Asian Fisheries Science 1(1987):65-74.
Asian Fisheries Society, Manila, Philippines

https://doi.org/10.33997/j.afs.1987.1.1.006

Bias in Estimating Growth Parameters Using Fabens' Mark-Recapture Procedure

Yew-Hu Chien

Department of Aquaculture
National Taiwan College of Marine
Science and Technology
Keelung, Taiwan

Richard E. Condrey

Coastal Fisheries Institute
Center for Wetland Resources
Louisiana State University
Baton Rouge, Lousiana 70803-7503
USA

Abstract

Simulations of a mark-recapture experiment are used to demonstrate the bias associated with use of a widely accepted equation for estimating growth rates from data where absolute age is unknown. The bias results in an overestimation of K, the growth rate constant, and an underestimation of L_∞, the theoretical maximum size. The bias appears to be associated with a failure to account for the redistribution of the error term when the basic growth equation is transformed to eliminate the necessity of estimating age.

Introduction

It is standard procedure in estimating the growth rate of wild populations to capture a portion of the stock, measure their lengths, mark and release them, and then remeasure their lengths upon recapture at some future time. Among the problems associated with the procedure are those of determining size at age and of analyzing data for animals at large for widely varying times. Normally the animals are assumed to have von Bertalanffy growth kinetics and are fitted by the linear transformations of Walford (1946) or Chapman

(1961) or the nonlinear procedure of Fabens (1965). All these procedures avoid the necessity of determining age and Fabens' procedure has the additional advantage of allowing the direct fitting of data on animals at large for differing times. These transformations, however, alter the assumptions concerning the distribution of the error.

In this article we will demonstrate the bias associated with the most versatile of these transformations, the Fabens' equation. The results are also applicable to the procedures of Walford and Chapman as these are special cases where the time at large is constant or treated as a constant.

Fabens' Derivation

The von Bertalanffy growth equation is of the form:

$$L_t = L_\infty (1\text{-exp } (\text{-}K(t\text{-}t_o))) \qquad ...1)$$

where L_t is the length at time t, L_∞ is the theoretical maximum length, t_o is the theoretical age at which the animal would have zero length if it always followed von Bertalanffy kinetics, and K is the growth rate constant. Let the age at marking and at recapture be m and r, respectively, and the time the organism is at large be dt: $t_r = t_m + dt$. The length at marking is:

$$L_m = L_\infty (1\text{-exp } (\text{-}K(t_m\text{-}t_o))) \qquad ...2)$$

which can be rearranged as

$$\exp (\text{-}K(t_m\text{-}t_o)) = (L_\infty\text{-}L_m) \cdot L_\infty^{-1} \qquad ...3)$$

Similarly defining the length at recapture and substituting into equation (3) we obtain

$$\begin{aligned} L_r &= L_\infty (1\text{-exp } (\text{-}K(t_r\text{-}t_o))) \\ &= L_\infty (1\text{-exp } (\text{-}K(t_m\text{-}t_o)) \cdot \exp (\text{-}K \cdot dt) \\ &= L_\infty (1\text{-}((L_\infty\text{-}L_m) \cdot L_\infty^{-1}) \cdot \exp (\text{-}K \cdot dt)) \\ &= L_\infty + (L_m\text{-}L_\infty) \cdot \exp (\text{-}K \cdot dt) \qquad ...4) \end{aligned}$$

(Fabens 1965). The growth parameters K and L_∞ can be solved by nonlinear regression with L_r, L_m and dt. For the special case where dt is a constant (i.e., all organisms are at large for the same time period), then a linear regression of L_m + dt against L_m will generate a slope of exp (-K·dt) and intercept of L_∞ (1-exp(-K·dt)) (Walford 1946).

We assumed a von Bertalanffy equation of the form

$$L_t = 100 \cdot (1\text{-exp}\ (-0.1 \cdot t))$$

where $t_0 = 0$. To create each data set, we randomly chose 200 ages between t = 0 and t = 30 (the latter being the age where 95% of L_∞ was attained) and solved for the expected length. In all, we generated 40 data sets by replicating the process ten times at each of four different standard deviations: 2.5, 5.0, 7.5 and 10.0. Each data set had a unique distribution of age and variation in size-at-age. To form the final data sets, we randomly paired the 200 observations into 100 'mark-recapture' pairs, assigning the smaller age of the pair to that at marking, and the larger to age at recapture.

Because the simulation resulted in a set of randomly chosen ages characterized by lengths with randomly chosen deviations, the equation of best fit for each data set was expected to be slightly different from that of the original 'population'. We derived an estimate of the equation of best fit for each data set by nonlinear regression of the randomly chosen lengths against their ages. We used this direct fit (DF) of the von Bertalanffy equation as a standard against which we measured the bias associated with Fabens' equation (FB). We fitted both types of equations with the Marquardt procedure available in SAS NLIN (Statistical Analysis Systems 1983).

Comparison of Growth Curves

To facilitate discussion we designate the estimates of K, L_∞ and SD derived from directly fitting von Bertalanffy's curve (DF) as K', L'_∞ and SE', and those derived with the Fabens' procedure (FB) K", L''_∞ and SE". The age at which maximum difference in length occurs between the average growth curves (tmax) derived with FB and DF methods was derived as follows.

The difference in predicted length (L'_t, L''_t) at time t of the two growth curves is

$$\begin{aligned} D &= L''_t - L'_t \\ &= L''_\infty (1\text{-exp}\ (-K''t)) - L'_\infty (1\text{-exp}\ (-K't)) \qquad \text{...5)} \end{aligned}$$

Taking the first derivative with respect to t,

$$dD/dt = L'_\infty K' \cdot \exp(-K't) - L''_\infty K'' \exp(-K''t) \qquad \ldots 6)$$

and setting $dD/dt = 0$ to solve for tmax, we have

$$L'_\infty K' \exp(-K' t_{max}) = L''_\infty K'' \exp(-K'' t_{max})$$
$$L'_\infty K' (L''_\infty K'')^{-1} = \exp((K'-K'')\, t_{max})$$
$$t_{max} = \ln(L'_\infty K' (L''_\infty K'')^{-1}) \cdot (K'-K'')^{-1} \qquad \ldots 7)$$

For the case where $K' = K''$, tmax will be undefined. For all other cases, tmax is derived from equation (7) and the maximum difference between the growth curves is solved by substituting tmax for t in equation (5).

Three major statistical analyses were performed:

- Z-tests conducted on the DF method to test differences between K' and 0.1, between L'_∞ and 100, and between SE's and SDs.
- F-tests (analysis of variances) to test for differences between K' and K'', L'_∞ and L''_∞, and SE' and SE''.
- Linear regressions to test the relationships between the (SE''-SD), (K''-0.1), and (100-L''_∞), and their respective SEs.

Results

Table 1 compares the estimates of L_∞, K and SE derived from DF and FB methods. SE''s were consistently greater than SE's. For 80% of the trials, K''s were greater than K'; whereas 80% of the L_∞"s were smaller than L'_∞s. The situation where $L''_\infty < L'_\infty$ generally corresponded to the case where $K'' > K'$ with only two exceptions. $L''_\infty < L'_\infty$ generally corresponded to the case where $K'' > K'$ with only two exceptions.

Table 2 emphasizes the effect of growth variation (SD) on the growth parameter estimates. The average K'' increased from 0.1015 to 0.1222 directly with SD. Additionally the average difference of $(K'' - K')$ and $(K'' - K')/K'$ increased with SD. On the other hand, growth variation had no effect on K'.

The average L''_∞ decreased as SD increased. $(L''_\infty - L'_\infty)$ increased directly with SD, as did the average relative difference, $(L'_\infty - L''_\infty)/L'_\infty$. There was no obvious effect of SD on the estimates of L_∞

Table 1. Standard errors and estimates of growth parameters K and L_∞ estimated from Fabens' method (SE", K", $L"_\infty$) and from direct fitting method (SE', K', L'_∞) when original length-age data generated with four standard deviations (2.5, 5.0, 7.5 and 10.0).

S.D.	SE"	SE'	K"	K'	$L"_\infty$	L'_∞
2.5	3.0504	2.6765	0.1033	0.0984	99.206	100.483
	2.6747	2.2811	0.0972	0.1011	100.248	99.917
	2.8109	2.3880	0.1018	0.1005	99.450	99.690
	2.9467	2.4251	0.0966	0.0986	101.035	100.684
	3.0917	2.7191	0.1008	0.0996	100.298	100.250
	2.7820	2.4249	0.1092	0.1017	98.752	99.696
	2.6636	2.4190	0.1074	0.1012	98.752	99.608
	2.6055	2.2738	0.0982	0.0989	100.649	100.336
	2.6825	2.4187	0.0992	0.1003	100.162	100.287
	2.8652	2.3875	0.1012	0.0997	99.887	100.352
5.0	4.9512	4.6365	0.1162	0.1015	97.100	99.807
	5.3706	4.6126	0.1126	0.1031	96.553	98.625
	5.7766	5.2691	0.1069	0.0982	99.668	100.531
	5.4284	5.2207	0.1308	0.0994	95.306	100.036
	5.5709	4.7977	0.1201	0.1012	96.679	99.569
	5.4676	4.8687	0.1064	0.1032	98.768	99.063
	5.2443	4.6772	0.1042	0.0977	99.296	101.060
	5.4076	4.7922	0.1114	0.1004	98.427	99.950
	5.9073	5.1048	0.1216	0.1034	95.493	99.462
	5.5070	4.7597	0.0929	0.0986	100.496	99.320
7.5	7.8722	7.5514	0.1282	0.1004	96.275	99.998
	7.7810	6.8453	0.1068	0.0988	97.022	98.839
	7.6545	6.9813	0.0937	0.1028	100.862	98.685
	8.4122	7.2332	0.0929	0.1031	103.406	100.116
	8.5957	7.6944	0.1247	0.0968	95.478	100.328
	7.9356	7.0263	0.1212	0.1092	93.520	95.481
	8.3734	7.4463	0.1170	0.1037	96.272	98.848
	8.6290	7.8720	0.1206	0.1037	96.540	99.908
	8.1820	7.4820	0.1204	0.0979	95.910	100.712
	8.9070	7.4094	0.1102	0.0981	96.966	100.443
10.0	11.2053	10.3339	0.1290	0.0944	97.652	102.929
	10.6071	9.2446	0.1321	0.0891	98.521	105.879
	11.1694	9.1527	0.1413	0.0973	94.254	100.753
	10.2763	9.3067	0.0937	0.1001	100.724	100.278
	11.2707	10.5521	0.1418	0.0955	96.837	101.563
	11.3179	9.8944	0.0980	0.0940	101.426	102.204
	11.3564	9.8152	0.1026	0.0969	100.302	100.849
	12.4038	11.3507	0.1345	0.0967	92.875	101.177
	10.0135	9.3269	0.1288	0.1062	95.658	99.360
	11.7725	10.0656	0.1179	0.1114	93.728	94.237

Table 2. The effects of growth variations of four standard deviations on the growth parameters K and L_∞ and standard errors estimated from direct fitting method (K', L'_∞, SE') and from Faben's method (K", $L"_\infty$, SE"). The values are the average of ten replications.

Growth variations	2.5	5.0	7.5	10.0
K'	0.1000	0.1007	0.1014	0.0985
K"	0.1015	0.1123	0.1136	0.1220
K"-K'	0.0015	0.0116	0.0121	0.0234
(K"-K')/K' (%)	1.4843	11.5415	12.1141	24.3773
L'_∞	100.1303	99.7423	99.3360	100.9228
$L"_\infty$	99.8439	97.7787	97.2252	97.1976
$L"_\infty-L'_\infty$	-0.2864	-1.9636	-2.1108	-3.7252
$(L"_\infty-L'_\infty)/L'_\infty$ (%)	-0.2866	-1.9675	-2.1165	-3.6511
SE'	2.4414	4.8739	7.3542	9.9043
SE"	2.8173	5.4631	8.2343	11.1393
SE'-SE"	-0.3759	-0.5892	-0.8801	-1.2350
(SE'-SE")/SE' (%)	-15.4482	-12.1632	-12.0148	-12.6120

derived from the DF method. The average relative difference of (K"-K') increased from 1.48 to 24% with increasing SD, but the average relative difference of (L'_∞-L''_∞) only increased from 0.29 to 3.65%.

Therefore, the growth variation (SD) had a higher bias-effect on estimates of K than on L_∞. The overall effect of the FB method (Fig. 1) is a growth equation which becomes more biased as the standard deviation increases.

The average SE"s derived from FB method were greater, whereas the average SE's derived from DF method were smaller than respective SDs. The difference of (SE"-SE') increased with increasing SD though the relative difference, (SE"-SE')/SE', did not demonstrate a clear trend.

There was little difference between the average growth curves obtained from the two methods when the standard deviation was low, 2.5 (Fig. 1a; Table 2). At SD = 2.5, the maximum length difference MLD = L''_{tmax} - L'_{tmax} of two average growth curves was 0.37, and the relative maximum difference (RMLD = MLD/L'_{tmax}) was 0.68% (Table 3), occurring at the age of 8.0 (26.7% of the maximum life span, 30). The MLD and RMLD increased with increasing SD and tmax decreased with increasing SD (Figs. 1a-1d; Table 3).

The Z-test (PROC UNIVARIATE NORMAL in SAS) indicated that the distributions of all 80 sets of estimates of K and L_∞ were not different from the normal distribution. The Z-tests showed that in the DF method there were no significant differences between K' and 0.1, or between L'_∞ and 100, or between SE' and SD. Since the residual error distribution in the DF method was in the same format as that in the generated data, and since the Marquardt nonlinear regression technique provides an unbiased least square fit, the estimates were expected to be unbiased.

Table 3. The effects of growth variations of four standard deviations on the ages, tmax, when maximal length differences occur. Tmax per cent is the fraction of the age to the life span. L''_{tmax} and L'_{tmax} are the lengths at tmax from growth curves estimated by Fabens' method and direct fitting method, respectively.

Growth variations	2.5	5.0	7.5	10.0
tmax	8.0162	7.6849	7.5518	7.5045
tmax (%)				
tmax/life span (%)	26.7207	25.6163	25.1727	25.0150
L'_{tmax}	55.2117	53.7387	53.1464	52.7323
L''_{tmax}	55.5891	56.5270	55.9961	58.2987
L''_{tmax} - L'_{tmax}	0.3774	2.7883	2.8497	5.5574
(L''_{tmax} - L'_{tmax})/L'_{tmax} (%)	0.6835	5.1886	5.3620	10.5389

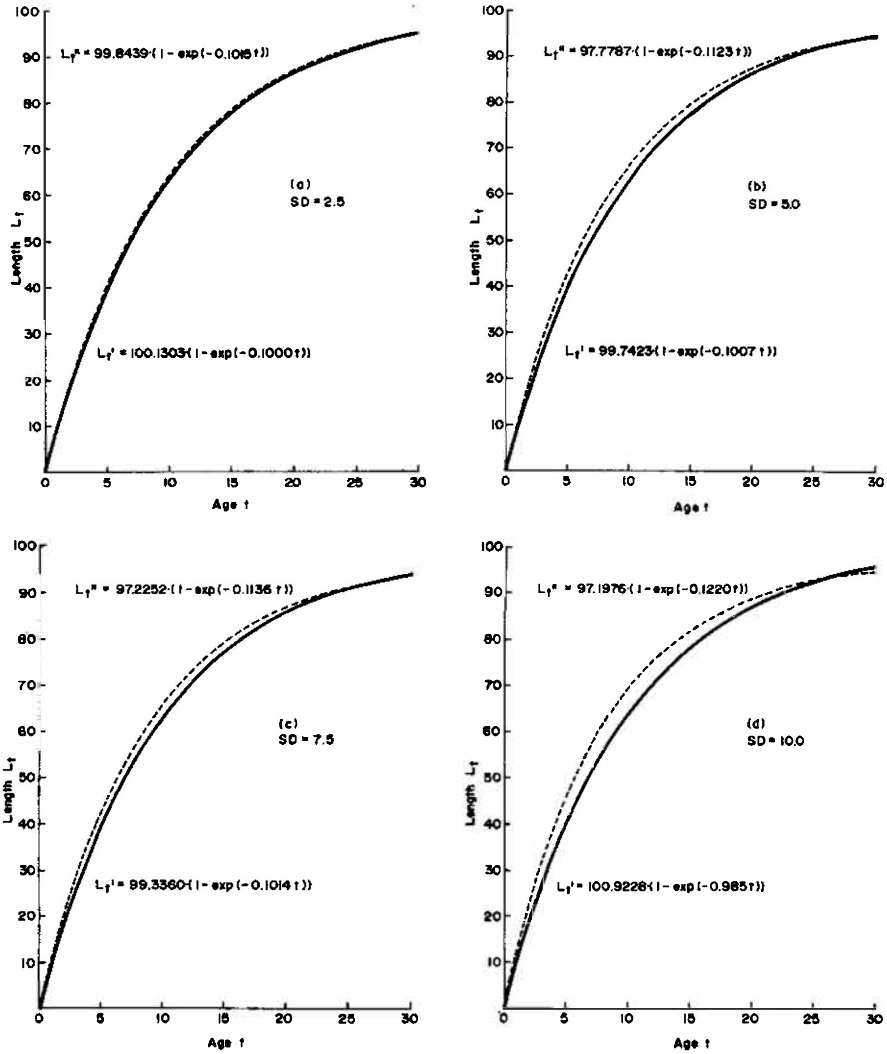

Fig. 1. The average growth curves derived from Fabens' method (broken line) or direct fitting method (solid line) with growth variation of (a) SD = 2.5, (b) SD = 5.0, (c) SD = 7.5, and (d) SD = 10.0.

The analysis of variance test showed that except when SD = 2.5, K" and L''_∞ were highly significantly ($P < 0.01$) or significantly ($P < 0.05$) different from K' and L'_∞. SE"s were highly significantly different from SEs in all four SDs in Table 1.

With the FB method, (SE"-SD), (0.1-K") and (L''_∞-100) increased as the SD increased. There was a significant relationship between (0.1-K") versus SEs (Fig. 2) and between (L"-100) versus SEs (Fig. 3).

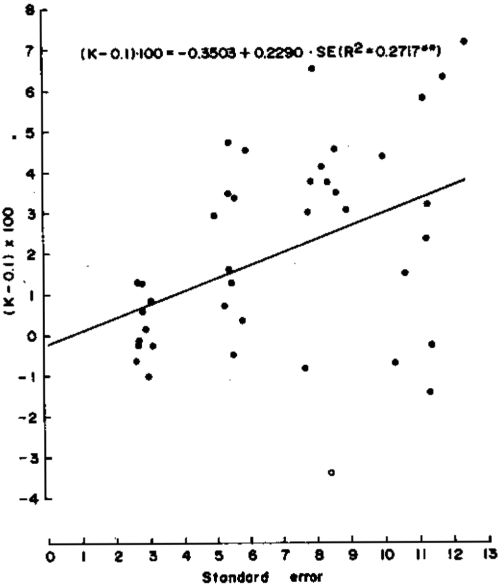

Fig. 2. The relationship of the deviation of K and 0.1 to the standard error. K and standard error were derived from Fabens' method. ** = highly significant ($P < 0.01$).

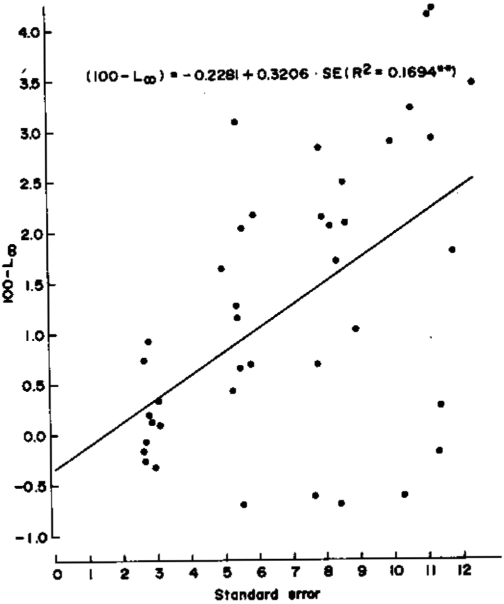

Fig. 3. The relationship of the deviation of 100 and L_{∞} to the standard error. L_{∞} and standard error were derived from Fabens' method. ** = highly significant ($P < 0.01$).

Discussion

In this study, we assumed $L_t = \tilde{L}_t + \varepsilon_t \sim \text{NID}(0, \sigma^2)$, that is the residual term ε_t is additive and normally and independently distributed with a mean of 0 and a constant variance of σ^2. When the von Bertalanffy equation is directly fitted to length-age data, this assumption is usually accepted since the curve is one of the class of transcendental asymptotic functions (Turner et al. 1961) of the form $f(x) = a + b \cdot \exp(-X \cdot C^{-1}) + \varepsilon$. Based on its general acceptance and simplicity we have employed this assumption and noted the bias associated with estimates of growth parameters derived by using Fabens' procedure. We can now search for the causes for this bias and provide possible remedies when using Fabens' equation.

When the length at marking is error free, then

$$\tilde{L}_r = L_\infty + (\tilde{L}_m - L_\infty) \cdot \exp(-Kdt) \qquad \text{...8)}$$

where $\tilde{L}_r$ and $\tilde{L}_m$ are the true values of L_r and L_m without variation. Since $L_r = \tilde{L}_r + \varepsilon_r$ and $L_m = \tilde{L}_m + \varepsilon_m$, equation (8) becomes

$$L_r = L_\infty + (\tilde{L}_m - L_\infty) \cdot \exp(-Kdt) + \varepsilon_r \qquad \text{...9)}$$

When L_r is nonlinearly regressed against $\tilde{L}_m$ and dt, as in equation (9), no bias estimation of K and L_∞ results since there is no variation associated with the independent variables L_m and dt. Equation (9) can be further developed as

$$\begin{aligned} L_r &= L_\infty + (L_m - \varepsilon_m - L_\infty) \cdot \exp(-Kdt) + \varepsilon_r \\ &= L_\infty + (L_m - L_\infty) \cdot \exp(-Kdt) - \varepsilon_m.\exp(-Kdt) + \varepsilon_r \qquad \text{...(10)} \end{aligned}$$

so that the actual of length at marking, L_m, the length at recapture, L_r, and the time at large, dt, can be used.

When using equation (10), two criteria of regression analysis are not met. First, the independent variable L_m is not error free. Second, the residual error is composed of $\varepsilon_m \cdot \exp(-Kdt)$ and ε_r. As such it is no longer normally distributed nor independent of the other variables in the equation since $\varepsilon_m \cdot \exp(-kdt)$ is a function of dt. Fabens' transformation of the von Bertalanffy equation ignores the element $\varepsilon_m \cdot \exp(-Kdt)$, which cannot be regressed because it is not normally distributed and is not independent of the independent variables. In effect, however, Fabens' treatment allows the element $\varepsilon_m \cdot \exp(-Kdt)$ to artificially augment the residual error. Therefore, the estimates of

K and L_∞ are biased and the SE"s derived are significantly greater than the SDs.

Since it is impossible to restore the source of bias $\varepsilon_m \cdot \exp(-Kdt)$ back to L_m, an unbiased solution of K and L_∞ cannot be obtained using a Fabens' type procedure. If, however, there is reasonable evidence to suggest that the organisms at marking belong to the same age group, then we suggest that the average size of the organisms at marking be used, instead of their individual sizes, in Fabens' equation. As we have pointed out in equation (9), no bias should result when the true value of $\tilde{L}_m$ is used. When the size variation is normally distributed, the average lm is an unbiased estimate of L_m.

Acknowledgements

We wish to thank Dr. Hsi-Chiang Liu of the Institute of Oceanography, National Taiwan University, for his critical review of the first draft of this report. Thanks are also due to Ms. Sue-Ming Sun for typing and Ms. Chin Yu-Ping for graphing. This work was partially funded by the Louisiana Sea Grant Program and the Louisiana Fisheries Initiative.

References

von Bertalanffy, L. 1938. A quantitative theory of organic growth (inquiries on growth laws. II). Hum. Biol. 10(2):181-213.

Chapman, D.G. 1961. Statistical problems in dynamics of exploited fisheries populations, p. 151-167. *In* Proc. 4th Berkeley Symp. Math. Stat. and Probability. Contr. Biol. and Probl. Med. 4:153-168, Univ. Calif. Press.

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

Gulland, J.A. 1969. Manual of methods for fish stock assessment, Part 1. Fish population analysis. Food and Agriculture Organization of the United Nations Manuals in Fisheries Science 4, FAO, Rome, Italy. 154 p.

Statistical Analysis Systems. 1983. SAS user's guide: Basics. SAS Institute Inc., Cary, North Carolina.

Turner, M.E., R.J. Monroe and H.L. Lucus, Jr. 1961. Generalized asymptotic regression and nonlinear path analysis. Biometrics 17(1):120-143.

Walford, L.A. 1946. A new graphic method of describing the growth of animals. Biol. Bull. 90:141-147.