

# A new likelihood for simultaneously estimating von Bertalanffy growth parameters, gear selectivity, and natural and fishing mortality

Nathan G. Taylor, Carl J. Walters, and Steven J.D. Martell

**Abstract:** Gear selectivity and the cumulative effects of size-selective fishing produce bias in the length-at-age samples used to estimate the von Bertalanffy growth parameters. In fished populations, fast-growing young fish and slow-growing old fish are overrepresented in size-age samples. To account for such effects, we treated size-at-age observations as multinomial samples, with expected catches in each size-age category dependent on growth parameters, growth variation, size selectivity, abundance at age, and the history of exploitation. Using simulated data sets, estimated growth parameters using the multinomial likelihood were unbiased when fishing mortality was not too high and the shape of the vulnerability function was correct. In contrast, estimated growth parameters using a least squares approach overestimated the metabolic growth coefficient ( $K$ ) and underestimated mean asymptotic length ( $L_{\infty}$ ). Models that do not explicitly account for the effects of fishing and size selectivity underestimated  $L_{\infty}$  and overestimated  $K$ . We estimate growth parameters for northern pikeminnow (*Ptychocheilus oregonensis*) as an example of the method and document a stunted "pigmy" population with an  $L_{\infty}$  of 175-mm fork length, attributing its small size to effects of high density and (or) a short growing season.

**Résumé :** La sélectivité des engins de pêche et les effets cumulatifs de la pêche sélective en fonction de la taille faussent les échantillons de longueurs en fonction de l'âge utilisés pour estimer les paramètres de croissance de von Bertalanffy. Dans les populations soumises à la pêche, les jeunes poissons à croissance rapide et les poissons âgés à croissance lente sont surreprésentés dans les échantillons de longueurs en fonction de l'âge. Pour tenir compte de ces effets, nous avons traité les observations de longueurs en fonction de la taille comme des échantillons multinomiaux, avec des captures attendues dans chaque catégorie de taille en fonction de l'âge dépendantes des paramètres de croissance, de la variation de la croissance, de la sélectivité en fonction de la taille, de l'abondance en fonction de l'âge et de l'exploitation faite au cours des ans. À l'aide de séries de données simulées, nous montrons que les paramètres de croissance estimés d'après la vraisemblance multinomiale ne sont pas faussés quand la mortalité due à la pêche n'est pas trop importante et lorsque la forme de la fonction de vulnérabilité est exacte. En revanche, le calcul des paramètres de croissance estimés par la méthode des moindres carrés surestime le coefficient de croissance métabolique ( $K$ ) et sous-estime la longueur moyenne à l'asymptote ( $L_{\infty}$ ). Les modèles qui ne tiennent pas compte de façon explicite des effets de la pêche et de la sélectivité en fonction de la taille sous-estiment  $L_{\infty}$  et surestiment  $K$ . À titre d'exemple de notre méthodologie, nous faisons l'estimation des paramètres de croissance du méné-brochet du nord, *Ptychocheilus oregonensis*, et nous présentons une population « pygmée » à croissance ralentie qui possède un  $L_{\infty}$  de 175 mm de longueur à la fourche, un effet que nous attribuons à la forte densité et (ou) à la courte saison de croissance.

[Traduit par la Rédaction]

## Introduction

Estimating parameters for the von Bertalanffy growth model assumes that a representative sample of lengths from each age-class has been collected and then fitting the growth function either to length-age data directly or to data on changes in length between the time of marking and time of recapture. This assumption fails because the sampling process is size selective and because of the cumulative effects of fishing on size-at-age. Growth rates vary among individu-

als (Sainsbury 1980; Burr and Doksum 1988) and intensive size-selective harvesting removes faster growing individuals. The result is that length-age samples are biased for larger fast-growing young individuals and biased for older slow-growing individuals that avoided capture (Lee 1912; Ricker 1969; Sinclair et al. 2002a). In our experience, almost all length-age data sets collected for growth analysis involve size-selective capture methods that favor fast-growing individuals. In addition, many come from populations that have historically been exposed to harvest so as to produce cumu-

Received 22 September 2003. Accepted 16 July 2004. Published on the NRC Research Press Web site at <http://cjfas.nrc.ca> on 10 March 2005.  
J17744

N. Taylor, C.J. Walters,<sup>1</sup> and S.J.D. Martell. Fisheries Center, University of British Columbia, 2259 Lower Mall, Vancouver, BC V6T 1Z4, Canada.

<sup>1</sup>Corresponding author (e-mail: [n.taylor@fisheries.ubc.ca](mailto:n.taylor@fisheries.ubc.ca)).

lative effects on size structure at age (Hansen and Chouinard 1992; Kristiansen and Svasand 1998). In both cases, the result is a downward bias in estimation of mean asymptotic length ( $L_\infty$ ), upward bias of the metabolic growth parameter ( $K$ ), and upward bias in the apparent size at hatching ( $t_0$ ).

Biased estimates of growth parameters will distort bias mortality rate estimates (where length–age keys are used to convert length to age), biological reference points for management such as yield per recruit (Ricker 1969; Parma and Deriso 1990), and management actions (quotas, for example). Changes in size selectivity over the course of a fishery complicate matters further, where size selectivity changes may be interpreted as changes in growth rates. Sinclair et al. (2002b) documented changes in size selectivity for the Atlantic cod (*Gadus morhua*) fishery in the southern Gulf of St. Lawrence, and Sinclair et al. (2002a) noted that historical effects of temperature on measured annual growth increment and density-dependent increases in growth rate were relatively small in comparison with the effect of size-selective mortality caused by the fishery.

In some cases, researchers wish to solely estimate growth parameters describing the current population available to harvest with the understanding that these growth parameters represent the remaining fish that were targeted. In this case, the effects of size-selective fishing do not need to be accounted for. However, growth parameters describing the harvested targeted population will change as vulnerability changes (as fishers target smaller fish, for example). If a researcher wishes to estimate the potential growth under no harvesting, then growth parameters representing the entire population are needed.

Recently, Laslett et al. (2002) derived a likelihood function for analysis of growth data from tagging experiments that accounts for individual variation in growth through variation in individual asymptotic lengths  $L_\infty$ , measurement errors, and possible changes in the von Bertalanffy metabolic parameter  $K$  with fish age. Unfortunately, they follow previous workers (James 1991; Palmer et al. 1991; Wang et al. 1995) in assuming away one of the most common sources of bias in growth data; they assumed independence of sampled individual  $L_\infty$  and age despite much evidence that higher  $L_\infty$  (fast-growing) individuals have a higher capture probability in fishing and sampling gear used to collect the length–age samples.

To deal with nonindependence of age and individual  $L_\infty$ , likelihood functions for the analysis of length–age samples should include parameters representing probabilities of capture-at-age and probabilities of capture-at-length (i.e., size selectivity). Where applicable, the cumulative effects of size-selective harvesting on apparent growth should also be considered. In this paper, we derive likelihood functions that include size selectivity, mortality, and growth parameters based on the assumption that size–age data are sampled from a multinomial distribution. We derive four likelihoods: (i) a full likelihood and (ii) a “reduced” likelihood using the conditional maximum likelihood estimates for numbers-at-age instead of estimating natural mortality  $M$ , (iii) a full likelihood that includes cumulative effects of fishing mortality  $F$ , and (iv) a reduced likelihood with cumulative effects of  $F$ . We test the performance of these formulations with 100 sim-

ulated data sets and show that they are better estimators of growth parameters than previous statistical models for growth data. Finally, we fit all four models to real length-at-age data from the pigmy northern pikeminnow (*Ptychocheilus oregonensis*) of south-central British Columbia.

## Methods

### Likelihood derivation

We assume that the basic data available for analysis consists of a matrix  $n_{l,a}$  with elements of numbers of fish sampled at (discrete) lengths  $l$  and ages  $a$  and that this is a random sample of the number of vulnerable fish of length  $l$  and age  $a$  available in the population to be sampled. In this case, the general form of the likelihood for  $n_{l,a}$  given the parameter vector  $\Theta$  is multinomial, and the log likelihood is the product of the data  $n_{l,a}$  and the log of the predicted proportions  $p_{l,a}$  of length  $l$  and age  $a$ :

$$(1) \quad \ln L_1(n|\Theta) = \sum_l \sum_a n_{l,a} \ln(p_{l,a})$$

How the  $p_{l,a}$  term of the likelihood is computed depends on which of the four likelihoods presented below is being used.

### Model 1: growth with no history of fishing

Where there has been no history of fishing, we calculate the predicted proportions in the matrix  $p_{l,a}$  as the proportion of vulnerable numbers of fish at age  $a$  and length  $l$  (in millimetres fork length) to the sum across all ages and lengths  $V_T$

$$(2) \quad p_{l,a} = V_{l,a}/V_T, \text{ where } V_T = \sum_l \sum_a V_{l,a}$$

We begin by assuming that  $V_{l,a}$  can be expressed as the product of three factors: a size-dependent vulnerability function  $v_l$  that depends on size (but not age), relative abundance  $N_a$  of age- $a$  fish, and the conditional probability  $P(l|a)$  of being in a discrete length interval  $l$  given age  $a$ :

$$(3) \quad V_{l,a} = v_l N_a P(l|a)$$

We assume that  $v_l$  has a functional form with increasing vulnerability-at-length:

$$(4) \quad v(l) = \frac{1}{(1 + e^{-\gamma(l-l_h)})}$$

We describe the length at 50% vulnerability ( $l_h$ ) as a fraction of  $L_\infty$  and a shape parameter ( $\gamma$ ) that describes the slope of the curve through  $l_h$  (Deriso et al. 1985). If we assume a stable recruitment in the population, then relative numbers-at-age ( $N_a$ ) can be calculated using

$$(5) \quad N_a = R e^{-M(a-1)}$$

Equation 5 computes the exponential decay of older cohorts relative to  $R$  individuals. The natural mortality rate  $M$  is estimated with the growth and vulnerability parameters using eq. 1. Where the assumption of a stable recruitment fails, it can be relaxed in the reduced likelihoods described below.

If we assume that variation in growth among individuals of age  $a$  is caused mainly by variation in individual asymptotic lengths  $L_{\infty,i}$ , then the von Bertalanffy prediction of size-at-age for any individual  $i$  is  $L_{\infty,i}f(a)$ , where  $f(a)$  is shared

by all individuals  $i$ . Using the von Bertalanffy growth model, this shared effect is described by

$$(6) \quad f(a) = 1 - e^{(-K(a-t_0))}$$

where  $K$  is the metabolic growth parameter and  $t_0$  is the theoretical time of zero length.

Assuming that individuals with  $L_{\infty,i}$  present at the start of cohort life are drawn from a normal distribution with mean  $L_{\infty}$  and variance  $\sigma_{\infty}^2$ , a subcohort of a specific  $L_{\infty,i}$  will follow a growth trajectory determined by the  $L_{\infty,i}$  that they were born with. This implies that the integral defining  $P(l|a)$  for each age  $a$  should be evaluated with mean  $l_a = L_{\infty}f(a)$  and variance in mean length-at-age  $\sigma_a^2$ :

$$(7) \quad P(l|a) = \frac{1}{\sigma_a \sqrt{2\pi}} \int_{l-d}^{l+d} \exp\left[-\frac{(l_a - l)^2}{2\sigma_a^2}\right] dl$$

where  $l$  is the length of the fish,  $d$  is half of the length interval width,  $l_a$  is the mean length-at-age, and  $\sigma_a^2$  is the variance in length-at-age. In most species of fish, the absolute value of  $\sigma_a^2$  increases with age  $\sigma_a^2$  (gets larger as  $a$  increases). To avoid estimating  $\sigma_a^2$  for each age, we assume that standard deviation at age  $\sigma_a$  can be expressed as a simple function of the mean length-at-age  $l_a$ . This function can be  $l_a$  multiplied by some coefficient of variation (CV), which is assumed to be constant across all ages or more complicated functions can be used. We chose to use eq. 8 derived by Fournier et al. (1991):

$$(8) \quad \sigma_a = \lambda_1 e^{\lambda_2 \left( -1 + 2 \frac{1 - \rho^{a-1}}{1 - \rho^{A-1}} \right)}$$

In this formulation,  $\rho$  is the Brody growth coefficient ( $\rho = e^{-K}$ ),  $A$  is the number of ages age  $a$ ,  $\lambda_1$  represents the magnitude of the standard deviations at age  $a$ , and  $\lambda_2$  determines the length-dependent trend in the standard deviations (if  $\lambda_2 = 0$ , the standard deviations are independent of length) (Fournier et al. 1991). This formulation has the disadvantage of adding two parameters over one if  $\sigma_a^2$  is expressed as  $CV \times l_a$ . The rationale for using it, however, is that it makes a less restrictive assumption about how  $\sigma_a$  changes as fish grow older. It also reduces the covariation in the parameters by making that standard deviation of length-at-age depend on one growth parameter ( $\rho = e^K$ ) rather than all three as it would if it were expressed as  $l_a \times CV$ .

### Model 2: reduced likelihood

An option for avoiding the assumption of stable recruitment is to construct the model using the conditional maximum likelihood estimates for numbers-at-age  $\hat{N}_a$  and using these in the predicted sample proportions-at-length and -age ( $p_{l,a}$ ) instead of the  $N_a$  predicted by eq. 5. This relaxes the assumption of stable age structure, which would be violated with highly variable recruitment. Differentiating the log likelihood with respect to  $N_a$ , setting this derivative to zero, and solving for  $N_a$  gives the conditional maximum likelihood estimates  $\hat{N}_a$ :

$$(9) \quad \hat{N}_a = \left( \frac{n_a}{n_T} \right) \left( \frac{V_T}{\bar{v}_a} \right)$$

where  $n_T$  is the total number aged in the sample,  $n_a$  is the total number in the sampled aged  $a$  years, and  $V_T$  the total vulnerable numbers from eq. 2 now computed without the numbers-at-age term ( $N_a$ ) and  $\bar{v}_a$ . The value of  $\bar{v}_a$  must first be calculated from the product of vulnerability-at-length and the conditional probability of being in length interval  $l$  given age  $a$ :

$$(10) \quad \bar{v}_a = \sum_l v_l P(l|a)$$

Note that this is not a true mean vulnerability-at-age. It represents the sum vulnerabilities-at-length in each age-class weighted by  $P(l|a)$ .

If the prediction of  $N_a$  from total mortality rate (eq. 5) is replaced in the likelihood by these conditional maximum likelihood  $\hat{N}_a$  estimates (allowing for any possible age structure in the population being sampled), the multiplicative term  $n_a/n_T$ , which does not vary with the parameters, can be dropped and  $p_{l,a}$  expressed as

$$(11) \quad p_{l,a} \propto \frac{v_l P(l|a)}{\bar{v}_a}$$

instead of eq. 2, so the reduced likelihood ( $L_2$ ) to be maximized without  $M$  becomes

$$(12) \quad \ln L_2 \propto \sum_l \sum_a n_{l,a} \ln \left( \frac{v_l P(l|a)}{\bar{v}_a} \right)$$

### Model 3: growth under harvesting

To model the effects of growth under harvesting, we again take the approach that  $n_{l,a}$  is a random sample from a population where  $V_{l,a}$  is the total number of vulnerable fish of length  $l$  and age  $a$ . The same general form of the likelihood function (eq. 1) applies but a historical fishing rate ( $F$ ) must now be incorporated into the estimated sample proportions  $p_{l,a}$ . To do this, we multiply each  $V_{l,a}$  element by a length- and age-specific survivorship  $\ell_{l,a}$  that accounts for a growth pattern that subjects individual fish to both fishing mortality and natural mortality.

Absent temporal information on past variation in recruitment, natural mortality rate  $M$ , and overall (fully vulnerable ages) fishing mortality rate  $F$ , we treat these factors as having been stable for enough years to result in a stable age-size distribution. When a population has been subjected to a fishery, the numbers at each age cannot be described by eq. 5 because each individual fish will be subject to a fishing mortality dependent on the growth trajectory of that individual. In this case, the  $N_a$  component of  $V_{l,a}$  in eq. 3 becomes  $N_{l,a}$ , which must be computed for each  $l$  and  $a$  combination. The  $N_{l,a}$  represents the survivors of fish that had individual  $L_{\infty,i}$  equal to  $l/f(a)$ , i.e., followed a growth pattern that subjected them to cumulative instantaneous mortality (now including the fishing mortality  $F$ ) at length and age:

$$(13) \quad Z_{l,a} = \sum_{a'} [M + v_{l(a')} F]$$

to give

$$(14) \quad N_{l,a} = R e^{-Z_{l,a}}$$

where the sum (or integral) over  $a'$  represents mortality over ages up to age  $a$  and where  $v_{l(a')}$  represents the time sequence of  $v_l$  vulnerabilities seen by fish that followed the growth trajectory  $l_{a'} = L_{\infty,i}f(a') = [1/f(a)]f(a')$  over ages  $a'$ . The initial recruitment can be set to  $R = 1$  for convenience, since only the ratios of the  $V_{l,a}$  to  $V_T$  appear in the likelihood function (the composition information  $n_{l,a}$  carries no direct information on total population size). For typical vulnerability functions  $v(l)$  describing  $v_l$ , there is no simple analytical solution for the sum or integral of  $v_{l(a')}$  values over ages  $a'$ , but these sums can be easily evaluated numerically given any assumed form for  $v(l)$ . This integral needs to be evaluated for all  $l, a$  combinations, since each such combination is assumed to have had a different  $L_{\infty,i}$  and hence a different exploitation history. With a history of exploitation, the vulnerable numbers at age and length  $V_{l,a}$  are calculated as

$$(15) \quad V_{l,a} = v_l R e^{-Z_{l,a}} P(l|a)$$

The  $p_{l,a}$  in eq. 1 is now calculated with  $V_{l,a}$  calculated using eq. 15:

$$(16) \quad p_{l,a} = \frac{V_{l,a}}{V_T} = \frac{v_l R e^{-Z_{l,a}} P(l|a)}{V_T}$$

so that the full likelihood ( $L_3$ ) that includes size-selective fishing mortality is

$$(17) \quad \ln L_3(n|\Phi) \propto \sum_l \sum_a n_{l,a} \ln \left( \frac{v_l R e^{-Z_{l,a}} P(l|a)}{V_T} \right)$$

The only numerically complex part of this growth under fishing formulation is computing the length-age-specific survivorship ( $e^{-Z_{l,a}}$ ) for each  $l, a$  element. Consider, for example, the function  $v(l)$  in eq. 4. Representing  $l(a')$  as  $l(a') = L_{\infty,i}f(a')$ , this age function can then be substituted for lengths in the vulnerability function for all length intervals  $l$ . For example, the substitution for the vulnerability function in eq. 4 is

$$(18) \quad v(a') = \frac{1}{(1 + e^{-\gamma(f(a') - l_h/L_{\infty,i})})}$$

which then has to be integrated over  $a'$  and then multiplied by  $F$  to predict cumulative fishing mortality effects on the "subcohort" of fish that started life with asymptotic size  $L_{\infty,i}$ . For alternative vulnerability functions that are difficult to integrate (i.e., dome-shaped curves), we recommend simply summing the function over ages  $a'$  from 1 to  $a$  using  $a'$  age steps of 1 year for long-lived fish and shorter steps (e.g., 0.5 year) for fish that live only a few years.

#### Model 4: reduced likelihood with fishing

To compute the reduced likelihood under fishing for the growth under harvest scenario we need to include the individual the length-age specific survivorship into the calculation of eq. 10 to give:

$$(19) \quad \bar{v}_a = \sum_l v_l R e^{-Z_{l,a}} P(l|a)$$

**Table 1.** Parameters used to generate fake data.

Parameter	Value
$N_T$	1000
$\ln L_{\infty}$	5
$K$	0.15
$t_0$	-0.25
$Z$	0.2
$l_h$	0.6
$\gamma$	0.1
$\lambda_1$	16
$\lambda_2$	0.5
$CV_R$	0.5

Once this is done,  $\bar{v}_a$  gets updated automatically with changes with  $F$ . Note that  $M$  still appears in eq. 13 but does nothing to change the likelihood, so is not estimated. The  $p_{l,a}$  matrix now is computed as

$$(20) \quad p_{l,a} \propto \frac{v_l R e^{-Z_{l,a}} P(l|a)}{\bar{v}_a}$$

and the reduced likelihood now including fishing becomes

$$(21) \quad \ln L_4 \propto \sum_l \sum_a n_{l,a} \ln \left( \frac{v_l R e^{-Z_{l,a}} P(l|a)}{\bar{v}_a} \right)$$

#### Simulations

To test the derivations above, we simulated data using known parameters (listed in Table 1). Variation in past recruitment anomalies was included as lognormal process error with a mean of 0 and standard deviation of 1 multiplied by a coefficient of variation,  $CV_R$ .

We compared the performance of the standard Fabens (after Fabens (1965)) method that assumes a representative sample of numbers-at-age and minimizes the sum of squares difference between predicted and observed lengths-at-age to the four models that we derive above using simulated data. We generated 100 data sets with a no fishing ( $F = 0$ ), with a sustainable fishery ( $F = K$ ), and an overfished population ( $F = 1$ ). This allows comparison of the estimated parameter values with those used to simulate the data.

Fisheries where the fishing mortality has been stable for a long time are uncommon. Using these same parameters (listed in Table 1) and a fully length-age-structured simulation model, we tested models 3 and 4 with simulated data from stocks where  $F$  increased in increments of 0.1/year for 5 years and where  $F$  was increased in increments of 0.1/year for 10 years. To do so, new recruits each year were assumed to have a fully representative distribution of individual asymptotic lengths (with mean  $L_{\infty,i}$  and variance  $\sigma_{\infty}$  computed with eq. 8 evaluated at an infinite age). Individuals of specific asymptotic lengths ( $L_{\infty,i}$ ) across all ages were then exposed to fishing mortality every year according to the size (and hence  $v_l$ ) that they were in that year. As in the simulations with stable fishing mortality, natural mortality is assumed constant and  $CV_R$  set to 0.5.



The sensitivities of all of the models to changes in parameters used to simulate data were tested by using extreme values. We conducted simulations with the length at half vulnerability  $L_h$  set to  $0.05 \times L_\infty$  and  $0.9 \times L_\infty$ , knife-edged vulnerability ( $\gamma = 0.9$ ), and with low variability in length-at-age ( $\lambda_1$  set to 5). Finally, we tested the models with a short-lived, fast-growing population with  $K = 0.7$  and  $M = 0.7$ .

### Growth parameter estimation for northern pikeminnow

We collected data for northern pikeminnow from Moose Pasture Lake located on the Bonaparte Plateau approximately 100 km north of Kamloops, British Columbia. We sampled during a depletion experiment in which fish were collected for aging. Fish were captured during six bouts of fishing using four hoop nets. Three of these nets consisted of 6 m  $\times$  1 m fiberglass hoops covered in 1-cm mesh with a 20-m center lead and 15-m side leads. The other consisted of 6 m  $\times$  0.7 m steel hoops covered in 0.5-cm mesh with a 15-m center lead and 10-m side leads. Lapilli otoliths were taken from each fish and 945 were aged. We cut larger otoliths along the ventral/dorsal axis using an Isomet Bueller slow-speed saw and burned and counted annuli. Otoliths too small to cut were burned and counted directly. For fish older than 5 years, failing to cut otoliths along the long axis resulted in significant underestimates of age.

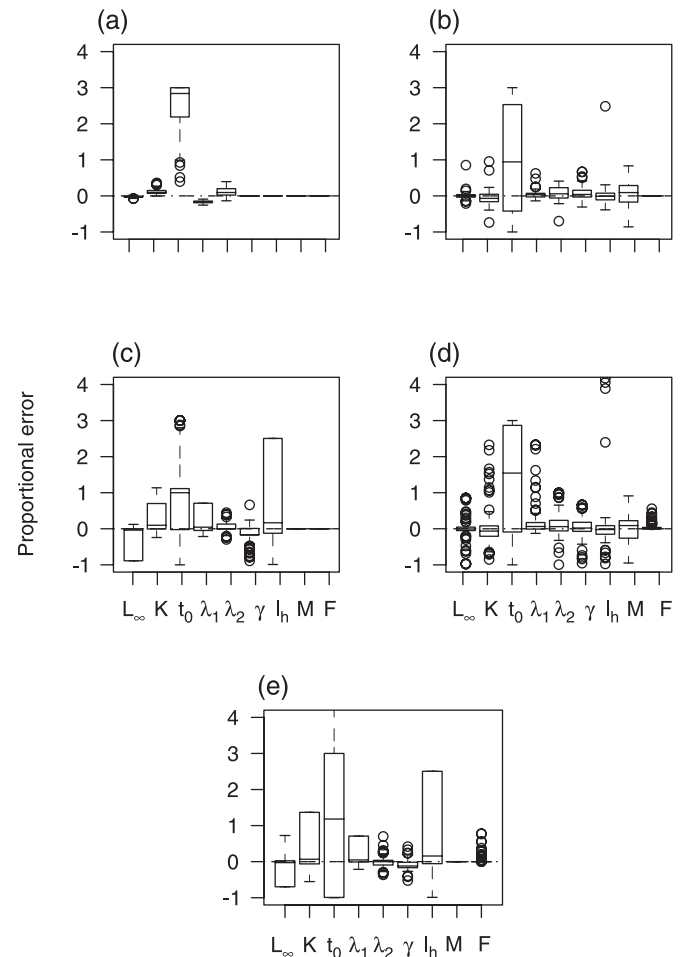
During the 3 months preceding the depletion (9–15 September 2002), 1004 fish were tagged using 5- and 15-mm Floy numbered tags during four bouts of mark-recapture conducted at approximately 2-week intervals. Only those fish that were released in perfect condition are included in this analysis. Because the functional form of the vulnerability-at-length is unknown for fyke nets and northern pikeminnow, we used these tagging data to directly estimate vulnerabilities-at-length. We divided the fish into 5-mm size bins and estimated the gear selectivity for each bin using the likelihood described by Myers and Hoenig (1997). We include this analysis both to ensure that the functional form of the vulnerability-at-length has been specified correctly and to compare how well the vulnerability function estimated using only the length-at-age data and how the multinomial likelihood(s) performs relative to the direct estimates of vulnerability computed from the tagging experiment.

## Results

With no fishing mortality  $F$ ,  $K$  was overestimated and  $L_\infty$  underestimated using the Fabens method (Fig. 1). These biases were small (0.20 for  $K$  and negligible for  $L_\infty$ ). There was still the common bias of a small  $t_0$  parameter, which was underestimated by a factor of 2. With no fishing, models 1, 2, 3, and 4 were unbiased for the growth and vulnerability parameters but slightly overestimated natural mortality  $M$  and  $F$ .

The bias of Fabens and models 1 and 2 increased with increasing  $F$ . With the Fabens method, the bias was strong for  $t_0$ , small for  $K$ , and negligible for  $L_\infty$  (Fig. 1). Models 1 and 2 still performed reasonably well. Model 1 accounted for  $F$  by overestimating  $M$  (Fig. 2) but could not account for the selective mortality on higher  $L_{\infty,i}$  individuals caused by fishing and so underestimated  $L_\infty$ . Although model 2 suffered

**Fig. 1.** Proportional error in parameter estimates for each likelihood formulation with  $F = 0$ . (a) Fabens; (b) model 1; (c) model 2; (d) model 3; (e) model 4. Circles represent outliers.



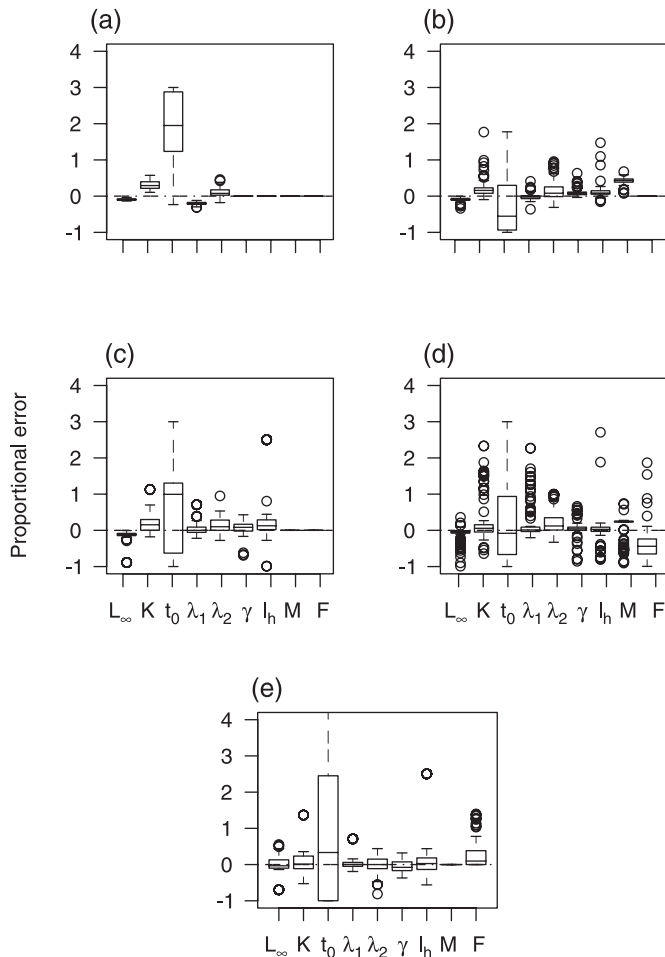
from the same bias in growth parameter estimates as model 1 did, models 3 and 4 were unbiased.

Biases in parameter estimates were much worse where  $F = 1.0$  for all but models 3 and 4 (Fig. 3). The bias in  $K$ ,  $L_\infty$ , and  $t_0$  was less severe for models 1 and 2 but in the same direction as Fabens. When  $F$  was low, the distortion in the sample caused by  $F$  was compensated for by high estimates of  $M$  (Fig. 2). This was not the case with high  $F$  (Fig. 3). The  $F$  distorted the sample so much that the few older fish left were the small  $L_{\infty,i}$  (slow-growing) individuals.

Models 3 and 4 correctly estimated the von Bertalanffy growth parameters (Fig. 1) at all levels of stable  $F$ . Even with moderate recruitment anomalies used to simulate the data ( $CV_R = 0.5$ ),  $F$  was overestimated. This was the case at low  $F$  (although it cannot be seen because of the scaling of Figs. 1 and 2) and much worse at higher  $F$  (Fig. 3).

The performance of models 3 and 4 suffered with variable fishing. All of the parameters were estimated well with 5 years of increased fishing (Figs. 4a, 4c, and 4e). With 10 years of increased fishing, model 5 (Fig. 4f) performed as poorly as Fabens (Fig. 4b) and model 4 slightly better (Fig. 4d).

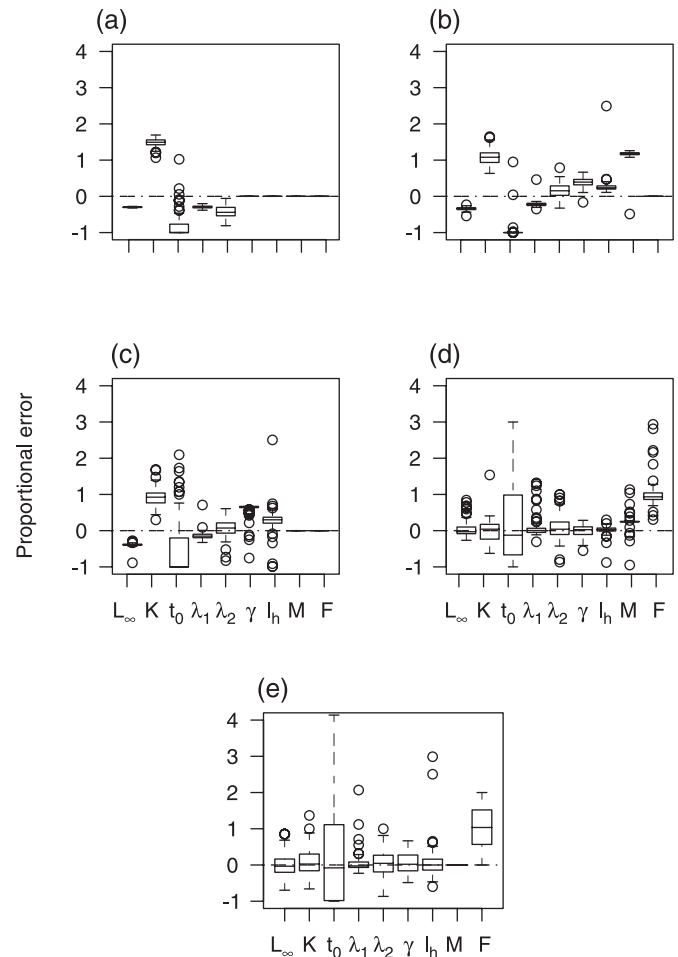
**Fig. 2.** Proportional error in parameter estimates for each likelihood formulation with  $F = 0$ . (a) Fabens; (b) model 1; (c) model 2; (d) model 3; (e) model 4. Circles represent outliers.



At all levels of fishing, models 2 and 4 were less precise than models 1 and 3. Models 2 and 4 do not assume a stable age distribution, and so, more parameter combinations can explain the data. The reduced likelihoods effectively estimated recruitment of each age- $a$  cohort  $R_t - a$ . This means that  $a$  additional parameters are estimated. While there is loss in precision of the parameter estimates, it is a much more honest acknowledgement of the uncertainty of them in cases where a stable age structure is doubtful or unknown.

All of the models are sensitive to low values of  $l_h$ . A very slowly sloping vulnerability ( $\gamma = 0.05$ ) function had the same effect. Distortion in the data of younger age-classes relative to fully represented ones needs to be present in the data to resolve the vulnerability parameters. For example, with no fishing, using the same growth parameters listed in Table 1 but with an  $l_h$  equal to  $0.05 \times L_\infty$ , the data do not contain any information about the vulnerability parameters with mean biases of 3 for  $l_h$  and 5.1 for  $\gamma$ . The  $M$  is confounded with the vulnerability parameters and is underestimated by 0.28. Fortunately, the growth parameters  $K$ ,  $L_\infty$ , and  $t_0$  are well estimated with a mean bias of less than 0.05. As expected, with little size selectivity, the Fabens method actually performs as well as the likelihoods presented here

**Fig. 3.** Proportional error in parameter estimates for each likelihood formulation with  $F = 0$ . (a) Fabens; (b) model 1; (c) model 2; (d) model 3; (e) model 4. Circles represent outliers.



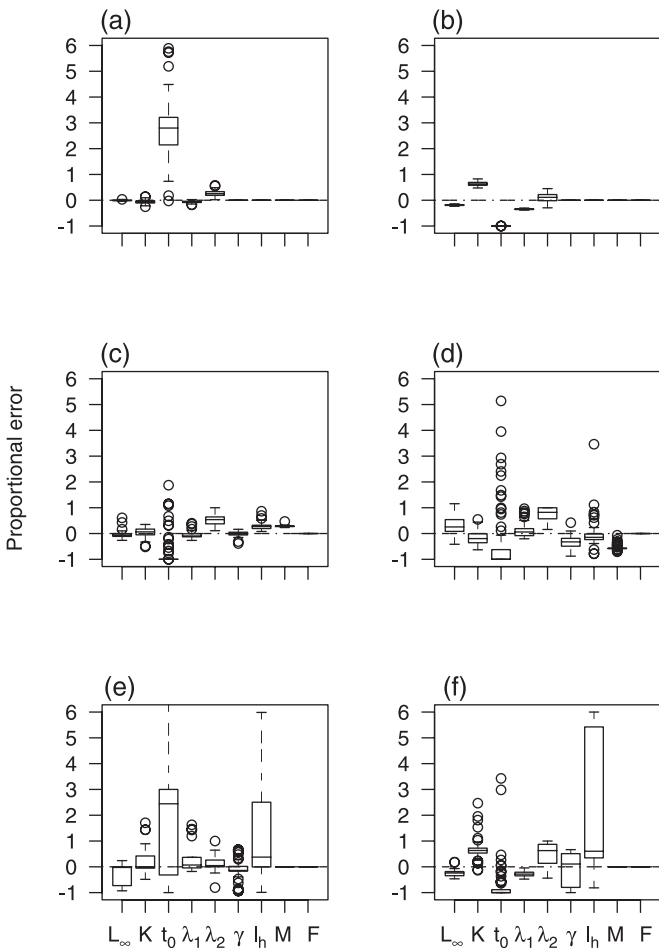
because there is not significant distortion in length-age samples from either fishing or gear selectivity.

With no fishing and parameters again set to those listed in Table 1 but making  $l_h = 0.9$ , the problem is that there are no fully represented age-classes in the length-age sample. The parameters  $L_\infty$ ,  $K$ ,  $\lambda_1$ ,  $\lambda_2$ , and  $l_h$  are all well estimated with biases less than 0.08, but  $t_0$ ,  $\gamma$ , and  $M$  are biased by 0.35, 0.21, and -0.40, respectively. As long as  $l_h$  is an intermediate value, then the model is robust to steeply sloping  $v_l$  ( $\gamma = 0.9$ ).

The models were robust to data simulated with faster growing (moderately high  $K = 0.5$ ) and faster dying (higher  $M = 0.6$ ) populations, but in populations where  $K$  and  $M$  are greater than about 0.7, fractional (half or quarter year) ages need to be used or the model performs badly and either underestimates  $K$  or fails to converge at all.

These models break down when the shape of the vulnerability curve is not specified correctly, if  $F$  is too large, or if  $M$  is size or age dependent and with smaller length-at-age samples. If the actual vulnerability function is dome shaped but an asymptotic function is assumed in the analysis,  $F$  is confounded with decreasing vulnerability of older/larger fish and becomes biased upward. If  $F$  is very high, it is very de-

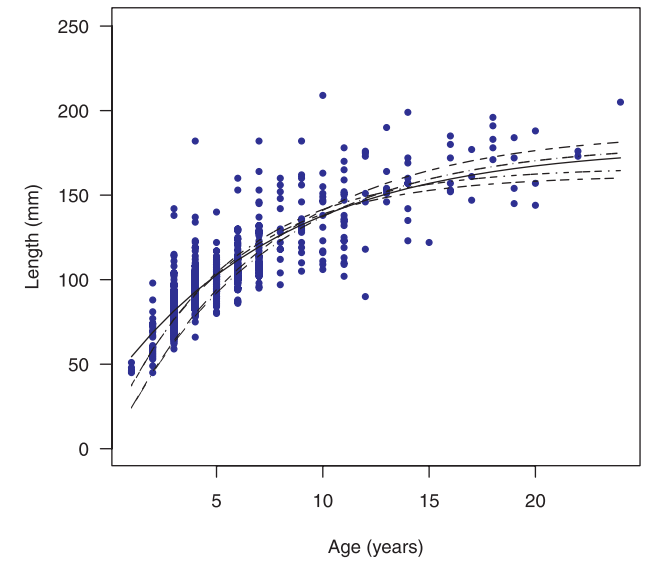
**Fig. 4.** Proportional error of (a and b) Fabens, (c and d), model 3, and (e and f) model 4, where  $F$  increases in increments of 0.1 to (a, c, e)  $F = 0.5$  and (b, d, f)  $F = 1.0$ ;  $F$  was estimated but the proportional bias is not included here for clarity. Circles represent outliers.



structive to the age structure of the population and no parameters can be estimated well. Where  $M$  is size or age dependent (where  $M$  decreases over age, for example), there is severe confounding of  $M$  and  $F$  with the parameters that describe  $v_l$ . At low sample sizes, it is possible to have no data by random chance from either the faster growing younger fish (that help resolve  $v_l$ ) or older fish (that resolve  $M$  and  $F$ ). With 25 age-classes, these models worked very poorly with fewer than approximately 500 samples.

Choosing between reduced or full likelihoods can be done empirically using various information criteria (Bayesian, Akaike, etc.). When the Akaike information criterion was used on simulated data, the reduced likelihood was almost always favored because even with the added number  $a$  estimated parameters, the value of the objective function at the minimum log likelihood was generally about three orders of magnitude smaller, thus overpowering any penalties for additional parameters. However, when the stock assessment scientist has some prior knowledge of past recruitment anomalies and mortality from the stock assessments, these

**Fig. 5.** Growth curves fit to northern pikeminnow data using all five models. Solid, Fabens; short-dashed, full likelihood; dotted, reduced likelihood; dot-dashed, full likelihood with fishing; long-dashed, reduced likelihood with fishing.



**Table 2.** Estimated parameter values for each model.

Parameter	Model				
	1	2	3	4	5
$L_{\infty}$	179.61	162.06	179.77	166.46	188.10
$K$	0.12	0.19	0.14	0.18	0.14
$t_0$	-1.95	-0.41	-0.00	-0.38	-0.00
$\lambda_1$	13.90	13.18	16.16	13.20	15.96
$\lambda_2$	0.69	0.65	0.60	0.61	0.57
$l_h$	—	0.45	0.49	0.44	0.47
$\gamma$	—	0.21	0.20	0.19	0.19
$Z$	—	0.38	—	0.26	—
$F$	—	—	—	0.18	0.15

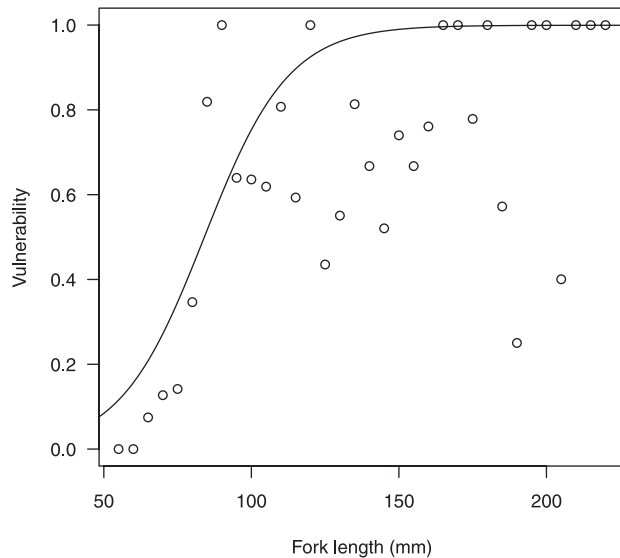
data could simply be included as terms in the likelihood calculation.

**Northern pikeminnow growth**

Each of the five model fits to the northern pikeminnow data is plotted in Fig. 5. Parameter estimates are very similar for all of the models used. The likelihood formulations that include fishing (models 3 and 4) actually estimate a small  $F$  in this population. Models 4 and 5 have a tendency to overestimate  $F$ , but a nonzero value for this parameter is not as unusual as it might seem. The population is exposed to a small-scale fishery from local lodge owners who have in the past conducted annual trapping programs during the spring spawn.

Parameter estimates from all four models are very similar (Table 2.). The most reasonable ones to choose, however, are those fit using model 5. We have no prior knowledge of the  $M$  of the population or of the stable age structure in the past.

**Fig. 6.** Plot of vulnerability curve (line) for northern pikeminnow estimated using model 3 and direct estimates of vulnerability using mark-recapture.



The vulnerability curve estimated by model 5 (Fig. 6) is very similar to that fit using the mark-recapture data. Fitting with model 5 appears to overestimate the vulnerability-at-length for smaller size-classes (Fig. 6) but otherwise seems to be a reasonable approximation. Note that the vulnerability-at-length for the higher size-classes is not well defined because of the small sample sizes of tagged animals in those size bins.

## Discussion

The estimated  $L_{\infty}$  of northern pikeminnow is tiny compared with other documented populations. While the von Bertalanffy growth parameter  $K$  for this population of northern pikeminnow is similar to those published in the literature, the  $L_{\infty}$  is much smaller. For this reason, we have dubbed this population “the pygmy pikeminnow”. Northern pikeminnow have been studied extensively owing to their predation on juvenile salmonids in a variety of systems such as Cultus Lake, British Columbia (Ricker 1941; Steigenberger and Larkin 1974), and they are presently managed in the Columbia basin to reduce their impact on juvenile salmon (Friesen and Ward 1999). Typical  $L_{\infty}$  values for this species are on the order of 400–600 mm total length (Parker et al. 1995) in the Columbia River basin, and in Idaho lakes, the average reported size of a sexually mature female is approximately 440 mm total length (20 in.) for females and 300 mm (15 in.) for males (Jeppson 1959). These values are well above even the maximum sizes that we observe in our study area. The reasons for their small size are presently unknown, but we hypothesize that it is mainly due to the short growing season (3 months or so) that these fish experience.

The simulations shown here are not unusual in showing the biases owing to gear selectivity and size-selective mortality in the estimation of growth parameters. These biases are well known and estimation procedures to deal with them have been developed (Deriso et al. 1985; Parma and Deriso

1990; Sinclair et al. 2002b). The methods that we present are original in using simulation studies to show these biases in estimates of the von Bertalanffy parameters and those describing the size selectivity and  $M$  and  $F$ . Unlike Parma and Deriso (1990) and Sinclair et al. (2002b), we make no effort to include environmental effects or, in the case of Sinclair et al. (2002b), density-dependent growth on size-at-age. Parma and Deriso (1990) showed that increasing the contribution of environmental factors had little effect on the variance in size-at-age relative to size-selective harvesting and that most fisheries data contain little information about such effects. Sinclair et al. (2002b) also determined that by far the largest effect on mean predicted  $L_{\infty}$  was size-selective mortality.

Using our method, we have shown it to be possible under a restricted set of assumptions to estimate the von Bertalanffy growth parameters,  $M$ , and  $F$  and the vulnerability parameters from only size-at-age data when other data from a fishery are not available. We caution that estimates of the vulnerability and mortality parameters should be considered very uncertain. Unfortunately, the sensitivity of the models to stable  $F$  and known vulnerability shape restricts the applicability of these models. Recall that the models were robust to an increasing fishing rate as long as  $F$  had not grown too rapidly, while the Fabens method performed almost equally as well in that case but did not tell the researcher anything about  $M$  or the vulnerability function  $v_l$ . Of particular concern is the breakdown of the models following a very rapid increase in  $F$ . Having accurate assessment of the true growth parameters of the stock is particularly needed to assess the rebuilding potential of the stock under rebuilding.

Fortunately, in many cases, the reason for estimating growth parameters in the first place is that it is part of a larger stock assessment where  $F$  (and sometimes  $v_l$  as in virtual population analysis) is already estimated. This method could be applied in more complex stock assessments where  $n_{l,a}$  tables can be provided for a collection of sample years. For each such table, the stock assessment model provides additional information on time-varying recruitments  $R_{t-a}$  and  $F$  and in some cases (like virtual population analysis)  $v_l$ . Assuming that the  $n_{l,a}$  data are collected independently each year, the log-likelihood term for each of these tables can simply be added to the overall log likelihood for the assessment model. The accounting (sum or integration) calculations for each  $Z_{l,a}$  (cumulative mortality by size and age) would require calculating the survival of those individuals growing along each  $L_{\infty,i}$  trajectory but may carry valuable information on changes over time in  $F$ . A potentially important advantage of this approach would be to correct the problem that has plagued some past length-based stock assessment methods of ignoring cumulative effects of fishing on length distribution patterns. Mean length and variability of length-at-age are explicitly variable and dependent on  $F_t$  in our method owing to variation in  $Z_{l,a}$ .

## Acknowledgements

This manuscript was greatly improved by thorough reviews by an anonymous reviewer and Ana Parma. We greatly appreciate the help of an excellent field crew in the collection of otoliths consisting of Shaun McGloughlin, Hannah



Murphy, Lew Coggins, Mike Melnychuk, and Eric Parkinson. We could not have gotten started without the expert advice from the Department of Fisheries and Oceans aging laboratory at the Pacific Biological Station in Nanaimo, British Columbia, or finished without Angela Danyluk who did the bulk of the aging. Financial support was provided by a Natural Sciences and Engineering Research Council of Canada operating grant and Pew Fellowship to C.J. Walters and by a Habitat Conservation Trust Fund grant to Eric Parkinson.

## References

- Burr, P., and Doksum, K. 1988. Inferring the distribution of the parameters of the von Bertalanffy growth model from length moments. *Can. J. Fish. Aquat. Sci.* **45**: 1179–1788.
- Deriso, R.B., Quinn, T.J., II, and Neal, P.R. 1985. Catch-age analysis with auxiliary information. *Can. J. Fish. Aquat. Sci.* **42**: 815–824.
- Fabens, A. 1965. Properties and fitting of the von Bertalanffy curve. *Growth*, **29**: 265–289.
- Fournier, D.A., Sibert, J.R., and Terceiro, M. 1991. Analysis of length frequency samples with relative abundance data for the Gulf of Maine northern shrimp (*Pandalus borealis*) by the MULTIFAN method. *Can. J. Fish. Aquat. Sci.* **48**: 591–598.
- Friesen, T.A., and Ward, D.L. 1999. Management of northern pikeminnow and implications for juvenile salmonid survival in the Lower Columbia and Snake Rivers. *N. Am. J. Fish. Manag.* **19**: 406–420.
- Hansen, J., and Chouinard, G. 1992. Evidence that size-selective mortality affects growth of Atlantic cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. *J. Fish Biol.* **41**: 31–41.
- James, I.R. 1991. Estimation of von Bertalanffy growth parameters from recapture data. *Biometrics*, **47**: 1519–1530.
- Jeppson, P.W. 1959. Ecology and control of the Columbia pikeminnow in northern Idaho lakes. *Trans. Am. Fish. Soc.* **88**: 197–202.
- Kristiansen, T., and Svasand, T. 1998. Effects of size-selective mortality on growth of coastal cod illustrated by tagging data and an individual-based growth and mortality model. *J. Fish Biol.* **52**: 688–705.
- Laslett, G.M., Eveson, J.P., and Polachek, T. 2002. A flexible maximum likelihood approach for fitting growth curves to tag-recapture data. *Can. J. Fish. Aquat. Sci.* **59**: 976–986.
- Lee, R.M. 1912. An investigation into the methods of growth determination in fishes. *Cons. Perm. Int. Explor. Mer*, **63**: 35.
- Myers, R.A., and Hoenig, J.M. 1997. Direct estimates of gear selectivity from multiple tagging experiments. *Can. J. Fish. Aquat. Sci.* **54**: 1–9.
- Palmer, M.J., Phillips, B.F., and Smith, G.T. 1991. Application of nonlinear models with random coefficients to growth data. *Biometrics*, **47**: 623–635.
- Parker, R.M., Zimmerman, M.P., and Ward, D.L. 1995. Variability in biological characteristics of northern pikeminnow in the Lower Columbia and Snake rivers. *Trans. Am. Fish. Soc.* **124**: 335–346.
- Parma, A.M., and Deriso, R.B. 1990. Dynamics of age and size composition in a population subject to size-selective mortality: effects of phenotypic variability in growth. *Can. J. Fish. Aquat. Sci.* **47**: 274–289.
- Ricker, W.E. 1941. The consumption of young sockeye salmon by predaceous salmon by predaceous fish. *J. Fish. Res. Board Can.* **5**: 293–313.
- Ricker, W.E. 1969. Effects of size-selective mortality and sampling bias on estimates of growth, mortality, production and yield. *J. Fish. Res. Board. Can.* **26**: 479–541.
- Sainsbury, K. 1980. Effect of individual variability on the von Bertalanffy growth equation. *Can. J. Fish. Aquat. Sci.* **37**: 241–247.
- Sinclair, A.F., Swain, D.P., and Hanson, J.M. 2002a. Disentangling the effects of size selective mortality, density and temperature on length at age. *Can. J. Fish. Aquat. Sci.* **59**: 372–382.
- Sinclair, A.F., Swain, D.P., and Hanson, J.M. 2002b. Measuring changes in the direction and magnitude of size-selective mortality in a commercial fishery. *Can. J. Fish. Aquat. Sci.* **59**: 361–371.
- Steigenberger, L.W., and Larkin, P.A. 1974. Feeding activity rates of digestion of northern pikeminnow (*Ptychocheilus oregonensis*). *J. Fish. Res. Board Can.* **31**: 411–420.
- Wang, Y.-G., Thomas, M.R., and Somers, I.F. 1995. A maximum likelihood approach for estimating growth from tag-recapture. *Can. J. Fish. Aquat. Sci.* **52**: 252–295.