SUMMARY PAPER

Reduction of Bias Generated by Age-Frequency Estimation Using the von Bertalanffy Growth Equation[1]

NORMAN V. BARTOO[2] and KEITH R. PARKER[3]

INTRODUCTION

Complex population dynamics techniques rely heavily on age structure information. For some species, accurate ageing methods have not been developed. Often the age structure of a fisheries catch (age-frequency) is estimated from sampled length-frequency (Majkowski and Hampton 1983), the age relationship being described by either an age-length key or a growth curve, such as the von Bertalanffy growth curve (Ricker 1958). The growth curve method is used when there are insufficient data to construct an age-length key. But as noted by Kimura (1977) and later demonstrated by Westrheim and Ricker (1978), under conditions of varying year-class strength and substantial overlap of lengths between ages, age-length keys can yield nearly useless estimates of numbers-at-age. Even with bias correction procedures, the construction of a sufficient key can present difficulties.

In this paper we deal specifically with the von Bertalanffy growth equation and the application of stochastic methods to reduce or eliminate biases. However, it should be noted that the method presented here may be applied to any growth equation, as well as to cases where no growth equation has been fitted or where growth is discontinuous, as in crustaceans.

The von Bertalanffy growth equation mathematically models the relationship between age and length, length being the dependent variable (see Equation (1)). As suggested by Gulland (1969), age can be estimated from length by algebraically rearranging the growth equation so that age is the dependent variable (see Equation (2)). Regardless of whether length or age is the dependent variable, the von Bertalanffy relationship is deterministic, i.e., there is a one-to-one correspondence between age and length.

Use of the von Bertalanffy growth equation for age-frequency estimation results in several types of biases (Powers 1983), different from those inherent in age-length keys. In this paper we document these biases and propose a method for their resolution.

AGE-FREQUENCY BIASES

When growth is modeled according to the von Bertalanffy age-length relationship (Brody 1945; Ricker 1958)

$$L_t = L_\infty \,(1\text{-exp}\,[-k(t-t_0)]), \qquad (1)$$

then age, t, can be converted to length:

$$t = t_0 + \ln\,(1 - L_t/L_\infty)/(-k) \qquad (2)$$

where L_t = length at age t
L_∞ = the asymptotic length
k = the rate at which length reaches L_∞, and
t_0 = hypothetical age at which fish would have zero length.

When computing numbers-at-age from Equation (2), estimation bias occurs from several sources. One bias is due to L_∞ being a fitted parameter. Thus, all numbers-at-length greater than L_∞ must either be eliminated or arbitrarily distributed to older ages. Bias also results when lengths approach L_∞ and are mathematically allocated to ages above those attainable by fish within the stock. As lengths (L) approach L_∞, Equation (2) will yield unreasonably old ages (i.e., ages greater than are known to occur).

Additional bias results from the deterministic nature of the von Bertalanffy equation. For example, back calculations of length to age from Equation (2), which are on a one-to-one basis, result in one determined age for any length. In reality, there can be a number of possible ages for any given length, the most probable age-at-length being that with the highest relative contribution of numbers-at-length. Since these back calculations are without probabilistic arguments, the determined age is not necessarily the most probable for the given length.

Back calculations of length to age also result in a mathematical estimation bias due to the substitution of independent and dependent variables in moving from Equation (1) to Equation (2). The degree of bias is likely to be a function of the amount of residual error in estimating length at age in fitting Equation (1). The bias will probably not be consistent between cases and the degree of bias will have to be considered separately for each case. Consequently, biases associated with equation transformation are not specifically dealt with here.

A computer model can demonstrate these biases. For von Bertalanffy parameters L_∞ = 90.0 units, t_0 = 0.0 units, and k = 0.30, predetermined numbers-at-age are assumed normally distributed with a standard deviation equal to 3 units about the von Bertalanffy length-at-age Equation (1), for ages (1) through (10). A length-frequency vector is then generated by: 1) Multiplying the number-at-age times the probability of age occurring within each 0.5 unit length interval, thus generating a vector of

[1]This paper is summarized from research conducted by the authors and appears in *Fishery Bulletin*, Vol. 81, No. 1, p. 91-96.

[2]Southwest Fisheries Center La Jolla Laboratory, National Marine Fisheries Service, NOAA, 8604 La Jolla Shores Drive, La Jolla, CA 92038.

[3]1837 Puterbaugh Street, San Diego, CA 92103.

Reprinted from: ***In*** **E. D. Prince and L. M. Pulos (editors), Proceedings of the international workshop on age determination of oceanic pelagic fishes: Tunas, billfishes, and sharks, p. 25-27. NOAA Tech. Rep. NMFS 8.**

number-at-length for length intervals between 0 and 100 units for each age, and 2) accumulating numbers-at-length for each length interval over all ages. The numbers-at-age are then deterministically estimated from Equation (2) by accumulating numbers-at-length over the length intervals at age.

The bias from this model is illustrated by input and back-calculated numbers-at-age and their differences, which are listed in columns 2, 3, and 4, respectively, of Table 1. The input numbers-at-age represent a sample age distribution with varying year-class strengths. The differences in column 4 indicate a strong bias which increases with overlap of length distributions at age. The estimated ages of 111 fish were greater than the maximum age, 10. Thirty-five had lengths greater than L_∞ and, consequently, were not classifiable.

Table 1.—Input and estimated numbers-at-age for both the deterministic (column 3) and stochastic (column 5) models, with the input numbers-at-age in column 1. The difference between the input numbers-at-age and the deterministic estimates are given in column 4.

Estimated	Numbers at age			
Age (1)	Input (2)	Deterministic (3)	Diff. (4)	Stochastic (5)
1	200	199	1	200
2	400	399	1	400
3	800	760	40	800
4	200	267	−67	200
5	600	441	159	600
6	300	378	−78	300
7	400	320	80	400
8	300	258	42	300
9	100	164	−64	100
10	100	68	32	100
>10	—	111	−111	—
Inf.	—	35	−35	—

STOCHASTIC MODEL

With estimated variance of length-at-age, a stochastic model can be built from the von Bertalanffy relationship (or any other growth relation): For any age the probability of a specified length interval is the probability of that interval taken over all length intervals containing that age. Thus, for all ages, a probability matrix ("P"-matrix) of dimension r by c can be computed, where r = the number of rows, or length intervals, and c = the number of columns, or ages, then $P\,(1,1) = P$ (max. length, min. age). If the number-at-age vector is "a" ($a_{(1)} = a$ (min. age)) and the number-at-length vector is L ($L_{(1)} = L$ (max. length)), then

$$P\,a = L. \tag{3}$$

And as long as $r > c$, then the numbers-at-age vector can be uniquely solved via least-squares:

$$a = (P'P)^{-1}P'L. \tag{4}$$

Applying this stochastic method (Equation (4)) to the previous example, the numbers-at-age generated from the number-at-length vector is given in column 5 of Table 1. Since the probabilities of the P-matrix are the same as those used to generate the number-at-length vector, it is not surprising that the solution yields unbiased results. This computed example illustrates that the stochastic method yields unbiased estimates of age-frequency.

DISCUSSION

Calculation of age from length via the von Bertalanffy growth equation results in several types of bias. The degree of bias is proportional to overlap in lengths-at-age and changes with weak or strong year-classes. When overlap increases with age, age-frequency estimates will generally be more biased for older ages than for younger ages. When overlap occurs, biases will always result, since the numbers-at-length will be allocated to unreasonably old ages. Any numbers-at-length for lengths greater than L_∞ will be undetermined in age estimation, resulting in downward biases for those ages contributing such lengths.

Age estimation biases can be effectively removed by creating a stochastic model based on a matrix of length interval probabilities at age. The probability matrix (P-matrix) is independent of year-class strength and will effectively remove all sources of estimation bias, except that due to random variation in length-frequency estimation. A probability model of the distribution of length-at-age with estimated parameters is necessary for estimating probabilities of length intervals at age for the P-matrix. As long as the von Bertalanffy growth parameters are correct, the stochastic method based on accurate estimates of variance in length-at-age will yield unbiased results.

There may be serious implications to the bias introduced by using the von Bertalanffy equation without bias correction. In fishery management, the overestimation of maximum age by the deterministic von Bertalanffy equation may produce underestimates of mortality rates, which may result in overestimates of population size and recruitment. Further, the deterministic method tends to "fill in" weak year-classes, which results in underestimates of year-class variability and overestimates of recruitment stability. In general, all of these affect accuracy of a stock assessment and contribute to improper advice for fishery management.

Application of the stochastic method shown here to cover other growth equations and situations, such as discontinuous growth, is handled by simply estimating appropriate elements in the P-matrix for each case.

ACKNOWLEDGMENTS

We thank D. Chapman and A. MacCall for helping to define the problem and evaluating the solution. M. Farber, J. Powers, L. Bledsoe, and G. Sakagawa provided critical reviews and comments for which the authors are grateful.

LITERATURE CITED

BRODY, S.
1945. Bioenergetics and growth, with special reference to the efficiency complex in domestic animals. Reinhold Publ. Corp., N.Y., 1023 p.
GULLAND, J. A.
1969. Manual of methods for fish stock assessment. Part 1. Fish population analysis. FAO (Food Agric. Organ. U.N.) Man. Fish. Sci. 4, 154 p.
KIMURA, D. K.
1977. Statistical assessment of the age-length key. J. Fish. Res. Board Can. 34:317-324.
MAJKOWSKI, J., and J. HAMPTON.
1983. Deterministic partitioning of the catch of southern bluefin tuna, *Thun-*

nus maccoyii, into age-classes using an age-length relationship. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 8:87-90.

POWERS, J. E.
1983. Some statistical characteristics of ageing data and their ramifications in population analysis of oceanic pelagic fishes. U.S. Dep. Commer., NOAA Tech. Rep. NMFS 8:19-24.

RICKER, W. E.
1958. Handbook of computations for biological statistics of fish populations. Fish. Res. Board Can., Bull. 119, 300 p.

WESTRHEIM, S. J., and W. E. RICKER.
1978. Bias in using an age-length key to estimate age-frequency distributions. J. Fish. Res. Board Can. 35:184-189.