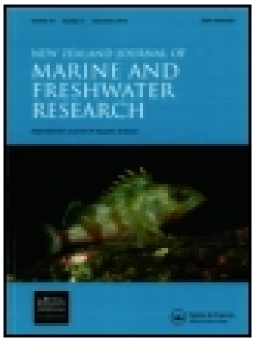

New Zealand Journal of Marine and Freshwater Research

ISSN: 0028-8330 (Print) 1175-8805 (Online) Journal homepage: http://www.tandfonline.com/loi/tnzm20

Maximum likelihood estimation of growth and growth variability from tagging data

R. I. C. C. Francis

New Zealand Journal of Marine and Freshwater Research, 1988, Vol. 22: 42–51
0028-8330/88/2201-0042$2.50/0

Maximum likelihood estimation of growth and growth variability from tagging data

R. I. C. C. FRANCIS
Fisheries Research Centre
Ministry of Agriculture and Fisheries
P.O. Box 297, Wellington, New Zealand

Abstract A maximum likelihood approach is described for the analysis of growth increment data derived from tagging experiments. As well as describing mean growth this approach allows the separate estimation of measurement error and growth variability, and uses mixture theory to provide an objective way of dealing with outliers. The method is illustrated using data for Pacific bonito *(Sarda chiliensis)* and the growth variability model is compared to other published models. The difference between growth curves derived from tagging and age–length data is emphasised and new parameters are given for the von Bertalanffy curve that have better statistical properties, and represent better the growth information in tagging data, than do the conventional parameters.

Keywords growth; growth variability; tagging; maximum likelihood; mixture theory; *Sarda chiliensis;* von Bertalanffy curve

Received 16 March 1987; accepted 6 May 1987
Fisheries Research Publication 577

INTRODUCTION

Growth increment data from tagging experiments have long been used to describe the average growth of fish stocks. Initially this was done using graphical methods (e.g., Walford 1946). As computers became readily accessible the method of Fabens (1965) was widely used to fit the von Bertalanffy curve. More recently, with the advent of efficient general purpose maximisation routines, the least squares/maximum likelihood method has come into use to fit a range of different growth curves (e.g., Shepherd & Hearn 1983; Kirkwood & Somers 1984). Kimura (1980) gave an excellent description of this technique and showed how it leads easily to tests of significance and the construction of confidence regions. Though his paper describes the fitting of the von Bertalanffy curve to age–length data the principles are easily transferred for use with tagging data and other growth curves.

This paper addresses three problems which are commonly encountered in the estimation of growth from tagging data.

Firstly, where the von Bertalanffy curve is used the standard parameters ℓ_∞ and k are unsatisfactory. Knight (1968) pointed out the inadequacy of ℓ_∞ as a growth descriptor when age–length data are not extensive enough to demonstrate asymptotic growth; for tagging data ℓ_∞ is not appropriate even if individual growth is asymptotic (see Discussion and Sainsbury 1980: fig. 3). Furthermore there is a high correlation between estimates of ℓ_∞ and k.

Secondly, the standard approach makes no attempt to model (or estimate) variability in individual growth rates. This is surprising in that information on the extent of variability is often clearly apparent in graphs that show an increase in the scatter of residuals (differences between observed and predicted growth increments) with increasing time at liberty (e.g., Sanders & Powell 1979: fig. 3). An obvious interpretation is that, for short times at liberty, the residuals are caused mostly by errors in the measurement of length at tagging and recapture, and as the time at liberty increases so does the contribution from growth variability.

The third problem is that of outliers, which are normally dealt with somewhat arbitrarily by removal of data.

The aims of this paper are, in the order in which they are addressed: (a) to describe new parameters for the von Bertalanffy curve which are more appropriate for use with tagging data; (b) to describe a simple model of growth variability and compare it to other models; (c) to present a more objective approach to data outliers; (d) to combine these three features in a maximum likelihood model for growth; and (e) to illustrate the use of this model on published tagging data.

NOTATION

To aid the reader the convention will be followed that tagging data and model parameters are represented by upper and lower case roman letters respectively, and greek letters are used for other variables.

The tagging data will be represented by T_1, T_2, L_1, and L_2 where T stands for time (i.e., date) and L for length, and the subscripts 1 and 2 refer to the moments of tagging and recapture, respectively. Increments in length and time will be written as ΔL and ΔT, respectively. The standard von Bertalanffy parameters will be written as ℓ_∞ and k.

VON BERTALANFFY PARAMETERS

The most widely used length growth curve in the fisheries literature is that of von Bertalanffy (Ricker 1975). In the context of tagging data this may be written as

$$\Delta L = (\ell_\infty - L_1)(1 - e^{-k\Delta T}) \qquad (1),$$

and interpreted as describing the growth expected from a fish of length L_1 over the time period ΔT.

The parameters advocated here for use with tagging data are g_α and g_β — the mean annual growth rates of fish of arbitrary lengths α and β — defined from (1) by substituting g_α (or g_β) for ΔL, α (or β) for L_1, and 1 for ΔT. This leads to

$$\ell_\infty = (\beta g_\alpha - \alpha g_\beta)\ (g_\alpha - g_\beta) \text{ and}$$
$$e^{-k} = 1 + (g_\alpha - g_\beta)\ /(\alpha - \beta)$$

so that (1) may be rewritten as

$$\Delta L = \left\{\frac{\beta g_\alpha - \alpha g_\beta}{g_\alpha - g_\beta} - L_1\right\} \left\{1 - \left[1 + \frac{g_\alpha - g_\beta}{\alpha - \beta}\right]^{\Delta T}\right\} \qquad (2).$$

The lengths α and β should be chosen so that g_α and g_β are well determined by the tagging data. This makes g_α and g_β directly descriptive of the growth information contained in the tagging data. As will be seen in the example below, it may also result in a lower correlation between the von Bertalanffy parameters than is normally found with ℓ_∞ and k. The time period of 1 year associated with g_α and g_β is arbitrary and could be replaced by any other period suggested by the data. The mean annual growth rate at any other length, γ, may be calculated from

$$g_\gamma = \big((\gamma - \alpha)\, g_\beta + (\beta - \gamma)\, g_\alpha\big)\ /(\beta - \alpha) \qquad (3).$$

In the example below a seasonal form of the von Bertalanffy curve is used in which ΔT in Equation 2 is replaced by $\Delta T + (\phi_2 - \phi_1)$ where

$$\phi_i = u\big(\sin(2\pi\,[T_i - w])\big)/(2\pi), \text{ for } i = 1, 2 \qquad (4).$$

This is a slightly different form of a curve proposed by Pitcher & MacDonald (1973). The parameter w describes the time of year at which growth rates are maximum and u ($\geqslant 0$) describes the extent of seasonality. If $u = 0$ there is no seasonal variation in growth, otherwise maximum and minimum instantaneous growth rates are in the ratio $1 + u$ to $1 - u$.

DESCRIBING GROWTH VARIABILITY

We will assume that the growth of a fish of length L_1 over time ΔT is normally distributed with mean μ and standard deviation σ. Growth variability is thus defined by σ and this must, since the model is built for tagging data, be a function of L_1 and ΔT alone (i.e., there can be no dependence on age). In this paper we will make σ a function of μ (which is itself a function of L_1 and ΔT). It seems reasonable that this function be increasing (the greater the expected growth, the greater the scope for variation in growth) so the simplest form for σ uses one parameter, ν, and is given by

$$\sigma = \nu\mu \qquad (5).$$

If, when this model is fitted, a plot of standardised residuals against expected growth shows it to be inadequate, then a second parameter, t, may be introduced and more complex relationships such as

$$\sigma = \min(\nu\mu,\ t) \qquad (6),$$

$$\sigma = t\,(1 - e^{-\nu\mu}) \qquad (7),$$

or

$$\sigma = \nu\mu^t \qquad (8)$$

may be tested.

OTHER MODELS OF GROWTH VARIABILITY

Cohen & Fishman (1980) devised a model in which the variability of growth increment is a function of ΔT alone (see their equations 2.3, 2.4). Their model may work reasonably well where growth is almost linear but where mean growth rate changes markedly with age it is counter-intuitive. Suppose, for example, the mean annual growth of one-year-old fish is 10 cm but older fish of the same species grow only 1 cm per year. It seems unrealistic to apply the same variance to these very different means and yet that is what Cohen & Fishman's model would do.

A very different approach was followed by Sainsbury (1980) who considered the possibility that all fish grow according to a von Bertalanffy curve but that each fish has its own parameters ℓ_∞ and k. By making certain distributional assumptions he was able to give formulae for the expectation and variance of ΔL (his equations 9 and 10). The way in which the variance in ΔL depends on L_1 and ΔT in this model depends on the parameters $\bar{\ell}_\infty$, $\bar{k}$, $\sigma_{\ell\infty}$, and σ_k — the population means and standard deviations of ℓ_∞ and k respectively. However, for at least some parameter values, widely different variances may be associated with similar expected values of ΔL (Table 1). Unfortunately, the data from which the parameter values used in Table 1 were estimated are too few ($n = 35$) to determine whether this disparity in variances is typical of the species or simply an artefact of the model. Kirkwood & Somers (1984) used this model with $\sigma_k = 0$. In that case ΔL depends on ΔT alone.

Sainsbury (1980) concludes that, if his model is correct, then conventional methods of fitting tagging data to the von Bertalanffy model may seriously underestimate the parameter k. This is certainly true if the object is to estimate the population means, $\bar{\ell}_\infty$ and $\bar{k}$, of the von Bertalanffy parameters. However, the usual aim is to estimate parameters which describe mean growth. This is not at all the same thing. To see this it is only necessary to note that the growth described by population mean parameter values will depend on which pair of parameters is used to describe the von Bertalanffy curve. For example, different results will be obtained if k is replaced by e^{-k} since $e^{-\bar{k}}$ is not the same as the mean of e^{-k}.

Two further models warrant passing mention, although, since they are not based on simple tagging data, they fall outside the scope of this paper. Krause et al. (1967) used a logistic curve and modelled multiple age-weight observations from chickens raised in a controlled environment, and McCaughran (1981) used a power curve to model tagging data with ages for Pacific halibut. What these models have in common with that of Sainsbury (1980) is the concept of assigning separate parameters to each individual. This seems more defensible for the chicken data, where it may be assumed that growth variability is (mostly) inherent in the individual, than it does for wild fish where much variability must stem from a heterogeneous environment.

Table 1 Means and variances of the length increment (ΔL) for various values of length at tagging (L_1) and time at liberty (ΔT) according to the model of Sainsbury (1980) with von Bertalanffy parameters ($\bar{\ell}_\infty = 131.3$, $\bar{k} = 0.6693$, $\sigma^2_{\ell\infty} = 380.6$, $\sigma_k^2 = 0.0482$) estimated for *Haliotis iris* from Sainsbury (1977: table 3.2).

L_1 (mm)	ΔT (year)	ΔL (mm) mean	ΔL (mm) variance
91	0.1	2.60	2.42
120	0.4	2.62	22.67

OUTLIERS

Outliers may arise in several ways, for example punching or transcription errors, or the recording of a recapture against the wrong tag number. They may also be true observations of freak individuals. However they arise, if they are not dealt with in some way they may exert an undue influence in the model fitting procedure and this may result in parameter values which are not typical of the bulk of observations.

We will assume there is a (small) probability p that the growth increment for any fish could be recorded as any number within the observed range of growth increments (and that all such numbers are equally likely). In the terminology of mixture theory, p is called the contamination probability.

The assumption that outliers are distributed uniformly over some range is quite arbitrary and not supposed to represent reality. As long as the incidence of outliers is low, this model will serve as well as any other to reduce the effect of extreme data points on growth parameter estimates.

THE MODEL

A maximum likelihood model for growth increment data is a means of describing the likelihood of the observed ΔL in terms of T_1, T_2, and L_1. In addition to growth variability and outlier contamination, the model described here has two further components: mean growth and measurement error. It remains to describe these, together with the likelihood function that links them.

Mean growth enters the model as an equation which relates μ, the expected value of ΔL, to T_1, T_2, and L_1. Any functional form (such as Equation 1) may be used, but it is the goodness of fit which should determine whether one form is better than another for a particular data set. Length measurement errors at tagging and recapture may not (without further information) be estimated separately so we combine them and assume that the net measurement error in ΔL is normally distributed with mean m and standard deviation s.

The log likelihood function may now be written as

$$\lambda = \Sigma_i \log [(1 - p)\, \lambda_i + p/R] \tag{9}$$

$$\text{where } \lambda_i = \exp \frac{-\frac{1}{2}(\Delta L_i - \mu_i - m)^2/(\sigma_i^2 + s^2)}{[2\pi\,(\sigma_i^2 + s^2)]^{1/2}},$$

R is the range (largest minus smallest) of observed growth increments, and the subscript i refers to the i-th data point (fish).

The fitting of the model simply involves searching for the set of parameters which maximises λ. A typical list of model parameters is given in Table 2.

TESTING SIGNIFICANCE

In general a model will fit better (or at least as well) if additional parameters are introduced. We require a statistical test to decide whether additional parameters give a significantly better fit. The conventional test for a maximum likelihood model is the standard likelihood ratio test (LRT) which is based on the fact that, under the null hypothesis, twice the difference between the log likelihood function values with and without the additional parameters is (asymptotically) distributed according to a χ^2 distribution with n degrees of freedom (Kendall & Stuart 1967: 231). Thus, for a significant improvement in fit (at the 5% level), the addition of one parameter must increase λ by at least 1.92; for two parameters the gain must be at least 3.00.

Table 2 Model parameters for the case where mean growth is described by Equation 2 and growth variability by Equation 5.

Parameter type	Symbols
Mean growth	g_α g_β
Growth variability	v
Measurement error	m s
Outlier contamination	p

The LRT may be used as long as there are no outliers ($p = 0$). However, the theory which supports this test does not, unfortunately, cover models such as this, in which the range of a distribution (in this case R, the range of the contamination distribution) is derived from the data. A reasonable approach, in the presence of outliers, is to apply the test over a plausible range of fixed values of p and R. If all these tests are significant then the additional parameters may be considered to provide a significantly better model fit.

This procedure is not applicable in testing whether there are outliers (i.e., testing the null hypothesis, $p = 0$, against the alternative, $p > 0$) — this is the problem of testing how many components there are in a mixture, a problem addressed recently by Aitkin & Rubin (1985). However, this test will not in general be of interest in the present context — our aim is to estimate growth parameters.

EXAMPLE

Campbell & Collins (1975: table 2) gave tagging data for 140 Pacific bonito *(Sarda chiliensis)*. Fish were tagged at lengths ranging from 310 to 745 mm and were at liberty for periods ranging from 0 to 397 days (plus one fish at liberty for 848 days). There are a number of (mostly minor) discrepancies in the published data between the first 4 columns of data (release and recovery dates and lengths) and columns 5 and 6 (days at liberty and growth). The data used here are from columns 1 to 4 but rows 46 and 47 were discarded because the discrepancies were large. For this example, lengths were converted to cm and times were converted to years with $T = 0$ corresponding to 1 January 1969.

The model fitted was a seasonal von Bertalanffy curve — Equations 2 and 4 — with growth variability according to Equation 5. To gain some idea of the relative importance of the parameters and how well the data are able to estimate them, a stepwise fitting procedure was followed. Initially a simple 3-parameter model was fitted and then parameters were added in the order determined by selecting the parameter which gave the greatest increase in log likelihood. Parameter estimates for this sequence of fits are given in Table 3, Models 1 to 4. At each stage the introduction of an additional parameter resulted in a significantly better model fit.

Model 1 is equivalent to the standard least squares fit. It is implausible (since $g_{35} < g_{55}$) but this is due to outliers. If it were desired to fit this simple model then the introduction of the outlier contamination parameter (Model 5) provides a

Table 3 Log likelihood function values and parameter estimates for 5 growth models fitted to the bonito tagging data of Campbell & Collins (1975). For a significant ($P < 0.05$) improvement in fit the introduction of one (two) extra parameter(s) must increase λ by at least 1.92 (3.00); * indicates parameters held fixed.

Parameter		Model				
Type	Symbol (unit)	1	2	3	4	5
Log likelihood	λ	−262.2	−227.3	−217.4	−205.3	−249.7
Mean growth rates	g_{35} (cm/year)	8.8	8.6	11.0	11.0	8.8
	g_{55} (cm/year)	10.1	6.2	8.3	7.5	6.8
Seasonal variation	u (year)	0*	0*	0*	0.66	0*
	w (year)	0*	0*	0*	0.65	0*
Growth variability	v	0*	0.44	0.43	0.35	0*
Measurement error	s (cm)	1.62	0.99	0.81	0.76	1.26
	m (cm)	0*	0*	−0.56	−0.73	0*
Outliers	p	0*	0*	0*	0.000	0.033

more objective alternative to the common approach of arbitrarily removing data considered as "improbable". The incidence of outliers in Model 5 ($p = 0.033$) is seen to be caused by a lack of fit since no outliers were detected in Model 4 ($p = 0.000$). In both Models 1 and 5 the parameter s is simply the root mean square error, made up of measurement error, individual growth variability, and lack of fit.

It is noteworthy that the first additional parameter selected was that describing growth variability. Also, the marked effect of the measurement bias parameter, m, on mean growth rates is of interest (cf. Models 2 and 3, Table 3). This bias is readily apparent in Campbell & Collins (1975: fig. 9).

No significant improvement in fit was achieved from introducing a further growth variability parameter according to any of Equations 6 to 8. Plots of standardised residuals against L_1, and ΔT showed no pattern (correlations were −0.09 and 0.02 respectively), which suggests both that the seasonal von Bertalanffy curve describes the mean growth well and that the assumption that growth variability depends only on mean growth is supported. The way that growth variability affects the residuals for Model 4 is shown in Fig. 1, and an example of the predicted growth is given in Fig. 2.

The seasonal parameter estimates ($u = 0.66$, $w = 0.65$) indicate peak growth in late August at a rate which is 4.9 times the minimum growth rate in late February. This agrees well with the conclusion, "The modal progression indicated very rapid growth from about May or June until October, and almost no growth from January through April" (Campbell & Collins 1975: 191). However, the mean growth rates of 11.0 and 7.5 cm/year (at 35 cm and 55 cm respectively) are substantially different from the values of 19.4 and 10.1 cm/year which may be calculated from their von Bertalanffy curve (derived from age–length data). Though there are difficulties in comparing growth derived from tagging and age–length data (see Discussion) it would seem that these differences do not support the assertion of Campbell & Collins (1975: 191) that their two data sets are "in good agreement".

Confidence intervals for the parameter estimates may be obtained via the information matrix (Kimura 1980). However, this method assumes that the sample size is large enough for asymptotic results to hold. A way of avoiding this assumption is to use simulations. New values of ΔL are simulated for the given values of T_1, T_2, and L_1 using the model assumptions and the estimated parameters. New parameter estimates are then made for the simulated data (using the maximum likelihood method above) and compared to the assumed parameters. The results from 100 such simulations on the bonito data are shown in Table 4.

These results may be interpreted as telling us how good the present data are for estimating the model parameters, assuming that the model is a true representation of reality. Since the distributions of all the parameter estimates were reasonably symmetric and not remarkably long- or short-tailed, normal approximations may be used to calculate approximate confidence intervals. Thus, for example, we may be approximately 95% confident that the growth rates g_{35} and g_{55} are estimated to within 1.6 cm/year and the time of peak growth rate is correct to within 0.06 year (= 3 weeks). Note that the correlation between the von Bertalanffy

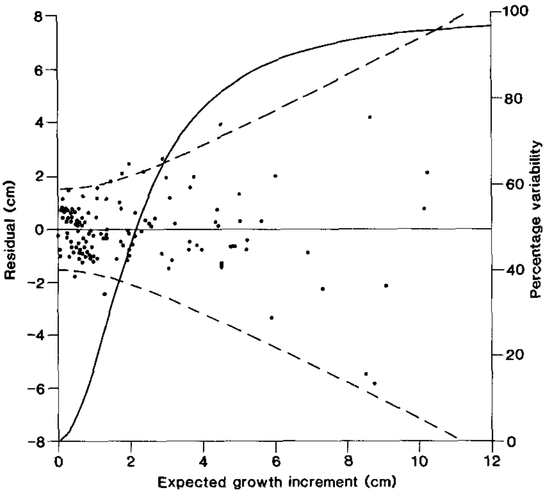

Fig. 1 Scatter plot of residuals (observed-expected growth) against expected growth for best fit (Table 3, Model 4) to bonito data. (One data point, $x = 21.7$, $y = 7.6$, omitted.) Broken lines are $\pm$ 2 standard deviations. Also shown (solid curve) is the percentage of residual which is attributed to growth variability as opposed to measurement error.

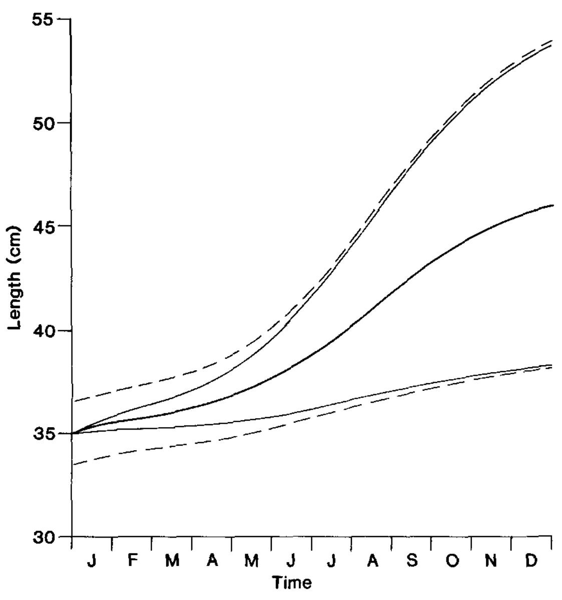

Fig. 2 Predicted mean growth of a 35 cm bonito over one calendar year (heavy line). The light lines show the range of growth and the broken lines show the observed range of growth (both $\pm$ 2 standard deviations) when measurement errors (of mean 0 and standard deviation 0.76 cm) are included.

parameters g_{35}, g_{55} is much lower than is commonly reported between ℓ_∞ and k (e.g., Sanders & Powell, 1979; Kirkwood & Somers 1984).

Mean growth rates at lengths other than 35 or 55 cm may be calculated from Equation 3, and their standard errors can be found using the associated formula

$$(\text{s.e.}\,(g_\gamma))^2 =$$
$$\{[(\gamma - \alpha)\,\text{s.e.}\,(g_\beta)]^2 + [(\beta - \gamma)\,\text{s.e.}\,(g_\alpha)]^2 +$$
$$2\,(\gamma - \alpha)(\beta - \gamma)\,\text{corr}\,(g_\alpha, g_\beta)\,\text{s.e.}\,(g_\alpha)\,\text{s.e.}\,(g_\beta)\}/(\beta - \alpha)^2.$$

Thus, for Model 4, g_{40} has estimated value 8.4 cm/year and s.e. of 0.7 cm/year.

In this example we were able to use the LRT, to test whether additional parameters significantly improved the model fit, since, in each case, no outliers were detected and so p was fixed at 0. To illustrate how this test could be used in the presence of outliers, a new data set was simulated. Values of T_1, T_2, and L_1 were as in the original data, but ΔL values were generated according to the parameters of Model 4 (Table 3) except that p was set at 0.05. Suppose now we wish to test whether the seasonal variation parameters significantly improve the fit for these data. The best fit including these parameters gives $p = 0.087$, and for these data $R = 29.1$. The LRT was applied repeatedly using all combinations of $R = 14.6$, 29.1, and 58.2, and $p = 0.044$, 0.087, and 0.174 (i.e., half and double the estimated values in each case). The test statistics ranged from 14.0 to 19.0, all well above the 5% significance level of 3.0. Thus we may affirm for these data that the addition of the seasonal parameters significantly improve the model fit (i.e., the data provide evidence of seasonal variation in growth rates). These tests involved nine fits of the model including the seasonal parameters. Amongst these fits the variation in the estimates of the seven parameters was less than that observed in the simulations reported in Table 4. Thus we may have confidence that, at least for these simulated data, parameter estimates were not substantially affected by choice of R and p.

A copy of GROTAG, the Fortran computer program used for the above estimations and simulations, is available on request from the author. It uses the NAG minimisation subroutine E04JBF but could readily be adapted to work with other subroutines.

Table 4 Results of 100 simulations based on the bonito data of Campbell & Collins (1975) and the parameter estimates of Model 4, Table 3 (reproduced here as "simulated values"). For reasons of stability p was restricted to be > 0.001 in estimating parameters for simulated data.

Parameter	Simulated values	Estimated values: Mean	s.d.	Correlations: g_{35}	g_{55}	u	w	v	s
g_{35}	11.0	10.9	0.8						
g_{55}	7.5	7.5	0.8	0.4					
u	0.66	0.64	0.12	−0.2	−0.1				
w	0.65	0.65	0.03	−0.5	−0.2	0.2			
v	0.35	0.35	0.05	−0.3	−0.3	−0.1	0.1		
s	0.76	0.76	0.07	0.0	−0.1	0.1	−0.2	−0.2	
m	−0.73	−0.71	0.10	−0.5	−0.5	−0.1	0.2	0.4	0.0
p	0.000	0.001	0.000						

DISCUSSION

Conventional methods of estimating growth from tagging data describe mean or expected growth. The approach described here provides a straightforward means of going one step further and estimating variability in growth. Thus, as well as saying that the expected annual growth of a 35 cm Pacific bonito is 11 cm, we may also say that about two-thirds of such fish will grow between 0.65 and 1.35 times the mean increment.

Ricker (1979: 701) said "The only criteria for choosing a growth curve that have proved valid are goodness of fit and convenience". The merit of the maximum likelihood method is that it allows the use of any growth curve and, as shown here, any functional form for growth variability. For example, this method may be used to choose between various models for seasonal growth: Pitcher & MacDonald (1973) gave two such models (one of which was used here) and a further one was given by Cloern & Nichols (1978). The key point in the present approach is that growth variability is assumed to depend on the expected growth alone. The basis for this assumption is what is intuitively reasonable rather than any physiological evidence. However, the lack of correlation between the standardised residuals and either ΔT or L_1, both in the above example and in four other data sets (author's unpubl. data), suggest that this assumption may provide at least a good first approximation to the truth. Another thing common to all these data sets was the existence of a significant seasonal variation in growth rates. That such variation exists was not surprising (at least for fish from temperate latitudes), but that it should be detectable amongst the measurement errors and individual growth variation was. Incorporating seasonal terms in the model also avoids wasting information as Davis & Kirkwood (1984) did in discarding early returns because seasonal variation would bias parameter estimates.

In the usual approach to estimating mean growth, data points are weighted equally, whereas when a growth variability component is introduced each data point is weighted according to the information it contains. That is part of the reason that there are differences between the growth rate estimates for Models 5 and 2 in the above example (Table 3).

Arbitrary exclusion of outliers, the conventional practice in fitting growth curves, is perfectly acceptable if only mean growth is being estimated, but this subjective procedure may lead to substantial bias in estimating growth variability. The model given here provides an alternative approach which deals with outliers objectively by means of a probabilistic model. Of course, if the true distribution of growth increments is more long-tailed than the normal distribution, then the assumption of normality will lead to an underestimate of growth variability. However, the characterisation of the normal distribution as being that arising from the sum of many small independent effects (Johnson & Kotz 1970: 45) suggests that the assumption of normality here may not be too far wrong. This characterisation is also grounds for treating measurement error as normally distributed.

This more objective approach to outliers should result in better growth parameter estimates. Estimation using a range of values of R and p will establish to what extent they affect these estimates. A disadvantage of the mixture theory approach is that the behaviour of the likelihood ratio test (LRT) for the general model is unknown, so that the ad hoc and slightly cumbersome procedure of evaluating the test statistic for a range of fixed values of p and R must be used. This is certainly a weakness of the model and an area where more theoretical work is needed — either to establish the behaviour of the LRT or to modify the distributional assumptions for outliers. I suspect that, as long as p is not too large (say < 0.20), its effect on the performance of the LRT is unlikely to be great. For much larger values of p one could have little confidence in estimates of p, s, and v without reliable information justifying the arbitrary assumption of a uniform distribution of outliers. Readers who doubt the legitimacy of my suggested approach to the use of the LRT in the presence of outliers may revert to the traditional practice of culling outliers by eye and then apply the model with $p = 0$.

It should be emphasised that, except on very large samples, no test should be thought of as providing a knife-edge division of all possible test results into significant and non-significant — results near to the significance value should be treated as indeterminate.

A distinction which is often ignored in the fisheries literature is that between mean growth as described by length increment data and that derived from length–age data. This distinction is easily overlooked because the same mathematical functions are used to describe both. Consider the form of the von Bertalanffy function given in Equation 1. For tagging data this should be interpreted as relating the expected growth increment for a fish of length L_1 over the time period ΔT; for age–length data it describes the difference in mean lengths between fish of ages T_1 and $T_1 + \Delta T$ where T_1 is the age at which the fish have mean length L_1. As Kirkwood (1983) realised when he tried to fit both types of data to the same curve, there is no reason to assume that one set of parameters will describe both aspects of growth. When estimated from age–length data, ℓ_∞ represents the asymptotic mean length-at-age so there may be many individuals with length $> \ell_\infty$. When ℓ_∞ is estimated from length increment data there will be few if any individuals with $L_1 > \ell_\infty$ since for these fish Equation 1 predicts negative growth. Thus, ℓ_∞ is inappropriate as a descriptor of the growth information contained in length increment data. This was the motivation for the derivation of the parameters g_α and g_β.

A further advantage of these parameters is that they have been shown to have desirable statistical properties. Ratkowsky (1986) considered eight parameterisations of the von Bertalanffy equation as it is used with age–length data, and showed that one was markedly superior to the others in the sense that the parameter estimators have close to linear properties. He has also shown, in a similar manner (Ratkowsky, pers. comm.), that estimators for the parameters given here for use with tagging data are also close to linear.

The approach given here is intended to provide a means of increasing the amount of growth information extracted from tagging data. It does not provide a complete model of growth. The growth rates of fish of a given length may depend on age, for example, but such dependence is not modelled here because tagging data, which contains no (unambiguous) age information, cannot estimate that dependence. A complete model of growth must unify the length- and age-based views of growth which arise from tagging and length–age data respectively. Until such a model is found and shown to be reliable (and the model of Sainsbury (1980) is a bold step in this direction) researchers must tread with caution in comparing growth curves derived from these two types of data.

ACKNOWLEDGMENTS

I am grateful to Bob Gauldie for interesting me in this problem, to Dave Fournier who suggested the mixture theory approach to outliers and whose maximum likelihood computer program (for length frequency analysis) provided the bones for GROTAG, and to David Ratkowsky for examining the statistical properties of the new parameterisation.

REFERENCES

Aitkin, M.; Rubin, D. B. 1985: Estimation and hypothesis testing in finite mixture models. *Journal of the Royal Statistical Society, series B 47*: 67–75.

Campbell, G.; Collins, R. A. 1975: The age and growth of the Pacific bonito, *Sarda chiliensis*, in the eastern north Pacific. *California fish and game 61(4)*: 181–200.

Cloern, J. E.; Nichols, F. H. 1978: A von Bertalanffy growth model with a seasonally varying coefficient. *Journal of the Fisheries Research Board of Canada 35*: 1479–1482.

Cohen, M.; Fishman, G. S. 1980: Modelling growth-time and weight-length relationships in a single year-class fishery with examples for North Carolina pink and brown shrimp. *Canadian journal of fisheries and aquatic sciences 37*: 1000–1011.

Davis, T. L. O.; Kirkwood, G. P. 1984: Age and growth studies on barramundi, *Lates calcarifer* (Bloch), in northern Australia. *Australian journal of marine and freshwater research 35*: 673–689.

Fabens, A. J. 1965: Properties and fitting of the von Bertalanffy growth curve. *Growth 29*: 265–289.

Johnson, N. L.; Kotz, S. 1970: Continuous univariate distributions. 1. Boston, Houghton Mifflin.

Kendall, M. G.; Stuart, A. 1967: The advanced theory of statistics. Volume 2: Inference and relationship. London, Griffin.

Kimura, D. K. 1980: Likelihood methods for the von Bertalanffy growth curve. *Fishery bulletin 77(4)*: 765–776.

Kirkwood, G. P. 1983: Estimation of von Bertalanffy growth curve parameters using both length increment and age-length data. *Canadian journal of fisheries and aquatic sciences 40(9)*: 1405–1411.

Kirkwood, G. P.; Somers, I. F. 1984: Growth of two species of tiger prawn, *Penaeus esculentus* and *P. semisulcatus*, in the western Gulf of Carpentaria. *Australian journal of marine and freshwater research 35*: 703–712.

Knight, W. 1968: Asymptotic growth: an example of nonsense disguised as mathematics. *Journal of the Fisheries Research Board of Canada 25(6)*: 1303–1307.

Krause, G. F.; Siegel, P. B.; Hurst, D. C. 1967: A probability structure for growth curves. *Biometrics 23*: 217–225.

McCaughran, D. A. 1981: Estimating growth parameters for Pacific halibut from mark-recapture data. *Canadian journal of fisheries and aquatic sciences 38*: 394–398.

Pitcher, T. J.; MacDonald, P. D. M. 1973: Two models for seasonal growth in fishes. *Journal of applied ecology 10*: 599–606.

Ratkowsky, D. A. 1986: Statistical properties of alternative parameterizations of the von Bertalanffy growth curve. *Canadian journal of fisheries and aquatic sciences 43*: 742–747.

Ricker, W. E. 1975: Computation and interpretation of biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada 191*.

——— 1979: Growth rates and models. *In:* Hoar, W. S.; Randall, D. J.; Brett, J. R., *ed.*, Fish physiology, Volume 8. Academic Press, New York.

Sainsbury, K. J. 1977: Population dynamics and fishery management of the abalone, *Haliotis iris*. PhD thesis, University of Canterbury, Christchurch, New Zealand.

——— 1980: Effect of individual variability on the von Bertalanffy growth equation. *Canadian journal of fisheries and aquatic sciences 37*: 241–247.

Sanders, M. J.; Powell, D. G. M. 1979: Comparison of the growth rates of two stocks of snapper *(Chrysophrys auratus)* in south-east Australian waters using capture-recapture data. *New Zealand journal of marine and freshwater research 13*: 279–284.

Shepherd, S. A.; Hearn, W. S. 1983: Studies on Southern Australian abalone (genus *Haliotis*). IV. Growth of *H. laevigata* and *H. ruber*. *Australian journal of marine and freshwater research 34*: 461–475.

Walford, L. A. 1946: A new graphic method of describing the growth of animals. *Biological bulletin 90(2)*: 141–147.