4 was comparable to values for predator locations.

Discussion

Comparison to previous analyses of the Utah chub data

Analysis of the Utah chub data based on a nonlinear hierarchical model gave similar conclusions to those obtained by Johnson and Belk (1999). Both analyses revealed significant differences in the growth patterns of Utah chub living in the presence of predators versus those living in locations without predators. Both analyses also indicated that location 2, which had no predators, had growth patterns similar to predator locations. Agreement between the current and previous results is expected because the previous approach (of fitting a model—von Bertalanffy growth model or two-phase regression—to the data for each fish and then applying ANOVA to the parameter estimates for the fish) is similar to a nonlinear hierarchical model. The main difference is that the former procedure is ad hoc and inefficient. The methodology of nonlinear hierarchical models combines the data from all of the fish optimally and produces explicit estimates of relevant variance–covariance parameters.

Our analysis selected a nonlinear function (function 3) of age that apparently has not been used previously in fish growth studies. The use of models based on function 3 in other size-at-age studies should be considered, at least if theoretical considerations do not suggest the use of any other function. Neither β nor γ has a natural biological interpretation, but α is the size at age 1. However, as noted above, relevant biological variables can be defined as functions of fixed parameters of the model and then estimated and compared statistically.

Analysis based on the nonlinear hierarchical model revealed that the relationship between growth patterns, sexes, and locations is very complicated for the Utah chub. The presence of predators was associated with higher growth rates in general, but the effect was quite variable. Apparently, other environmental factors also affect growth rates. Additionally, there was a significant male–female difference in growth rates at some locations, but this effect was not consistently associated with the presence or absence of predators. The results could suggest interesting evolutionary and ecological hypotheses.

Nonlinear hierarchical models for annulus-based size-at-age data

Using nonlinear hierarchical models in the analysis of the annulus-based size-at-age data for the Utah chub was labor-intensive. We had to carefully study the user instructions for the NLMIXED procedure, including numerical op-

tions and optional statements such as those giving starting values and bounds for the parameters estimates. Trial and error were required to obtain appropriate starting values and successful numerical methods. Even after resolving these computational problems, several model-fitting runs were needed to compare various candidate nonlinear functions and statistical models and to estimate and compare the desired functions of the parameters. In many cases, it would be advisable to obtain expert statistical help when analyzing size-at-age data with nonlinear hierarchical models.

Nonetheless, the approach had several advantages. Nonlinear hierarchical models easily handled unbalanced and incomplete repeated measures data. Models based on alternate nonlinear functions were readily compared, and a statistical model of an appropriate degree of complexity was objectively selected. Estimates of variance and covariance components were used in calculations of tests and confidence intervals for interesting functions of the parameters.

The availability of software and technology for nonlinear hierarchical models allows annulus-based size-at-age data to be analyzed using the general statistical approach of model selection followed by use of the selected model to answer scientific questions. This approach simplifies the statistical analysis, leads to valid and defensible statistical results, and simplifies the reporting of methods and results. We recommend the general use of these methods for analyzing annulus-based size-at-age data.

Acknowledgments

We thank Jerald Johnson for the use of data from his M.S. thesis.

References

- Akaike, H. 1974. A new look at statistical model identification. IEEE (Inst. Electr. Electron. Eng.) Trans. Automatic Control, AC-19: 6.
- Beal, S.L., and Sheiner, L.B. 1982. Estimating population kinetics. CRC Crit. Rev. Biomed. Eng. **8**: 195–222.
- Beal, S.L., and Sheiner, L.B. 1988. Heteroskedastic nonlinear regression. Technometrics, **30**: 327–338.
- Belk, M.C. 1998. Age and growth of June sucker (*Chasmistes liorus*) from otoliths. Great Basin Nat. **58**: 390–392.
- Campana, S.E. 1990. How reliable are growth back-calculations based on otoliths? Can. J. Fish. Aquat. Sci. **47**: 2219–2227.
- Chen, Y., Jackson, D.A., and Harvey, H.H. 1992. A comparison of von Bertalanffy and polynomial functions in modeling fish growth data. Can. J. Fish. Aquat. Sci. **49**: 1228–1235.
- Davidian, M., and Giltinan, D.M. 1995. Nonlinear models for repeated measurement data. Chapman & Hall, London.
- Jennings, M.J., and Philipp, D.P. 1992. Reproductive investment and somatic growth rates in longear sunfish. Environ. Biol. Fishes, **35**: 257–271.
- Johnson, J.B., and Belk, M.C. 1999. Effects of predation on life-history evolution in Utah chub (*Gila atraria*). Copeia, **4**: 948–957.
- Johnson, J.B., Belk, M.C., and Shiozawa, D.K. 1995. Age, growth, and reproduction of leatherside chub (*Gila copei*). Great Basin Nat. **55**: 183–187.
- Lehmann, E.L. 1997. Elements of large-sample theory. Springer-Verlag, Inc., New York.

- Lindstrom, M.J., and Bates, D.M. 1990. Nonlinear mixed effects models for repeated measures data. *Biometrics*, **46**: 673–687.
- Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. 1996. SAS system for mixed models. SAS Institute Inc., Cary, N.C.
- McCullagh, P., and Nelder, J.A. 1989. Generalized linear models. 2nd ed. Chapman and Hall, New York.
- McRoberts, R.E., Brooks, R.T., and Rogers, L.L. 1998. Using nonlinear mixed effects models to estimate size–age relationships for black bears. *Can. J. Zool.* **76**: 1098–1106.
- Osenberg, C.W., Werner, E.E., Mittelbach, G.G., and Hall, D.J. 1988. Growth patterns in bluegill (*Lepomis macrochirus*) and pumpkinseed (*L. gibbosus*) sunfish: environmental variation and the importance of ontogenetic niche shifts. *Can. J. Fish. Aquat. Sci.* **45**: 17–26.
- Pinheiro, J.C., and Bates, D.M. 1995. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *J. Comput. Graphical Stat.* **4**: 12–35.
- Pinheiro, J.C., and Bates, D.M. 2000. Mixed effects models in S and S-Plus. Springer-Verlag, Inc., New York.
- Potvin, C., Lechowicz, M.J., and Tardif, S. 1990. The statistical analysis of ecophysiological response curves obtained from experiments involving repeated measures. *Ecology*, **71**: 1389–1400.
- Racine-Poon, A. 1985. A Bayesian approach to nonlinear random effects models. *Biometrics*, **41**: 1015–1023.
- Ratkowsky, D.A. 1990. Handbook of nonlinear regression models. Marcel Dekker, New York.
- SAS Institute Inc. 1999. The NL MIXED procedure, SAS/STAT® online documentation. Ver. 8. SAS Institute Inc., Cary, N.C.
- Schwarz, G. 1978. Estimating the dimension of a model. *Ann. Stat.* **6**: 461–464.
- Sheiner, L.B., and Beal, S.L. 1985. Pharmacokinetic parameter estimates from several least squares procedures: superiority of extended least squares. *J. Pharmacokinet. Biopharm.* **13**: 185–201.
- Sheiner, L.B., Rosenberg, B., and Melmon, K.L. 1972. Modeling of individual pharmacokinetics for computer-aided drug dosing. *Comput. Biomed. Res.* **5**: 441–459.
- Tesch, F.W. 1968. Age and growth. In *Methods for assessment of fish production in fresh waters. Edited by W.E. Ricker.* Blackwell Scientific, Oxford, U.K. pp. 93–123.
- von Bertalanffy, L. 1957. Quantitative laws in metabolism and growth. *Q. Rev. Biol.* **32**: 217–231.
- Vonesh, E.F., and Carter, R.L. 1992. Mixed effects nonlinear regression for unbalanced repeated measures. *Biometrics*, **48**: 1–18.
- Vonesh, E.F., and Chinchilli, V.M. 1996. Linear and nonlinear models for the analysis of repeated measurements. Marcel Dekker, New York.