SCRS/1981/053

Col.Vol.Sci.Pap. ICCAT, 17 (1): 35-43 (1982)

STOCHASTIC AGE-FREQUENCY ESTIMATION USING THE VON BERTALANFFY GROWTH EQUATION

N. Bartoo, K. Parker

NMFS-Southwest Fisheries Center

SUMMARY

The traditional method of estimating age-frequency from length-frequency via the von Bertalanfy growth equation is deterministic and yields biased results. Most of the bias can be removed by incorporating a stochastic element in the von Bertalanfy relationship. The stochastic element is based on estimated probabilities of lengths by intervals at age, the probabilities being estimated from variances in lengths-at-age. Based on age-length samples from the Pacific bonito fishery the stochastic method gives improved age-frequency estimates over those obtained by the deterministic method.

RESUME

La méthode traditionnelle d'estimation des fréquences d'âge à partir des fréquences de longueur au moyen de l'équation de croissance de von Bertalanffy est déterministe et donnes des résultats faussés. La plupart des influences peuvent être éliminées en incorporant à la relation de von Bertalanffy un élément stochastique. Ce dernier se fonde sur les probabilités estimées de longueur par intervalles d'âge, ces probabilités étant à leur tour estimées à partir des variances de la taille à un âge donné. A partir d'échantillons âge-longueur de la pêcherie de "bonito" du Pacifique, la méthode stochastique a fourni des estimations meilleures que celles qui avaient été obtenues par la méthode déterministe.

RESUMEN

El método tradicional para calcular la frecuencia de edad de la frecuencia de talla a partir de la ecuación de crecimiento de von Bertalanfy es determinista, y produce resultados sesgados. El sesgo puede eliminarse en su mayor parte, incorporando un elemento estocástico en la relación de von Bertalanfy. El elemento estocástico se basa en probabilidades estimadas de tallas por intervalos de edad, calculando las probabilidades a partir de las variantes talla por edad. Basándose en muestras edad/talla de la pesquería de bonito en el Pacífico, el método estocástico facilita estimaciones de frecuencia de edad sobre las obtenidas por el método determinista.

INTRODUCTION

Complex population dynamics techniques rely heavily on age structure information. Frequently, appropriate assessment techniques for a stock require an estimate of the age-frequency of that stock. For example yield-per-recruit analysis (Ricker 1975) is computed on the dynamic relationship between growth and mortality: mortality rates when computed via cohort analysis (Murphy 1965) are based on estimated age-frequency.

For some species accurate ageing methods are not available. When feasible, determining the age of fish and consequently computing an age-frequency, is most accurately accomplished by visual inspection of scales, otoliths or other structures (Ricker 1975). Such visual inspection is time consuming and often expensive. To reduce the cost and time of estimating the age structure of a fisheries catch, age-frequency is usually estimated from sampled length-frequency, the age-length relationship being described in terms of either an age-length key or a von Bertalanfy growth curve (Ricker 1975). The growth curve method is used when there are insufficient data to construct an age-length key.

Age-length keys work on the principle that age can be estimated from length using information contained in a previously aged sample from the population. As long as the proportion of length-at-age remain the same for all ages then the age-length key will yield unbiased estimates of age for any sampled lengths from that population. However since the estimated parameters of an age-length key - proportions of age-at-length - are dependent on the sampled population used to construct the key, the application of the key to the population under altered age structures can yield inaccurate results. Kimura (1977) and later Westerhiem and Ricker (1978) demonstrate that 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.

Clark (1981) effectively removes age-length key bias by first proportioning numbers in length intervals at age over time and then using the matrix of these proportions standardized over time to compute least-squares estimates of age-frequency from the vector of length-frequency. Effective applications of many stock assessment growth and mortality based methods require that ages are expressed in fractions of years (Ricker 1975; Lenarz et al. 1974). The large number of aged fish required to construct a sufficient key for a large number of ages is difficult and expensive to attain. Even with Clark's bias correction procedure, the construction of a sufficient key can present difficulties due to data needs.

The von Bertalanfy growth equation mathematically models the relationship between age and length, length being the dependent variable (see equation (1)). Age is estimated from length by algebraically rearranging the growth equation so that age is the dependent variable (see equation (2)). Regardless of which, length or age, is the dependent variable, the von Bertalanty relationship is deterministic: there is a one-to-one correspondence between age and length.

For the von Bertalanfy growth equation, age-frequency is estimated from a length sample as follows:

1) For each length compute the corresponding age.

2) For each age interval, usually the interval between midpoint ages of adjacent ages, sum the number of aged fish falling within the interval.

3) The number-at-age, age-frequency, is then the total number of aged fish falling within each age interval.

Use of the von Bertalanfy growth equation for age-frequency estimation results in several types of bias, different from those inherent to age-length keys. This paper documents these biases and proposes a method for their resolution.

BIAS

When growth is modeled according to the von Bertalanfy age-length relationship (Brody 1945; Ricker 1975)

$$L_T = L_\infty (1-\exp(-k(t-t_0))), \qquad (1)$$

then age, t, can be converted to length:

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

where L_T = length at age "t",

L_∞ = the asympotic length,

k = the rate of growth decreases,

t_0 = time axis intercept.

When computing numbers-at-age from equation (2) non-mathematical estimation bias occurs. First, is due to L_∞ being a fitted parameter, measured lengths greater than L_∞ yield infinite ages. Thus, all numbers-at-length greater than L must either be eliminated or arbitrarily distributed to older ages. Also bias results when longer, presumably older, fish are allocated to ages above those attained by fish within the stock. As lengths (L) approach L_∞, equation (2) will yield unreasonably old ages.

Additional bias results from the deterministic nature of the von Bertalanfy equation: back calculations of length to age, Equation (2), are on a one-to-one basis. Thus, for any length there is a determined age. 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 switching of independent and dependent variables in going from equation (1) to equation (2). The degree of bias is likely to be a function of the amount of residual error in fitting equation (1). Bias will probably not be consistent between cases and the degree of bias will have to be considered separately for each case. Consequently this bias is not specifically dealt with in this paper.

A computer model example readily demonstrates bias. For von Bertalanfy parameters

L_∞ = 90.0 units,

t_o = 0.0 units, and

k = 0.30,

predetermined numbers-at-age are distributed normally with a standard deviation equal to 3 units about the von Bertalanfy length-at-age, equation (1), for ages (I) through (X). 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 for each age generating a vector of number-at-length for length intervals between 0 and 100 units, and then 2) accumulating numbers-at-length for each length interval over all ages. The numbers-at-age are then deterministically estimated (equation (2)) by accumulating numbers-at-length over the length intervals at age.

Bias is illustrated in Table 1. Input and back-calculated numbers-at-age and their differences are listed in cols. 2, 3, and 4 respectively. The input numbers-at-age model a sample age distribution where either catchability, recruitment, or mortality, or some combination thereof, are age-class variant. Differences, col. 4, indicate a strong bias which increases with overlap of length distributions at age. One hundred and eleven fish were aged to be greater than the maximum age, 10. Thirty five had lengths greater than L and, consequently, were not classifiable.

BIAS RESOLUTION

With estimated variance of length-at-age a stochastic model can be built from the von Bertalanfy relationship: 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 (min. age, max. length). 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 to the previous example, the least-squares solution, Eq. (4), of the numbers-at-age generated from the number-at-length vector is given in col. 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 least-squares solution yields unbiased results. This computed example serves as a simple proof that the stochastic method yields unbiased estimates of age-frequency.

PACIFIC BONITO

For the Pacific bonito (<u>Sarda chiliensis</u>) of the Eastern Tropical Pacific, Campbell and Collins (1975), using ages estimated from otoliths, estimate the von Bertalanfy growth parameters:

L_∞ = 76.87 cm.

t_o = -0.785 cm.

k = 0.6215.

Numbers-at-length for 1 cm. intervals for ages (I) through (V) are shown in Fig. 1 with the corresponding length-frequency plot in Fig. 2. These numbers represent the 1973 catch from California waters and are a random subset of the data used to estimate the von Bertalanfy parameters.

The length-frequency information and von Bertalanfy parameters are used to generate both deterministic and stochastic estimates of numbers-at-age. The estimated length-at-age and sample standard deviations are:

Age	Length	Standard Deviation
I	51.5 cm.	2.7 cm.
II	63.3 cm.	2.1 cm.
III	69.5 cm.	2.3 cm.
IV	72.9 cm.	2.0 cm.
V	74.8 cm.	1.9 cm.

Deterministic estimates were made on lengths rounded to the nearest tenth centimeter. From equation (2) the deterministic numbers-at-age are shown in col. 3 of Table 2 with the difference between the true and estimated numbers in col. 4. While the estimates are reasonably close over the first two ages they become increasingly disparate for older ages. Thirteen fish had lengths greater than those at the maximum age, (V). Seven had lengths greater than L_∞ and consequently were unclassifiable.

Quarter centimeter intervals were used to compute the stochastic estimates of age frequency. The least-squares results are shown in col. 6 of Table 3, with the difference between true and estimated in col. 7. For all ages, especially the older ages, the least-squares estimates are closer and less biased than those of the deterministic method.

Some insight into the improvement of the stochastic age estimates over the deterministic age estimates can be gained by inspection of Fig. 1. Lengths of age (I) fish overlap those of age (II) fish and vice versa. Since the deterministic cutoff point for age (I) fish is 58.7 cm. (1.5 years), all overlap is lost in the deterministic model. In contrast, for the stochastic model, overlaps in lengths-at-age are shared between ages, the degree of sharing being relative to the probabilities of length intervals at the respective ages.

With increasing age, the extent of relative overlap and, consequently, mis-ageing increases for the deterministic model; allocation of lengths to ages becomes more sensitive. Only if the degree of overlap between adjacent ages is equal do accurate estimates of numbers-at-age result from the deterministic model. In the present example varying yearclass strength and random variability in lengths-at-age offset this sensitive compensatory mechanism needed for accurate estimation with the deterministic model.

Fish lengths above 75.3 cm., the length at age 5.5 years, are misclassified either as older or of infinite ages for the deterministic model. Since for the stochastic model probabilities of length intervals at age exist for all ages and lengths, even for lengths above L_∞, fish at lengths above the 75.3 cm. cutoff point are distributed to all ages relative to their respective probabilities for length intervals.

DISCUSSION

Age-frequency back-calculated from length-frequency via the von Bertalanfy growth equation results in several types of bias. The degree of bias is proportional to overlap in lengths-at-age. When overlap increases with age, age-frequency estimates will generally be more biased for older than younger ages. When overlap occurs, the upper length boundary for the last age attainable, biases will always result, since the numbers-at-length will be allocated to unreasonably old ages. Any numbers-at-lengths for lengths greater than "L_∞" will be lost in age estimation, resulting in downward biases for those ages contributing such lengths.

Age estimation bias 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 non-mathematical estimation bias except that due to random variation in length-frequency estimation. As long as the von Bertalanfy growth parameters remain the same over time, the stochastic method based on accurate estimates of variance in length-at-age will always yield unbiased results.

Estimates of variance in length-at-age are necessary for estimating probabilities of length intervals at age for the P-matrix. As with the Pacific bonito, a sample of age-determined fish can be used to estimate the variances. If age information is unavailable then variances can be estimated from visually separable length-frequency modes. In the case where modes are separable for the first few ages only, there will be a problem in estimating variances for older ages: a model relating the variance in length-at-age with age can be used in estimating variances for these older ages. Ricker (1969) proposes that while distributions in lengths-at-age remain normal, variances increase during the first few years, stabilize, and then decrease over the final years. The trend in variances with age for a similar species might also be substituted in cases where variances are unavailable.

The principal strength of the stochastic method lies in that few fish are required to be aged to estimate the P-matrix and in the ability to utilize existing von Bertalanfy growth relations. Accurate estimates of variance in length-at-age can probably be achieved with as few as 20 to 30 fish per age, which is likely to be a much smaller number of fish than needed to estimate accurate proportions of age-at-length necessary to construct an age-length key.

For many species von Bertalanfy growth parameters have been estimated. Since most stocks have variant year class strength, overlaps in lengths-at-age, and lengths exceeding the upper bound for the last age attainable, conversion to a stochastic model may be necessary if unbiased estimates of age-frequency are desired. Re-examination of age-length data used to estimate the von Bertalanfy parameters may be useful in estimating variances in lengths-at-age for the P-matrix. Taking additional age-length samples may prove a cost effective way of improving age-frequency estimation.

LITERATURE CITED

Brody, S. 1945. Bioenergetics and growth. Reinhold Publishing Corp., New York, NY 1023p.

Campbell, W.G. and R.A. Collins. 1975. The age and growth of the Pacific bonito, *Sarda chiliensis*, in the Eastern North Pacific. Calif. Fish and Game, 61(4):181-200.

Clark, W.G. 1981. Restricted least squares estimate of age composition. Can. J. Fish. Aquat. Sci. 38: 297-307.

Kimura, D.K. 1977. Statistical assessment of the age-length key. J. Fish. Res. Board Can. 34: 317-324.

Lenarz, W.H., W.W. Fox, G.T. Sakagawa, and B.J. Rothschild. 1974. An examination of the yield per recruit basis for a minimum size regulation for the Atlantic yellowfin tuna, *Thunnus albacares*. Fish. Bull.: Vol. 72, No. 1.

Murphy, G.I. 1965. A solution of the catch equation. J. Fish. Res. Board Can., 9:405-91.

Ricker, W.E. 1969. Effect of size-selective mortality and sampling bias on estimates of growth, mortality, production and yield. J. Fish. Res. Board Can. 26:409-434.

______________. 1971. Methods for assessment of fish populations in fresh water. IBP Handbook No. 3. 2nd ed. Blackwell Scientific Publications, Oxford and Edinburgh. 348p.

______________. 1975. Handbook of computations for biological statistics of fish populations. Bull. Fish. Res. Board Can. 119: 300p.

Westerheim, 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.

	NUMBERS-AT AGE			
AGE (1)	INPUT (2)	DETERMIN-ISTIC (3)	DIFF. (4)	STOCHASTIC (5)
I	200	199	1	200
II	400	399	1	400
III	800	760	40	800
IV	200	267	-67	200
V	600	441	159	600
VI	300	378	-78	300
VII	400	320	80	400
VIII	300	258	42	300
IX	100	164	-64	100
X	100	68	32	100
>X	---	111	-111	---
INF.	---	35	-35	---

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

NUMBERS-AT-AGE

AGE (1)	TRUE (2)	DETER-MINISTIC (3)	DIFF. (4)	% DIFF (5)	STOCHAS-TIC (6)	DIFF. (7)	% DIFF (8)
I	424	411	13	3.1	415	9	2.1
II	158	167	-9	-5.7	162	-4	-2.5
III	54	39	15	27.8	49	5	9.3
IV	80	71	9	11.3	85	-5	-6.3
V	21	29	-8	-38.1	26	-5	-23.8
>V	--	13	-13	INF.	--	-	0.0
INF.	--	7	-7	INF.	--	-	0.0

TABLE 2. Deterministic (col. 3) and stochastic (col. 6) estimates of numbers-at-age with their respective differences from the true number-at-age in cols. 4 and 7 for the Pacific bonito (Sarda chiliensis) from 1973 California landings, Collins and Clark ([illegible]).

LENGTH IN CM.	I	II	III	IV	V
80					
79					
78				1	
77				2	3
76				5	2
75			1	8	2
74			9	11	5
73			11	14	3
72			8	20	5
71			8	12	1
70		1	6	4	
69		2	6	2	
68		1	2	1	
67		1	2		
66		3			
65		3			
64		9			
63		15			
62	3	38			
61	1	41			
60	8	25			
59	2	15			
58	1	3			
57		1			
56	1				
55	13				
54	26				
53	44				
52	48				
51	71				
50	77				
49	74				
48	37				
47	13				
46	3				
45	2				
44					
43					
42					
41					

AGE

Fig. 1. Numbers-at-age by length in 1 cm. intervals for the Pacific bonito (Sarda chiliensis).Data from 1973 California landings (Campbell and Collins 1975)