Aquatic Sciences 56/3, 1994

1015-1621/030206-14 $ 1.50 + 0.20/0

Estimating fish length and age at 50% maturity using a logistic type model

Y. Chen and J. E. Paloheimo

Department of Zoology, University of Toronto, Toronto, Ontario, Canada M5S 1A1

Key words: Length and age at 50% maturity, logistic model, arcsin square root transformation, nonlinear least squares.

ABSTRACT

We propose a two-parameter logistic type model for estimating fish length and age at 50% maturity using the nonlinear least-squares method. The independent and dependent variables in the model are length (or age) and the corresponding arcsin-square-root transformed proportion of mature fish (P_i). The two parameters in the model are length (or age) at 50% maturity (L50 or A50) and instantaneous rate of maturation (K). A simulation study was conducted to examine the statistical behaviour of the proposed model in estimating L50 and K. The L50 was well estimated with a small bias (<1%) using the proposed model in all simulation cases. The K was well estimated (bias <1%) when its true value and the variance in P_i were small. The proposed model was found to have relatively smaller bias than the probit and Lysack's methods in estimating L50 (or A50). To examine whether there are significant differences in maturation patterns between different groups of fish, we propose fitting the maturation data to the proposed model, and then conducting an analysis of residuals sum of squares to test whether there are significant differences in the fitted models.

Introduction

Length and age at maturity are important parameters of fish life history. It is usually represented by length and age at 50% maturity (L50 and A50; e.g. Trippel and Harvey 1991). Many methods exist in the literature for estimating the two parameters, from the probit method dating back to Leslie et al. (1945) to the recent multivariate model by Richards et al. (1990). Seldom were the statistical properties of estimators examined by methods such as simulation studies. Most of the studies only used real data to test the models, hence it was impossible to "prove" that the model yielded unbiased estimates. In addition, some parameters in the existing models do not have reasonably biological interpretations. Many of the methods need two steps in computing L50 and A50, beginning with estimating the model parameters and then calculating L50 and A50 (e.g. Bowering 1983, Welch and Foucher 1988). There may be large differences in the estimated L50 and A50 using different methods (Trippel and Harvey 1991). The estimates of the uncertainties for L50 and A50 may also differ greatly using different estimation methods.

Fisheries research involving fish maturation studies usually includes the comparison of maturation patterns between different groups of fish (Colby and Nepszy 1981), e.g. to see whether maturation patterns have been significantly changed due to heavy fishing (Anonymous 1983). L50 and A50 are usually used as an index to permit such comparisons (e.g. Colby and Nepszy 1981, Anonymous 1983, Trippel and Harvey 1991). This may be inadequate for examining the differences in the overall maturation patterns, considering the fact that fish can attain the same L50 or A50 at different rates. It may be more appropriate to compare the statistical models estimated from the maturation data. However, few such comparisons for the overall maturation patterns have been done in fisheries studies.

Logistic type models have been applied as the best simple models available in fitting maturity-at-age or maturity-at-length data (Anonymous 1983). Two formulations of the logistic models have been proposed to estimate L50 and A50. One is the equation used by Schnute and Richards (1990) expressed as

$$P_i = \frac{G}{1 + e^{A - B * X_i}} \tag{1}$$

The other used by Lysack (1980) can be written as

$$P_i = \frac{G}{1 + e^{-K * (X_i - C)}} \tag{2}$$

where P_i is the proportion of the mature fish at length or age X_i, G is maximum attainable proportion of the mature fish in the analysis, and A, B, K, and C are the parameters to be estimated. For most fish populations in nature, all individuals will attain maturity after a specific lenght or age. In this case, parameter G in both equations equals 1, and C in Equation (2) is the L50 or A50. K (or B in Equation (1)) is the parameter representing the instantaneous rate of fish maturation, a large K (or B) implies that fish mature during a short time interval.

It is apparent that (2) can be derived from re-parameterizing (1). However, the statistical properties of the estimator differ. Ratkowsky (1990) compared the statistical properties of Equations (1) and (2), and concluded that the nonlinear least squares (NLS) estimation properties for (2) tended to be better than those for (1). Two steps are needed in estimating L50 and A50 if (1) is used: estimating A and B, and then L50 or A50 by dividing A with B, but L50 or A50 equals C in Equation (2) and can be directly estimated.

Logistic regression has been suggested to be a preferred method of estimation for proportional data (McCullagh and Nelder 1983). However, for fish maturity data, a potential problem with this method occurs when all fish are mature or immature in an age class or length interval, the fitted logit or logistic transformation (ln[P/(1-P)] or ln[(1-P)/P]) approaching infinity. This can cause computational problems in the logistic regression (Statistix 1983, SAS 1987). To prevent this problem, cases having proportions of 0 and 1 are proposed to be dropped from the estimation (Statistix 1983). However, for many fish species or populations, maturation is finished within a short period, and hence there exist few age classes or length intervals in which there are both mature and immature fish. This may exclude the chances of applying the logistic regression method in estimating L50 and A50 in practice.

The main concern of the NLS method on the estimation of L50 and A50 is the nature of binomially distributed proportion-of-maturity (PM) data. Because of the essential assumption of normality for the dependent variable, no theoretical justification could be made on the reliability of the NLS estimates when the PM data are not normally distributed. Thus, Equation (2) can not be applied directly to fish maturity data. Two types of transformation are usually used to normalize the proportional data: arcsin√ (arcsine-square-root, ASR) transformation (Zar 1984) and the logistic transformation (Lysack 1980). However, the latter shares the same problem with the logistic regression when proportional data are 0 or 1, and hence is inappropriate to fish PM data.

In this study, a new model is proposed. A simulation study was conducted to examine the estimator properties of the proposed model. Differences in estimating L50 were compared among the proposed models, Equation (2), and the standard probit model. A statistic was proposed to facilitate the comparison of the differences in the overall maturation patterns betwen different groups of fishes.

Methods

Statistical model

Assuming there are n_i mature fish in age class or length interval X_i with the sample size of N_i, thus the proportion of the mature fish in X_i, $P_i = n_i/N_i$. The proposed model can be written as

$$ASR\,(P_i) \;=\; \frac{ASR\,(maximum\;P)}{1 + e^{-K*(X_i - C)}} + \varepsilon_i \tag{3}$$

where X_i is the i^{th} age class or the median of the i^{th} length interval, and C is L50 or A50 when maximum P = 1 (C is a median value dependent on the length interval or age-class selected for X_i). ASRX(P_i) is measured in radians (i.e. ranging from 0 to $\pi/2$ corresponding to P_i from 0 to 1) and is the ASR-transformed proportion of the mature fish corresponding to X_i; i.e.

$$ASR(P_i) \;=\; \arcsin\sqrt{P_i} \tag{4}$$

ASR(maximum P) in (3) equals $\pi/2$ when the maximum $P_i = 1$. ε_i in (3) is an error term. Parameters K and C in (3) can be estimated using the NLS method, and the theoretical justification can be made on the estimated uncertainties of these two parameters if the ε_i meets the Gauss-Markov conditions and the normality assumption. It is known from the statistical theory that proportional data follow a binomial, rather than a normal distribution, and the deviation from normality is great for small or large percentages (i.e. 0–30% and 70–100%. However, the ASR-transformed PM data have an underlying distribution which is nearly normal, with values ranging from 0 to $\pi/2$ (Zar 1984). Thus, the dependent variable in (3) for given X_i is expected to follow the normal distribution if P_i follow the binomial distribution.

The standard probit model can be written as

$$\Phi^{-1}\,(Y) + 5 \;=\; A + B * X \tag{5}$$

where Φ^{-1} is the inverse of the cumulative distribution function of the standard normal distribution, and A and B are parameters to be estimated (SAS 1987). The maximum-likelihood estimates are calculated for parameters A and B. It is obvious that maturation rate K = B and length or age at 50% maturity C = –A/B.

Simulation study

Performance of the models given in Equations (2), (3), and (5) were compared on the simulated data. For simplicity, only the estimation of L50 was examined. Similar results can be expected for estimating A50.

Two PM data sets were simulated (Table 1). For data set I, the midpoint values of length intervals, L_i, ranged from 5 cm to 14 cm with an interval of 1 cm. Parameter L50 was fixed at 10 cm. K had two values, 0.5 and 2. For data set II, L_i (i.e. midpoint values) varied from 17 cm to 35 cm with an interval of 3 cm. Two values were assigned to L50, 19 cm and 27 cm. K was also designated two values, 0.5 and 2. Based on the known K and L50 the "expected" proportions of maturity (EPM) were calculated using Equation (2). A sample of size N_i was drawn at random from a binomial distribution with a given P_i. The "observed" proportions of mature fish (OPM) were then calculated from the number of successes in the random draw. The variance associated with the OPM can be manipulated by changing N_i (as var[OPM]=EPM[1-EPM]/N_i). For both data sets, three levels of variance in the OPM (i.e. A, B, and C) were derived by changing N_i (Table 1). One hundred runs were conducted for each simulation case. Such a design in the simulation study makes it possible for the estimation of K and L50 to be examined systematically with respect to their sensitivity to different patterns of maturation with fish length.

The Marquardt method (SAS 1987) was employed in the NLS estimation of K and L50 using equations (2) and (3). The maximum-likelihood method was used for Equation (5). The NLS estimation was weighted by the sample size at each length interval (i.e. N_i in Table 1, see SAS 1987). The initial assigned values for L50 and K required by the NLS estimation procedures were listed in Table 1. Three methods based on Equations (2), (3), and (5) were referred to as ORIGINAL, ASR, and PROBIT, respectively.

For the simulated data in Table 1, the true values of parameters K and L50 were known *a priori*. Thus, it was possible to determine whether these estimation methods were able to locate the true solutions and how the models were influenced by the random noise incorporated into the simulated data. To measure to what extent the estimates of parameters Q (i.e. L50 and K) differ from the true values of the Q, the following indices were computed based on the 100 runs. They are the mean estimated value,

$$\bar{Q} = \frac{\sum_{k=1}^{100} \hat{Q}_k}{100},$$

and the mean squared error (MSE)

$$MSE\,(\hat{Q}) = \frac{\sum_{k=1}^{100} (\hat{Q}_k - Q)^2}{100}.$$

The latter quantity can be decomposed into a squared estimation bias and an among-run variance, i.e.

$$MSE(\hat{Q}) = (\bar{Q} - Q)^2 + \frac{\sum_{k=1}^{100} (\hat{Q}_k - \bar{Q})^2}{100} .$$

It is thus possible to examine what are the main sources of contribution to the mean squared errors of the estimates.

Application

Length-maturity data were obtained from Richards et al. (1990) for both female and male lingcod (*Ophiodon elongatus*) from stock 3c along the coast of British Columbia, Canada. Age-maturity data were taken from Colby and Nepszy (1981) for female walleye (*Stizostedion vitreum vitreum*) from Lake Winnipeg N., Manitoba, Canada. All three models (i.e. ORIGINAL, PROBIT, and proposed ASR) were applied to estimate K and L50 or A50. The estimation procedure for the PROBIT requires sample size at each length interval or age group. However, no sample size was provided in the original papers (Colby and Nepszy 1981, Richards et al. 1990). Two sample size were used, N = 10 and 100, for each length interval or age group. This permits an examination of the effects

Table 1. Simulated fish lenght data with different sample size at each length interval and with different values of parameters K and L50

# of i[th] length interval	Data set I				Data set II			
		Sample size				Sample size		
	L_i[a]	A	B	C	L_i	A	B	C
1	5	2	10	50	17	4	20	100
2	6	3	14	70	20	6	30	150
3	7	5	20	100	23	6	30	150
4	8	6	24	120	26	6	30	150
5	9	6	24	120	29	9	45	225
6	10	5	20	100	32	13	65	325
7	11	7	28	140	35	12	60	300
8	12	6	24	120				
9	13	5	20	100				
10	14	3	14	70				
L50		10				19 or 27		
K		0.5 or 2				0.5 or 2		
Init. L50[b]		9				10		
Init. K[c]		0.35				0.35		

[a] L_i = midpoint of the length interval (cm)
[b] Init. L50 = initial given value of parameter L50 in nonlinear LS
[c] Init. K = initial given value of parameter K in nonlinear LS

of differences in sample size on the parameter estimation with the PROBIT model. The distributions of residuals computed from fitting the data to the three models were examined with respect to the normality using the Shapiro-Wilk normality test (Proc Univariate, SAS, 1987).

Comparison of maturation patterns

In some fisheries studies it may be desirable to compare the difference in the maturation patterns between different groups of fishes (e.g. female and male fish). Because the proposed logistic model has a nonlinear formulation, an analysis of covariance (ANCOVA) can not be applied. An analysis of the residual sum of squares (ARSS) was suggested in

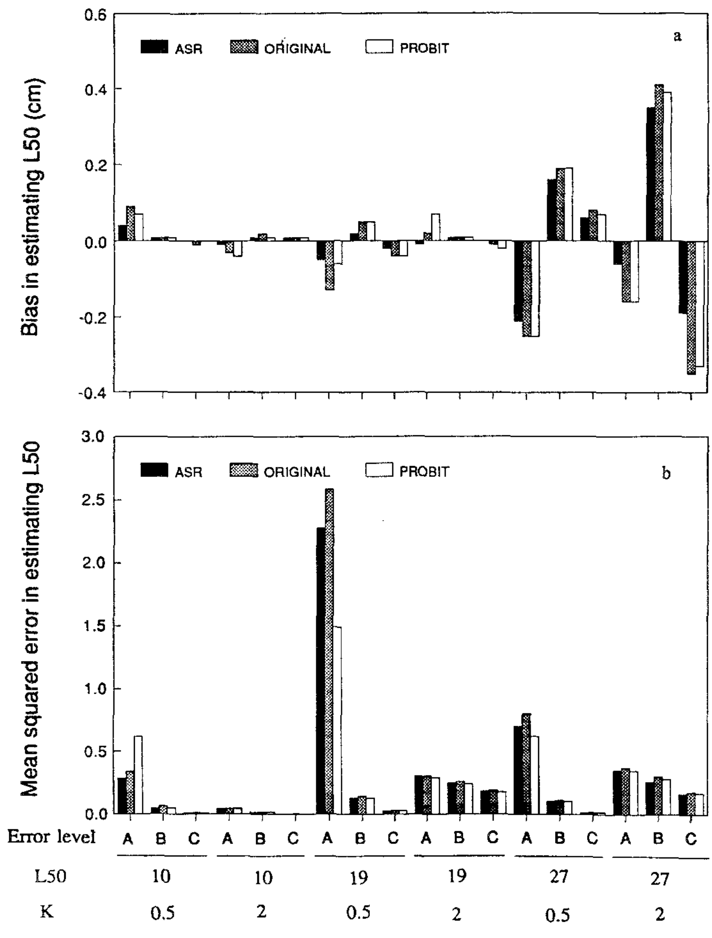

Figure 1. Errors in estimating length at 50% maturity (L50) using the three models (ASR, ORIGINAL, and PROBIT) for simulated data: (a) estimation bias for the estimated L50 calculated as the difference between the average L50 over 100 simulation runs and the true value of L50; (b) mean squared error (MSE) for the estimated L50

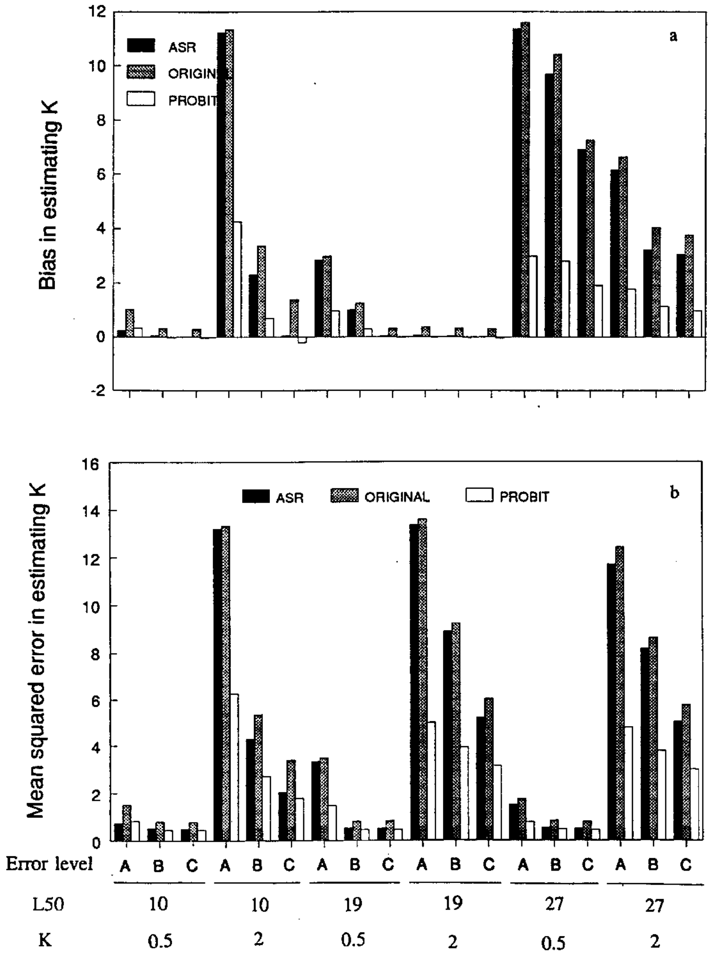

Figure 2. Errors in estimating instantaneous rate of maturation (K) using the three models (ASR, ORIGINAL, and PROBIT) for simulated data: (a) estimation bias for the estimated K calculated as the difference between the average K over 100 simulation runs and the true value of K; (b) mean squared error (MSE) for the estimated K

this study to compare the differences in the maturation patterns between different groups of fishes. An example was given for comparing the maturation patterns between the female and male lingcod. Procedures of the ARSS have been described by Chen et al. (1992).

Results

Simulation study

For the estimate of L50, the ORIGINAL method had the largest bias and MSE in estimating L50 among the three estimation methods for all simulation cases (Fig. 1 a, b). ASR had

Table 2. Frequency of normality in 100 runs of simulation for simulated data defined in Table 1

Data	K	L50	P_i			Arcsin $\sqrt{P_i}$		
			A	B	C	A	B	C
Set I	0.5	10	53	72	90	93	100	100
	2	10	0	0	0	0	0	0
Sett II	0.5	19	0	0	0	0	0	0
	0.5	27	30	44	78	65	100	100
	2	19	0	0	0	0	0	0
	2	27	0	0	0	0	0	0

a smaller bias than PROBIT for all simulation cases (Fig. 1a). However, by the MSE criterion, the differences between ASR and PROBIT depended on the true K value and sample size (N_i), ASR had a larger MSE than PROBIT when the sample size was small (i.e. the OPM data had a large variance; error level A), but with increased sample size (i.e. a decrease in the variance of OPM) ASR tended to have a smaller MSE of L50 than PROBIT (Fig. 1b).

In estimating K, the ORIGINAL method had the largest bias and MSE among the three estimation methods (Fig. 2a, b). The differences between ASR and PROBIT varied with the true values of K and the sample size, ASR tended to have a smaller bias and MSE than PROBIT when K = 0.5 and the sample size was large, but PROBIT had a smaller bias and MSE than ASR when K = 2 (Fig. 2a, b). Both the ASR and ORIGINAL methods tended to over-estimate K when K = 2 (Fig. 2a).

The effectiveness of the ASR transformation in converting proportional data from a binomial into a normal distribution was related to the values of K and the location of L50 in the data (Table 2). A large K value (e.g., K = 2 in the simulation study) implies that fish transform from an immature to a mature condition in a short length interval. In this case, the transformation did not perform well in approximating normality with the proportional data as the observed PM data were either 0 or 1 in almost all length intervals. However, if K had a small value (e.g. K = 0.5 in the simulation study), the ASR transformation greatly improved the data quality with respect to the normality (Table 2). For data set II, L50 was 19 cm, only slightly greater than the smallest length, 17 cm (see Table 1). Thus, L50 was far away from the median of L_i's. As a result, there was no 0 in the "observed" PM data. In this case, the transformation did not do well in normalizing PM data of binomial nature. However, when L50 was 27 cm, a value close to the median of L_i's, the ASR transformation performed much better in normalizing the data (e.g. when K was 0.5, Table 2).

Application

The estimates of L50 and K and their associated standard errors using the three models were summarized in Table 3. The observed and predicted maturation data were plotted in Figure 3 for lingcod and in Figure 4 for walleye. The L50 and K estimated from the proposed ASR model tended to have values between PROBIT- and ORIGINAL-estimated

L50 and K (Table 3). For the PROBIT model, the sample sizes in length intervals or age groups appeared not to affect the estimates of L50 and K, but the estimated standard errors reduced greatly with increase of the sample sizes (Table 3). The proposed ASR tended to have smallest standard error for the estimated L50 among the three models. When sample size in each length interval or age group was large ($N = 100$), the PROBIT had smallest standard error for the estimated K among the three models, however the proposed ASR had smallest standard error for K when the sample size was small ($N = 10$; Table 3). All the estimated parameters differed significantly from 0.

For all three data sets (i.e. female and male lingcod and female walleye), the distribution of residuals derived from the ASR model did not differ significantly from the normality ($P > 0.10$, Shapiro-Wilk normality test). Based on the ARSS, there were significant differences in the ASR-estiamted curve between the female and male lingcods (< 0.001; the calculated $F_{2,12} = 16.7$). Similar results were observed for the ORIGINAL and PROBIT models. However, the Shapiro-Wilk test of the null hypothesis that the distribution of residuals calculated from these two models is normal tended to have smaller values of P compared with these for the ASR model. This indicates that residuals from the ASR model are more likely normally distributed relative to those from the ORIGINAL and PROBIT.

Table 3. Summary of the length (L50) ad age (A50) at 50% maturity and maturation rate (K) estimated using the three methods, ASR, ORIGINAL, and PROBIT for female and male lingcod (data from Richards et al. 1990) and female walleye (data from Colby and Nepszy 1981). $PROBIT_{10}$ and $PROBIT_{100}$ are the PROBIT method with sample sizes being 10 and 100 in estimation, respectively

Data	Parameter		ESTIMATION MODEL			
			ASR	ORIGINAL	$PROBIT_{10}$	$PROBIT_{100}$
Female lingcod	L50 (cm)	Estimate	61.73	61.86	61.36	61.36
		SE	0.551	0.608	1.413	1.013
	K (1/cm)	Estimate	0.22	0.36	0.16	0.16
		SE	0.026	0.045	0.031	0.019
Male lingcod	L50 (cm)	Estimate	55.64	55.76	55.43	55.43
		SE	0.556	0.612	1.288	0.730
	K (1/cm)	Estimate	0.29	0.32	0.21	0.21
		SE	0.043	0.048	0.048	0.022
Female walleye	A50 (year)	Estimate	5.60	5.58	5.63	5.63
		SE	0.136	0.147	0.273	0.082
	K (1/year)	Estimate	1.03	1.38	0.88	0.88
		SE	0.129	0.117	0.178	0.056

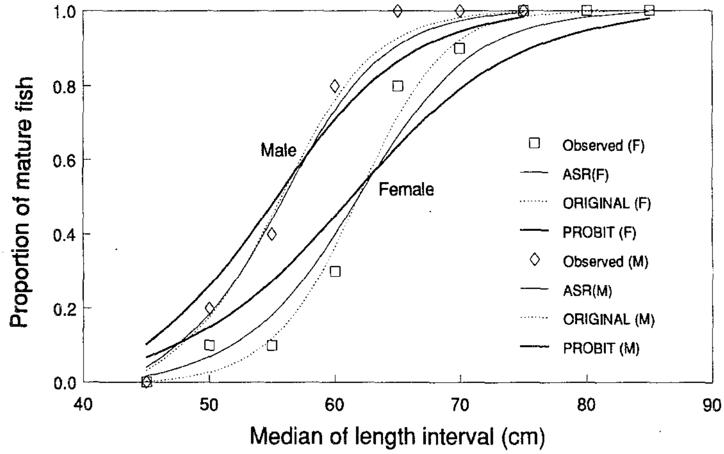

Figure 3. Plot of proportion of mature fish against the corresponding length interval for female (F) and male (M) lingcod from stock 3c along the coast of British Columbia, Canada. Observed = the observed data, and Predicted = the predicted data based on three models (ASR, ORIGINAL, and PROBIT)

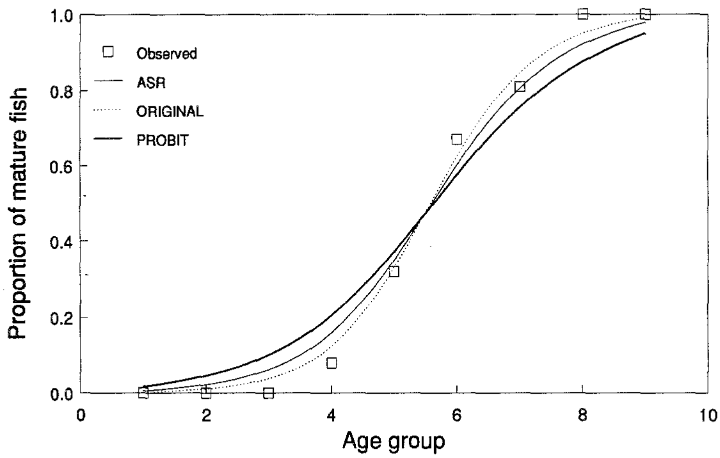

Figure 4. Plot of proportion of mature fish against the corresponding age classes for female walleye from Lake Winnipeg N., Manitoba, Canada. Observed = the observed data, and Predicted = the predicted data based on the three models (ASR, ORIGINAL, and PROBIT)

Discussion

Based on the simulation study, it can be concluded that the ASR is most efficient with the least bias in estimating L50 or A50 while the PROBIT is more suitable in estimating K among three estimation methods.

The ASR model was very robust in estimating L50 with respect to the distribution of the fish PM data in the simulation study. For the simulated PM-at-length data with small

Table 4. Frequency of binomial distribution in 100 runs of simulation for simulated data defined in Table 1

Data	K	L50	P_i		
			A	B	C
Set I	0.5	10	0	11	93
	2	10	40	87	97
Set II	0.5	19	29	96	100
	0.5	27	37	88	98
	2	19	100	96	100
	2	27	72	93	100

sample sizes (e.g. A in Table 1), variances in the PM data were so great that most of the 100 runs of the simulated PM data differed significantly from the binomial distribution. With increased sample size, more PM data sets in 100 runs followed the binomial distribution (Table 4). However, the estimation bias for L50 tended to be similar among different cases of simulation (Fig. 1 a). This may imply that the estimated L50 value is not sensitive to the distribution of the PM data. For the PM data simulated using the same K and L50, the number of the ASR-transformed data sets in 100 runs of simulation following the normal distribution increased with increased sample size (Table 2), so did the number of the PM data sets following the binomial distribution (Table 4). Apparently, the ASR transformation is much less effective in normalizing the non-binomially distributed PM data (e.g. case A) than in normalizing the PM data which have no significant differences from binomiality (e.g. case C; Tables 2 and 4). Based on the simulation study, we conclude that the normalization of the fish PM data following the binomial distribution can be achieved with an ASR transformation. However, for a fish PM data set distributed significantly different from the binomial distribution, the ASR transformation is ineffective in normalizing them.

Equation (1) is one of the most frequently used logistic type models for fitting sigmoidal responses with a lower asymptote of 0 and an upper asymptote of 1 (Ratkowsky 1983, 1990; Schnute and Richards 1990). This model has been found to have good estimation properties (Ratkowsky 1990). As mentioned before, Equation (2) can be obtained by reparameterizing Equation (1). Ratkowsky (1990) found that the estimation properties of parameter L50 tend to be close-to-linear (thus, L50 is an expected-value parameter; Ratkowsky 1983, 1990), and are better than the estimation properties of parameters in Equation (1). In this study, the asymptotic correlation coefficient between K and L50 in the NLS analysis was very small ($|r| < 0.1$) in all simulation runs and the analyses of lingcod and walleye maturation data. This indicates that the estimate of L50 is more or less independent of the K, a desired properties for the NLS estimator (see Ratkowsky 1990).

Four types of maturation patterns have been documented by Trippel and Harvey (1991): the abrupt transition to maturity (type I), the successive increase in proportion of mature fishes (type II), the nonsuccessive increases in proportion of mature fishes (type III), and the nonattainment of 100% maturity (type IV). All four types of maturation patterns can be described by Equation (3) with different magnitudes of K: a large K for type I, a small K for type II, a small K with a large variance for type III, and an extremely small K for type IV.

The type I pattern is found frequently in short-lived fishes (Nikolskii 1969). Because fish finish the transformation from immaturity to maturity abruptly, a large K must be associated with this type of maturation pattern. According to the simulation study, the ASR- and ORIGINAL-estimated K's may not be reliable, and the ASR transformation may not be effective in normalizing the PM data. However, L50 should still be a good estimate. Because of the great chances of non-normality for P_i (due to a large value of K), caution should be used in the interpretation of the asymptotic standard error or 95% confidence intervals of L50. For the sake of certainty, we suggest using some Monte Carlo resampling procedures (e.g. jackknife and bootstrap; Diaconis and Efron 1983, Meyer et al. 1986) to estimate the 95% confidence limits which are independent of the normality assumption of ε_i in Equation (3). Parameter K can be estimated using the PROBIT method (Equation 5).

In practice, for a fish population which is not or is lightly stressed by environmental variables, the type II pattern may be commonly seen if: (1) there is no segregation between mature and immature fish schools within a population; (2) sampling is random; and (3) the sample size is sufficiently large. In this case, the proportion of mature fish can be observed to increase gradually with the length or age. A small magnitude of K can be expected for this maturation pattern. According to the simulation study, the ASR-estimated K and L50 and their variances are credible. The type III pattern may be observed with physiologically stressed fishes. It may also be seen when the quality of the samples is affected by other factors such as small sample size for a certain length or age, size-selectivity of fishing gears, and non-random sampling. For this type of maturation pattern, the proportion of mature fishes varies irregularly over fish length, rather than increasing with fish lenght as for the type II pattern. However, the overall pattern of PM data is still to increase from 0 to 1 with length or age. Similar to the type II, a small K value can be expected for the type III maturation pattern. However, as a result of the irregular variation in P_i over some length intervals or age classes, this pattern is similar to the situation in the simulation study when the variance level in the OPM is A and K is 0.5 (see Table 1). Thus, the variance associated with estimates of K is expected to be large. However, based on the simulation study, it may be reasonable to postulate that the estimate of L50 is credible. The type IV pattern may be observed in fish populations with high density of heavily populated environments (Trippel and Harvey 1991). It may also be observed when there is bias in sampling, which excludes the large and mature fishes. If the proportion of mature fish tends to increase to more than 0.5 with length or age, LS-estimated K and L50 (or A50) can be well estimated. However, if there is no more than 50% mature fish in any length interval or age class, the estimation of L50 and A50 is meaningless.

It is well known that variation in abiotic and biotic environmental variables can cause great variation in fish growth patterns, resulting in great differences in length at age among the individuals (Nikolskii 1969, Brett 1979). In this case, the multivariate model developed by Richards et al. (1990) is a good approach as it considers the effect of both fish length and age on maturation. However, for many fish populations, the environmental stresses may not be so great that the length at age varies greatly among individual fishes. In this case, employing the multivariate model may not be necessary. Moreover, for the multivariate model, there are more parameters to be estimated. It is desirable to minimize the number of parameters for a maturation data set as the number of the cases being used in the estimation of the model is usually small. It is, thus, possible to overfit the model.

Our estimates of L50 differ greatly from those computed in the study of Richards et al. (1990) for the same data. This may result mainly from the differences in the length intervals used in two studies. Richards et al. (1990) used 1 cm as the interval whereas in our study 5 cm was used (because maturity data were only presented in the length interval of 5 cm in Richards et al. 1990). Because the estimated L50 using the proposed model is a median value within 5 cm length interval, the estimated L50 ranges from L50 – 2.5 cm to L50 + 2.5 cm in addition to its associated variance. These ranges certainly include the L50 estimates of Richards et al. (1990).

The distribution of residuals calculated using the proposed model was not different significantly from normality for both lingcod and walleye. This implies that the estimated uncertainties for L50 and K are reliable. Although the estimated L50's were almost the same using the 3 models, the respective asymptotic standard erros differed. The proposed ASR model tended to have smaller uncertainties for the estimated L50's. L50's estimated from the PROBIT model were the same using different sample sizes (i.e. 10 versus 100; see Table 3), but the uncertainties estimated for L50's differed greatly. This may imply that the PROBIT model is not suitable for maturation data of small sample sizes. The ORIGINAL model had larger estimates of K for both lingcod and walleye compared with the other two models. As shown in the simulation study, these estimates might be overestimated with the ORIGINAL for lingcod and walleye.

In practice, we often need to compare differences in maturation data or patterns between different groups of fish. This can be readily done with the proposed model using an analysis of residuals sum of squares (ARSS, Chen et al. 1992) as shown by the comparison of the female and male lingcods. We propose the following procedures to compare the overall maturation patterns between different groups of fish: fitting fish maturation data to the proposed model (i.e. ASR) using the NLS method, and then conducting an ARSS to test whether there are significant differences between the fitted models. If significant differences are observed in the ARSS between the models, we conclude that there are significant differences in the maturation patterns between the fish groups.

Compared with the recently developed methods of estimating the length at 50% maturity, the proposed model in this study is relatively simple and straightforward. Both parameters in the model have the biological explanations. The validity of the estimates of variance in length or age at 50% maturity and its variance can be evaluated by looking at the value of K. Based on this study, we suggest that Equation (3) be used to estimate the length at 50% maturity with the NLS method. The estimate of K can be used as an index in examining the validation of the LS estimation. If the estimate of K is greater than 1, we suggest that a Monte Carlo method (such as bootstrap and jackknife methods) be used to compute the nonparametric 95% confidence limits for the LS estimate of L50 or A50.

ACKNOWLEDGEMENT

This study was partially supported by a Natural Sciences and Engineering Research Council (NSERC) operating grant to J. E. P., and an NSERC postgraduate scholarship to Y. C. We thank Drs. D. A. Jackson, L. J. Richards, B. Shuter, and an anonymous referee for their helpful comments.

REFERENCES

Anonymous, 1983. The identification of overexploitation. Report of SPOF Working Group Number 15. Ontario Ministry of Natural Resources, Ontario.

Bowering, W. R., 1983. Age, growth, and sexual maturity of Greenland halibut, *Reinhardtius hypoglossoides* (Walbaum), in the Canadian Northwest Atlantic. Fish. Bull. U.S.A. 81: 599–611.

Brett, J. R., 1979. Environmental factors and growth. In: Fish Physiology. Vol. VIII ed. by W. S. Hoar, D. J. Randall and J. B. Brett. Academic Press, New York.

Chen, Y., D. A. Jackson and H. H. Harvey, 1992. A comparison for von Bertalanffy and polynomial functions in modelling fish growth data. Canadian Journal of Fisheries and Aquatic Sciences 49: 1228–1235.

Colby, P. J. and S. J. Nepszy, 1981. Variation among stocks of walleye (*Stizostedion vitreum vitreum*): management implicaton. Canadian Journal of Fisheries and Aquatic Sciences 38: 1814–1831.

Diaconis, P. and B. Efron, 1983. Computer-intensive methods in statistics. Scientific American 248: 116–132.

Leslie, P. H., J. H. Perry and J. S. Watson, 1945. The determination of the median bodyweight at which female rats maturity. Proceedings of Zoological Society, London 115: 473–488.

Lysack, W., 1980. 1979 Lake Winnipeg fish stock assessment program. Manitoba Department of Natural Resources, Canada. MS Report No. 30.

McCullagh, P. and J. A. Nelder, 1983. General Linear Models. Monographs on Statistics and Applied Probability. Chapman and Hall, London.

Meyer, J. S., C. G. Ingersoll, L. L. McDonald and M. S. Boyce, 1986. Estimating uncertainty in population growth rates: Jackknife vs Bootstrap techniques. Ecology. 67: 1156–1166.

Nikolskii, R. V., 1969, Theory of Fish Population Dynamics. Oliver & Boyd, Edinburgh, UK.

Ratkowsky, D. A., 1983. Nonlinear Regression Modelling, a Unified Practical Approach, Marcel Dekker, New York, NY.

Ratkowsky, D. A., 1990. Handbook of Nonlinear Regression Models. Marcel Dekker, New York, NY.

Richards, L. J., J. T. Schnute and C. M. Hand, 1990. A multivariate maturity model with a comparative analysis of three lingcod (*Ophiodon elongatus*) stocks. Canadian Journal of Fisheries and Aquatic Sciences 47: 948–959.

SAS. 1987. SAS/STAT Guide for Personal Computers. Version 6 Edition. SAS Institute. Cary, NC.

Schnute, J. T. and L. J. Richards, 1990. A unified approach to the analysis of fish growth, maturity, and survivorship data. Canadian Journal of Fisheries and Aquatic Sciences 47: 24–40.

Statistix, 1983. An Interactive Statistical Analysis Program for Microcomputers. NH Analytical Software, Roseville, MN.

Trippel, E. A. and H. H. Harvey, 1991. Comparison of methods used to estimate age and lenght of fishes at sexual maturity using populations of white sucker (*Catostomus commersoni*). Canadian Journal of Fisheries and Aquatic Sciences 48: 1446–1459.

Welch, D. W. and R. P. Foucher, 1988. A maximum likelihood methodology for estimating lenght-at-maturity with application to pacific cod (*Gadus macrocephalus*) population dynamics. Canadian Journal of Fisheries and Aquatic Sciences 45: 333–343.

Zar, J. H., 1984. Biostatistical Analysis. Second Ed. Prentice-Hall, Englewood Cliffs, NJ.

Received 1 October 1993;
revised manuscript accepted 16 March 1994.