

Abstract.—Size at 50% maturity is commonly evaluated for wild populations, but the uncertainty involved in such computation has been frequently overlooked in the application to marine fisheries. Here we evaluate three procedures to obtain a confidence interval for size at 50% maturity, and in general for $P\%$ maturity: Fieller's analytical method, nonparametric bootstrap, and a Monte Carlo algorithm. The three methods are compared in estimating size at 50% maturity ($I_{50\%}$) by using simulated data from an age-structured population, with von Bertalanffy growth and constant natural mortality, for sample sizes of 500 to 10,000 individuals. Performance was assessed by using four criteria: 1) the proportion of times that the confidence interval did contain the true and known size at 50% maturity, 2) bias in estimating $I_{50\%}$, 3) length and 4) shape of the confidence interval around $I_{50\%}$. Judging from criteria 2–4, the three methods performed equally well, but in criterion 1, the Monte Carlo method outperformed the bootstrap and Fieller methods with a frequency remaining very close to the nominal 95% at all sample sizes. The Monte Carlo method was also robust to variations in natural mortality rate (M), although with lengthier and more asymmetric confidence intervals as M increased. This method was applied to two sets of real data. First, we used data from the squat lobster *Pleuroncodes monodon* with several levels of proportion mature, so that a confidence interval for the whole maturity curve could be outlined. Second, we compared two samples of the anchovy *Engraulis ringens* from different localities in central Chile to test the hypothesis that they differed in size at 50% maturity and concluded that they were not statistically different.

Estimation of size at sexual maturity: an evaluation of analytical and resampling procedures

Rubén Roa

Departamento de Oceanografía
Universidad de Concepción
Casilla 160-C, Concepción, Chile
E-mail address: rroa@udec.cl

Billy Ernst

School of Fisheries, WH-10
University of Washington, Seattle, Washington 98195

Fabián Tapia

Departamento de Oceanografía
Universidad de Concepción
Casilla 160-C, Concepción, Chile

Size at 50% maturity ($I_{50\%}$) is commonly evaluated for wild populations as a point of biological reference (see Table 1). To estimate $I_{50\%}$, a sample of organisms known to have just reached sexual maturity could be available and their arithmetic mean size can be used as an estimator. However, the sample needed to obtain such a design-based estimator (Smith, 1990) for wild populations might be too expensive and would involve time-consuming histological procedures. Fisheries biologists prefer to conceive size at first maturity as the average size at which 50% of the individuals are mature. With this conception, the estimator is not based on a sampling design but on a model of the relation between body size and the number of individuals that are mature from a total number at each of many size intervals. The variance of a design-based estimator is determined by sampling design (Thompson, 1992). The variance of a model-based estimator is not as easily obtained. A sample of published works in the fisheries literature provides a measure of the frequency with which

statistical uncertainty of the model-based $I_{50\%}$ is ignored (Table 1). In this work, we show three alternative procedures: an analytical method derived from generalized linear models (McCullagh and Nelder, 1989), nonparametric bootstrap (Efron and Tibshirani, 1993), and a Monte Carlo algorithm developed in our study. We show by simulation the behavior of the three methods for sample sizes of 500 to 10,000 individuals, concluding that they are similar in terms of bias, length, and shape of confidence intervals but that the Monte Carlo method outperforms the other two methods in percentage of times that the confidence interval contains the true parameter, which remained close to the nominal 95% at all sample sizes.

The problem

In regression analysis, we are usually interested in assigning confidence bounds to the response variable at specified levels of the predictor variable. However, in maturity modeling the attention is turned to the converse problem of

Table 1

An example of published analyses on size at maturity in crustacean and fish populations. CW = carapace width; CL = carapace length; TL = total length; m = males; f = females; ^g = gonadal maturity; ^m = morphometric maturity.

Paper	Species	Fitting method	$I_{50\%}$ (mm)	CI 95%	CI estimation method
Somerton (1980)	<i>Paralithodes camtschatica</i> <i>Chionoecetes bairdi</i>	Weighted nonlinear least-squares	102.8 (CL, m)	Not reported	Random partition of data into subsets and computation of var ($I_{50\%}$) among the N independent estimates of $I_{50\%}$
			101.9 (CL, f)	Not reported	
			114.7 (CW, m)	Not reported	
Campbell and Robinson (1983)	<i>Homarus americanus</i>	Nonlinear least-squares	108.1 (CL, f)	—	—
			92.5 (CL, f)	—	—
			78.5 (CL, f)	—	—
Somerton and MacIntosh (1983)	<i>Paralithodes platypus</i>	Weighted nonlinear least-squares	80.6 (CL, f)	79.4–82.6	Random partition of data into subsets and computation of var ($I_{50\%}$) among the N independent estimates of $I_{50\%}$
			96.3 (CL, f)	95.7–96.9	
			93.7 (CL, f)	92.9–94.5	
			87.4 (CL, f)	86.4–88.4	
Campbell and Eagles (1983)	<i>Cancer irroratus</i>	Nonlinear least-squares	62.0 (CW, m)	—	—
			49.0 (CL, f)	—	—
Somerton and Otto (1986)	<i>Lithodes aequispina</i>	Weighted nonlinear least-squares	97.7 (CL, f)	Values not reported	Bootstrap samples were drawn from the original data set for obtaining independent estimates of $I_{50\%}$, and then computing var($I_{50\%}$) among them.
			99.0 (CL, f)		
			110.7 (CL, f)		
			92.0 (CL, m)		
			107.0 (CL, m)		
130.0 (CL, f)					
Gaertner and Laloé (1986)	<i>Geryon maritae</i>	Nonlinear least-squares	82.8 (CL, f)	—	Confidence regions for the parameters of a logistic function were computed.
Comeau and Conan (1992)	<i>Chionoecetes opilio</i>	Nonlinear least-squares	34.2 (CW, m)	—	—
Armstrong et al. (1982)	<i>Lophius americanus</i> (Pisces: Lophiiformes)	Linear regression of Prop. Mature (arcsine-square root transformed) on Total Length	368.9 (TL, m)	—	—
			458.3 (TL, f)	—	—
Lovrich and Vinuesa (1993)	<i>Paralomis granulosa</i>	Probit	50.2 (CL, m) ^g	—	—
			60.6 (CL, f) ^g	58.3–62.9	Not reported
			57.0 (CL, m) ^m	53.9–60.1	Not reported
			66.5 (CL, f)	63.4–69.5	Not reported
Roa (1993a)	<i>Pleuroncodes monodon</i>	Maximum likelihood	27.2 (CL, f)	24.2–30.2	Ratio of parameter estimates confidence limits
González-Gurriarán and Freire (1994)	<i>Necora puber</i>	Maximum likelihood	54.8 CW, m) ^g	—	—
			49.8 (CW, f) ^g	—	—
			53.3 (CW, m) ^m	—	—
			52.3 (CW, f) ^m	—	—

setting a confidence interval for the size at which a fixed proportion of individuals in a population are sexually mature. That is, we need a procedure for

estimating uncertainty in the predictor variable because management decisions are framed in terms of body size, and hence the uncertainty in estimation

must be transferred to this variable. The first part of the problem is the selection of the maturity model. The available data consist of size (normally length) and maturity status, which will be assumed to take only two values: mature or immature. The predictor variable is continuous and the response variable is dichotomous. With such variables, model errors distribute binomially. Welch and Foucher (1988) recognized this aspect of modeling maturity and showed an efficient procedure based on the principle of maximum likelihood that takes advantage of the binomial nature of the errors.

For dichotomous data modeled as a function of a continuous variable, the following simple logistic function is a consequence of the assumption of a linear relationship between the logit link function and a single predictor variable (Shanubhogue and Gore, 1987; Hosmer and Lemeshow, 1989; McCullagh and Nelder, 1989):

$$P(l) = \frac{\alpha}{1 + e^{\beta_0 + \beta_1 l}}, \tag{1}$$

where $P(l)$ = proportion mature at size l ; and α , β_0 , and β_1 = asymptote, intercept, and slope parameters, respectively (see also Eq. 3).

The estimates of these parameters, given a data set, are chosen from the point at which the product of binomial mass functions of all data points (the likelihood of the data under the model) is a maximum, or equivalently when the negative of the log likelihood

$$-\ell(\alpha, \beta_0, \beta_1) = -\sum_j [(h_j) \ln(P(l)) + (n_j - h_j) \ln(1 - P(l))] \tag{2}$$

is a minimum,

where h = the number of mature individuals; and
 n = sample size at l ;
 $P(l)$ = Eq. 1; and

where a constant term that does not affect the estimation is omitted.

Given the nonlinear nature of normal equations, the minimum is found by an iteration algorithm. The parameters estimated by minimizing Equation 2 are maximum likelihood estimates (MLE). In practical situations, the logistic model may be modified from its original form to allow more biological reality (Welch and Foucher, 1988).

The result from fitting the model (Eq. 1) to the data by using the objective in Equation 2, is a vector of parameter estimates and a covariance matrix, which represents the uncertainty associated to them. With

these results, we may undertake the converse problem of estimating size at fixed $P\%$ maturity, which takes the form

$$l_{P\%} = \frac{1}{\beta_1} \ln \left[\frac{1}{P} - 1 \right] - \frac{\beta_0}{\beta_1}. \tag{3}$$

In Equation 3 it is assumed that the asymptote parameter (α) from Equation 1 is fixed at 1. This assumption is justified on the basis of several published works on size at maturity, showing that all individuals were mature above a given size during the reproductive season (Table 1). Furthermore, if $\hat{\beta}_0$ and $\hat{\beta}_1$ are MLE of β_0 and β_1 and they are used to compute $l_{P\%}$ from Equation 3, then $\hat{l}_{P\%}$ is also MLE. We show below three procedures to perform this task and then test them by generating data from Monte Carlo simulation of the age-size structure and maturity progression of individuals of a hypothetical population.

Analytical estimation

The logistic model in Equation 1 belongs to a class of generalized linear models studied by McCullagh and Nelder (1989). These authors consider the problem of building approximate confidence intervals for the level of the predictor variable that gives rise to a fixed proportion in the response variable. They suggest the use of Fieller's (1944) theorem, according to which the linear combination

$$\beta_0 + \beta_1 l_{P\%} - g(P_0) = 0, \tag{4}$$

where $l_{P\%}$ = the value of the predictor variable for a fixed proportion; and $g(P_0) = \ln(P_0 / (1 - P_0))$ (the logit link function) is approximately normal with mean zero and analytical variance given by

$$v^2(l_{P\%}) = \text{var}(\hat{\beta}_0) + 2l_{P\%} \text{cov}(\hat{\beta}_0, \hat{\beta}_1) + l_{P\%}^2 \text{var}(\hat{\beta}_1). \tag{5}$$

The $100(1-\alpha)\%$ confidence interval is the set of values defined by

$$\frac{1}{\hat{\beta}_1} \left(-\hat{\beta}_0 + g(P_0) \pm z_{\alpha/2} v(l_{P\%}) \right), \tag{6}$$

where $z_{\alpha/2}$ = a quantile of the normal distribution.

Other link functions like probit, common in the field of toxicology (Finney, 1977), are not investigated in this paper.

Bootstrap estimation

Bootstrap is not a uniquely defined concept (Efron and Tibshirani, 1993). This means that bootstrap

samples may be obtained by conceptually different resampling procedures. In the context of logistic regression, it is possible to resample the observational pair (I_r, h_r) , or the semiobservational pair $(I_r, P(l) + \varepsilon_r)$, with ε_r as a realization from the residual distribution of the logistic model. To be valid, this second resampling unit needs the assumption of independence between ε_r and I_r . As stated by Efron and Tibshirani (1993), this is a strong assumption that can fail even when the model $P(l)$ is correct. These authors remark that bootstrapping the observational pair is less sensitive to assumptions than bootstrapping residuals. Therefore, in our work an observation to be resampled with replacement is defined as the pair (length, maturity status). For each and all bootstrap samples, a resampled frequency distribution for $I_{P\%}$ is obtained by fitting the maturity model in Equation 1 with the objective function in Equation 2 and by computing $I_{P\%}$ with Equation 3. The confidence interval is obtained by application of the bias-corrected and accelerated (BCa) method, recommended by Efron and Tibshirani (1993).

Monte Carlo estimation

In Monte Carlo resampling, a model is assumed for the distribution of the estimator and then data are generated computationally to assess the amount of variation (Manly, 1997). In our case, we consider a Monte Carlo resampling of maturity parameters from the modeled joint probability distribution of the estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ for computing $\hat{I}_{P\%}$ from Equation 3. In contrast to the bootstrap approach, the implementation of this approach needs only one fitting of the logistic maturity model and then uses the asymptotic distribution of estimated parameters of the model to generate the probability distribution of the derived statistic $I_{P\%}$. These parameter estimates, $\hat{\beta}_0$ and $\hat{\beta}_1$, distribute asymptotically bivariate normal, with mean vector equal to the population parameters and variance given by their covariance matrix (for nonlinear least-squares: Johansen, 1984; for maximum-likelihood estimates: Chambers, 1977). The bivariate normal distribution of $\hat{\beta}_0$ and $\hat{\beta}_1$ has a strong covariance component, which is the same as to say that these estimates are highly correlated. This also means that much of the variance in one estimate is given by the variance in the other one. Ignoring such correlation would lead to an overestimation of the variance of $I_{P\%}$. In a Monte Carlo setting, the correlation between parameter estimates may be considered in the computation by making the resampling of one estimate conditional on the resampling of the other one. In this work we develop such a technique using the theory of least-squares

estimates of two linearly related normal variates (Draper and Smith, 1981). This approach is justified by the asymptotic nature of standard errors. If $\hat{\beta}_0$ and $\hat{\beta}_1$ are two normal random variables that are linearly related, then we may write the linear equation

$$\hat{\beta}_1 = \hat{b}_0 + \hat{b}_1 \hat{\beta}_0 \tag{7}$$

This equation may be reversed by writing $\hat{\beta}_0$ as a linear function of $\hat{\beta}_1$ because both are random variables. It can be shown that (Draper and Smith, 1981)

$$\hat{b}_1 = \hat{r}_{\hat{\beta}_0, \hat{\beta}_1} \left[\frac{\hat{S}_{\hat{\beta}_1}}{\hat{S}_{\hat{\beta}_0}} \right] \tag{8}$$

where r is the estimated linear correlation coefficient between $\hat{\beta}_0$ and $\hat{\beta}_1$, and $S_{\hat{\beta}_0}$ and $S_{\hat{\beta}_1}$ are the respective standard errors. Furthermore, from Equation 7

$$\hat{b}_0 = \hat{\beta}_1 - \hat{b}_1 \hat{\beta}_0 \tag{9}$$

Therefore, the high correlation coefficient between both maturity parameters can be accounted for by free sampling from the marginal distribution of one parameter estimate (for example, $\hat{\beta}_0$) in each Monte Carlo trial and by computing the other by using

$$\begin{aligned} \beta_{1,j} &= \hat{\beta}_1 - \hat{r}_{\hat{\beta}_0, \hat{\beta}_1} \left[\frac{\hat{S}_{\hat{\beta}_1}}{\hat{S}_{\hat{\beta}_0}} \right] \hat{\beta}_0 + \hat{r}_{\hat{\beta}_0, \hat{\beta}_1} \left[\frac{\hat{S}_{\hat{\beta}_1}}{\hat{S}_{\hat{\beta}_0}} \right] \beta_{0,j} \\ &= \hat{\beta}_1 + \hat{r}_{\hat{\beta}_0, \hat{\beta}_1} \left[\frac{\hat{S}_{\hat{\beta}_1}}{\hat{S}_{\hat{\beta}_0}} \right] [\beta_{0,j} - \hat{\beta}_0] \end{aligned} \tag{10}$$

which is obtained by replacing Equations 8 and 9 in Equation 7. For each trial (indexed by j), a β_0 value is selected from the normal probability distribution defined by its estimate and standard error, and then the mean β_1 value is computed by using Equation 10.

The variance of the estimate is the $\hat{\beta}_0$ variance due to the linear relationship with $\hat{\beta}_0$ plus a residual variance not explained by the relationship. The variance due to the relationship is directly transferred from $\hat{\beta}_0$ to $\hat{\beta}_1$ through the Monte Carlo resampling of $\hat{\beta}_0$ and its mapping onto $\hat{\beta}_1$ by using Equation 10. The residual variance must be added in each trial with

$$\hat{S}_{\hat{\beta}_1, residual}^2 = \hat{S}_{\hat{\beta}_1}^2 \left(1 - \hat{r}_{\hat{\beta}_0, \hat{\beta}_1}^2 \right) \tag{11}$$

where $\hat{r}_{\beta_0, \hat{\beta}_1}^2$ = the proportion of variance due to the linear relationship.

Note in Equation 10 that when $r=0$, the mean of the $\hat{\beta}_{1,j}$ for all j would just be the $\hat{\beta}_1$ estimate, which means that β_0 and β_1 values are independently selected in each trial; note also in Equation 11 that the resampling variance of the estimate would be its total variance. On the other hand, when $r=|1|$, Equation 11 shows that the resampling variance of β_1 would be totally due to the mapping of $\beta_{0,j}$ onto $\beta_{1,j}$ which is expected when the linear relationship between two variables is deterministic. In this case, the algorithm presented here would only perform one Monte Carlo simulation, that on β_0 . Therefore, the algorithm is flexible enough to cover the whole range of correlation between both parameter estimates.

A confidence interval for $I_{P\%}$ may be obtained by the percentile method (Casella and Berger, 1990; Efron and Tibshirani, 1993), for which two computational alternatives are available. If the resampling through the bivariate normal distribution is unbounded, then the $100(1-\alpha)\%$ confidence interval is obtained by ordering the $I_{P\%,j}$ from smallest to largest, and taking as bounds the values at positions $N_{MC}(\alpha/2)$ and $N_{MC}(1-(\alpha/2))$, where N_{MC} is number of Monte Carlo trials. If the resampling through the bivariate normal distribution is bounded, with bounds $\alpha/2$ and $1-\alpha/2$, then the $100(1-\alpha)\%$ confidence interval limits are obtained as the first and last quantiles when ordering the $I_{P\%,j}$ from smallest to largest.

Monte Carlo simulation

To test the performance of the three procedures in estimating $I_{P\%}$ for different sample sizes, we carried out a simulation analysis of a model population with known size-at-age structure, maturity-at-size, and mortality parameters (Table 2). We explored only the behavior of the methods for median (50%) size at maturity ($I_{50\%}$). Performance was evaluated by using four criteria. First, as the proportion of times that confidence intervals did contain the true (and known) parameter ($I_{50\%}$), which we call success:

$$\begin{aligned}
 \text{success} &= 1 - \text{failure} \\
 &= 1 - \frac{\text{number}\{(true - lower)(upper - true) < 0\}}{\text{number of iterations}}
 \end{aligned}
 \tag{12}$$

Our second criterion was bias, evaluated as the average, over trials, of the sufficient statistic:

Parameter	Value
Size-at-age σ^2	4
Growth L	44.55
k (/yr)	0.179
t_0 (/yr)	-0.51
Maturity α	1
β_0	13.648
β_1	-0.502
$I_{50\%}$ (mm)	27.2
Natural mortality M (/yr)	0.6

$$\text{bias} = \frac{\text{resampled median}}{\text{true}}
 \tag{13}$$

which is 1 for an unbiased estimator. The third criterion was the length of confidence intervals:

$$\text{length} = \text{upper} - \text{lower}
 \tag{14}$$

and the fourth and final criterion was the shape of the interval (Efron and Tibshirani, 1993):

$$\text{shape} = \frac{\text{upper} - \text{median}}{\text{median} - \text{lower}}
 \tag{15}$$

which measures asymmetry around the median. In all four measures of performance, “upper” and “lower” refer to the bounds of the confidence interval, “median” is the median $I_{P\%}$, and “true” refers to the true value. The deterministic and stochastic features of our simulation were chosen for a population with features like those previously reported for the squat lobster (*Pleuroncodes monodon*) from the continental shelf off central Chile (Roa, 1993a, 1993b).

To accomplish this task, we implemented the following three-step algorithm, which we called MATSIMVL: step 1, generation of $N_{iter}=5000$ random samples of maturity-at-size data of sample size $N_{sample}=500, 1000, 3000, 5000,$ and $10,000$ individuals (Eqs. 16–19); step 2, estimation of the parameter vector and covariance matrix for each one of these samples (Eqs. 1 and 2); and step 3, running each of the three methods to obtain the 2.5%, 50%, and 97.5%

percentiles of $I_{50\%}$ (Eqs. 3–11) In bootstrap, for each sample size and each of the 5000 trials, we obtained $N_{boot}=5000$ bootstrap samples. With the Monte Carlo method, we resampled parameter estimate values from unbounded normal distributions with $N_{MC}=5000$. For completeness, $N_{Fieller}=1$.

In step 1, the deterministic size structure of the population was conceived as a mixture of normal probability distributions, each normal distribution corresponding to an age class. The proportion of individuals at each size interval was characterized by the following expression

$$P_{n_l} = \frac{\sum_{t=0}^9 \frac{P_{n_t}}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{1-\mu_t}{\sigma}\right)^2\right]}{\sum_{l=0}^{40} \sum_{t=0}^9 \frac{P_{n_t}}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{l-\mu_t}{\sigma}\right)^2\right]}, \quad (16)$$

where the sum is over 10 age classes (0 to 9) and 41 size classes (0 to 40), and where μ_t is determined by a growth equation

$$\mu_t = \mu_\infty \left(1 - e^{-k(t-t_0)}\right) \quad (17)$$

with known parameters (Table 2). Variance of size-at-age (σ^2) is known and constant through age (Table 2), and the proportion of individuals at age (P_{n_t}) is given by a simple exponential mortality model

$$P_{n_t} = \frac{e^{-Mt}}{\sum_{t=0}^9 e^{-Mt}}, \quad (18)$$

where the mortality rate (M) is known and constant through age (Table 2).

Random variability came from two sources. First, samples of the specified sizes were drawn, for each trial, from a uniform probability distribution and compared with the cumulative distribution of Equation 16, accumulating the scores in the respective size intervals. This computation yielded a sample of relative size frequencies p_{nl} . Next, we introduced the second source of uncertainty by assessing the maturity status (mature or immature) of individuals belonging to each size class. This random assignment of maturity status came from resampling the binomial probability distribution

$$P(n = n_{rand}) = \binom{n_{l,rand}}{n_{rand}} P(l)^{n_{rand}} \left(1 - P(l)^{(n_{l,rand}-n_{rand})}\right), \quad (19)$$

where $P(l)$ was computed from the logistic model (Eq. 1) with known maturity parameters (Table 2) and

n_{rand} is the random number of mature individuals out of $n_{l,rand}=N_{sample} \times p_{nl}$ individuals in the size interval l . In this way, step 1 was completed by randomly assigning two properties to each data individual: a size (continuous variable) and a maturity status (dichotomous variable). With these data, step 2 was completed by using a nonlinear parameterization of the logistic model (Eqs. 1 and 2) for obtaining estimates of β_0 and β_1 , and their covariance matrix, by means of the SIMPLEX algorithm (Press et al., 1992). Having this information in hand, step 3 was completed by obtaining 2.5%, 50%, and 97.5% percentiles by each of the three methods. We programmed the MATSIMVL algorithm using Microsoft FORTRAN for PowerStation 4.0 (Microsoft Corp., 1995).

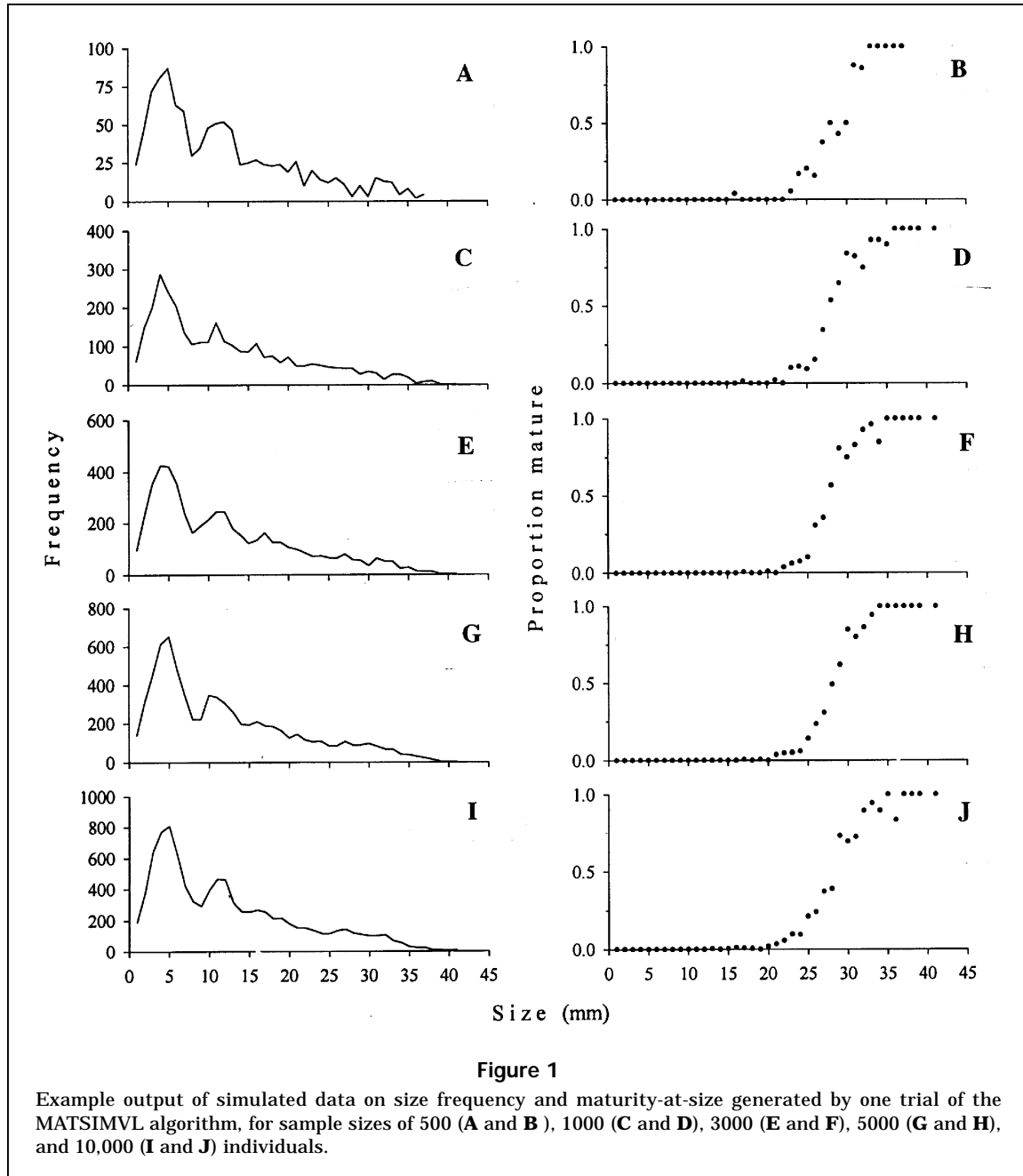
In the case of the Monte Carlo algorithm, we also investigated the effect of the natural mortality parameter, by varying its level in simulation at $M=0.2$, $M=0.4$, $M=0.6$, and $M=0.8$, for sample sizes of $N_{sample}=1000$ and 5000 individuals. N_{iter} and N_{MC} were both kept at 5000.

Finally, we introduce real data to show two applications of the Monte Carlo method developed here. First, we estimate $I_{P\%}$ ($N_{MC}=5000$) for a single population of the galatheid decapod *Pleuroncodes monodon*. In this application, we estimate size confidence intervals for percentages of maturity from 10% to 90% at steps of 10%. In this way a confidence interval for the whole maturity curve is outlined. Second, we compare samples of female anchovy *Engraulis ringens* from two localities 3° of latitude apart ($N_{MC}=5000$) to test the null hypothesis of equal $I_{50\%}$ between them.

Results

The simulation analysis with MATSIMVL yielded size-at-age and maturity-at-size data with the appropriate behavior as N_{sample} increased: size-frequency distributions became smoother and maturity data more closely followed a logistic curve, as shown by one example output of MATSIMVL data-simulating routines (Fig. 1).

A summary of the simulation results is presented in Fig. 2. It shows that, under the simulation conditions, the Monte Carlo method outperformed the bootstrap and the Fieller methods in proportion of success at all sample sizes and that it remained very close to the nominal 95%; the bootstrap method succeeded 94% or less at all sample sizes, whereas the Fieller method was unstable between sample sizes of 500 to 5000, with a minimum of 93% success at 3000 (Fig. 2A). All three methods showed negligible

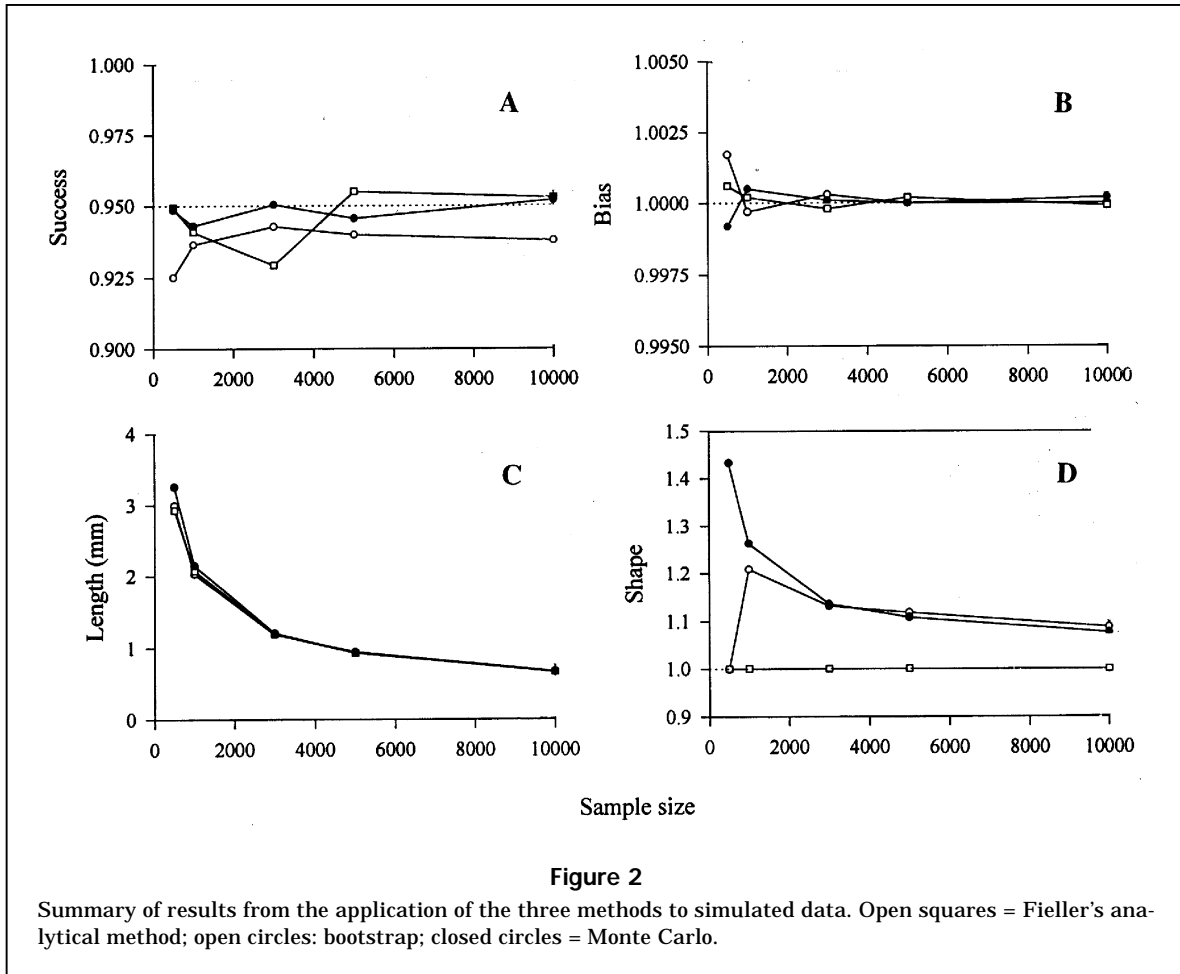


bias at low sample sizes and converged neatly to null bias at large sample sizes (Fig. 2C). Likewise, the three methods behaved exactly the same in length of confidence interval, decaying exponentially as sample size increased (Fig. 2C). Finally, both resampling methods yielded asymmetrical (right-tailed) confidence intervals, converging to the same shape value (ca. 1.1) as sample size increased (by definition, Fieller's method yields a symmetrical confidence interval with shape=1).

Percentage success and bias of the Monte Carlo method under our model for data generation were

fairly insensitive to changes in natural mortality M (Fig. 3) for sample sizes of 1000 and 5000 individuals. Percentage success remained close to the nominal 95% and bias was negligible. The length of the confidence interval and asymmetry, however, increased with increasing mortality, showing that estimation variance was directly proportional to natural mortality rate.

Results of the Monte Carlo algorithm with real data on female squat lobster are shown in Figure 4. The Monte Carlo confidence interval for $I_{50\%}$ was fairly narrow (25.86 to 28.51 mm carapace length),



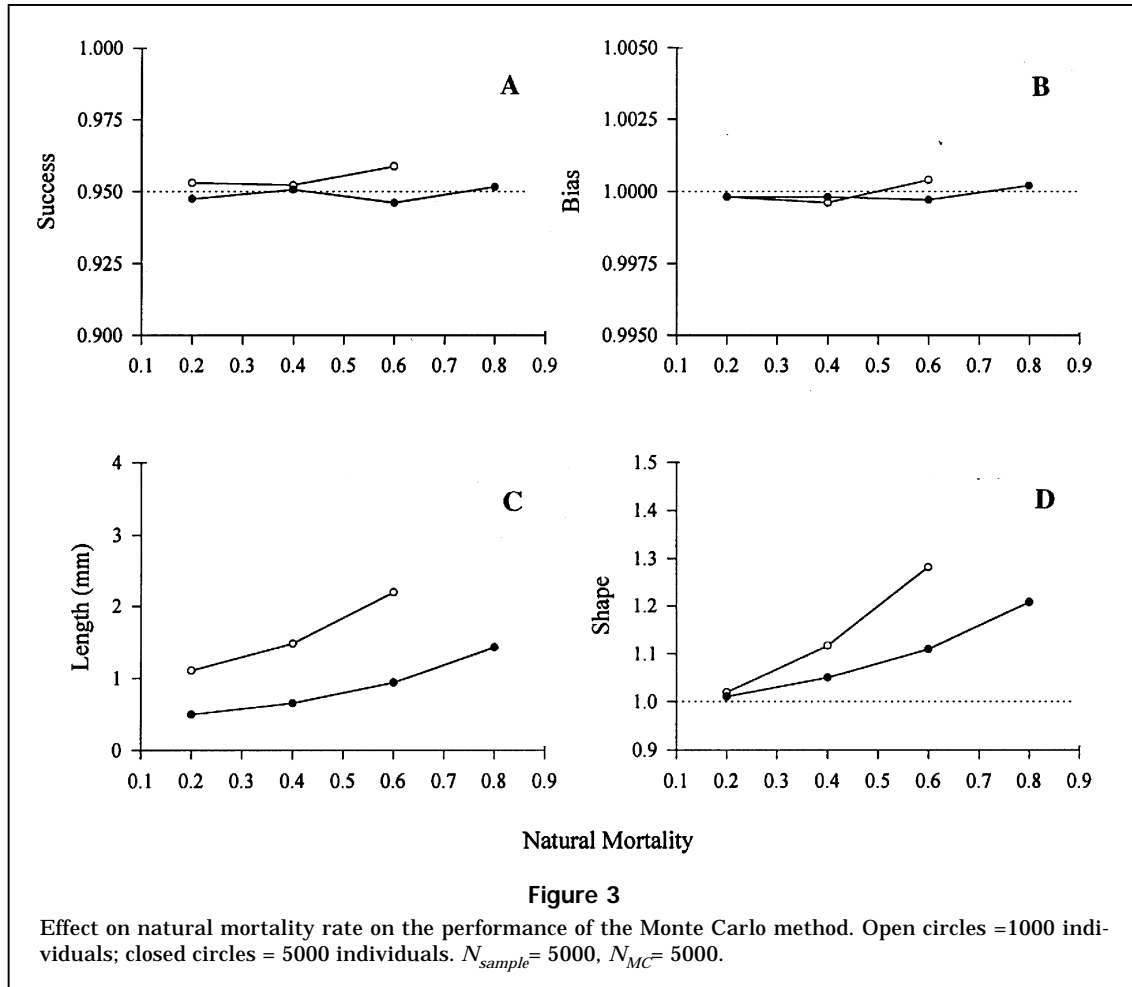
reflecting the effect of large sample size ($N_{sample} = 4458$). The Monte Carlo median $I_{50\%}$ (27.19 mm carapace length) coincided with the MLE of $I_{50\%} = -MLE(\beta_0) / MLE(\beta_1)$. The amplitude of the confidence interval for $I_{P\%}$ showed an increasing trend towards extreme values of P (Fig. 4), a reflection of the algebraic structure of Equation 3. Results of the second application with two samples of female anchovies are shown on Figure 5. Although different in their lengths, probably due to different sample sizes, confidence intervals from the two samples overlapped and have the same upper limit. This result provides support to the hypothesis of equal maturity schedules between female anchovies from the two localities.

Discussion

The logistic model is universally used as a mathematical description of the relation between body size and sexual maturity. To model residuals, however, two different approaches emerge: to consider them

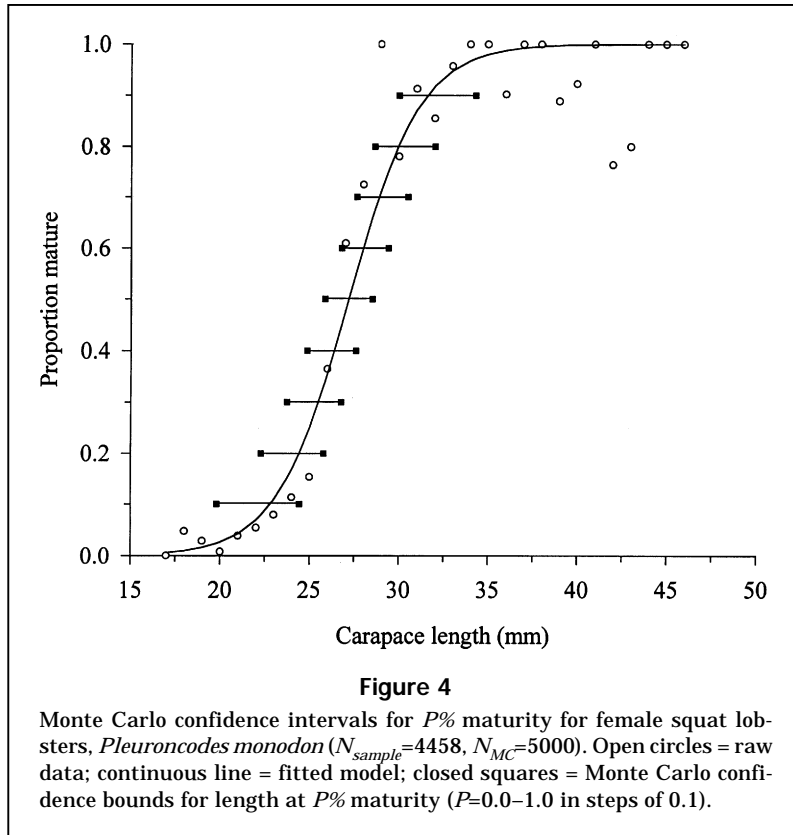
normally or binomially distributed, and closely related to this, to use the data as proportions or as counts. In the first (as far as we know) formal treatment of the problem, Leslie et al. (1945) used the data as proportions, transformed to probit scores, and assumed the normal distribution. Current researchers have not employed probit transformations (but see Lovrich and Vinuesa, 1993) but have continued using data as proportions and the normal distribution for residuals (Table 1). In this work however, and following the arguments by Welch and Foucher (1988), we emphasize the need to estimate the maturity model using the data as counts and therefore to consider residuals as binomially distributed. Under this approach, the standard procedure for fitting the maturity model is logistic regression (Hosmer and Lemeshow, 1989).

When the model has been fitted, the problem of setting a confidence interval for the level of the predictor variable (size) that gives rise to a fixed proportion of maturity is not trivial. We have explored here three approaches: one analytical method based



on Fieller's (1944) theorem and the logit link function, and two computationally-intensive methods based on nonparametric bootstrap of the observational pair (I_p, h_p) and Monte Carlo resampling of parameter estimates from the logistic model. Although this is not an exhaustive set of methods (for example, we did not explore likelihood profiles), they represent a set of conceptually different alternatives to be tested against the model we used to generate simulated data. In particular, both resampling methods are especially useful to obtain the distribution of a function of estimated parameters, such as in Equation 3, mainly because of their mathematical simplicity, which comes at the expense of extensive computation. Our results indicate that the three methods to estimate size at $P\%$ maturity perform almost equally well in terms of bias, length, and shape of the confidence interval, but that Monte Carlo performed better in containing the true parameter within its confidence bounds with the nominal 95% rate. This greater accuracy is accompanied by $N_{boot}-1$ times less computation than bootstrap.

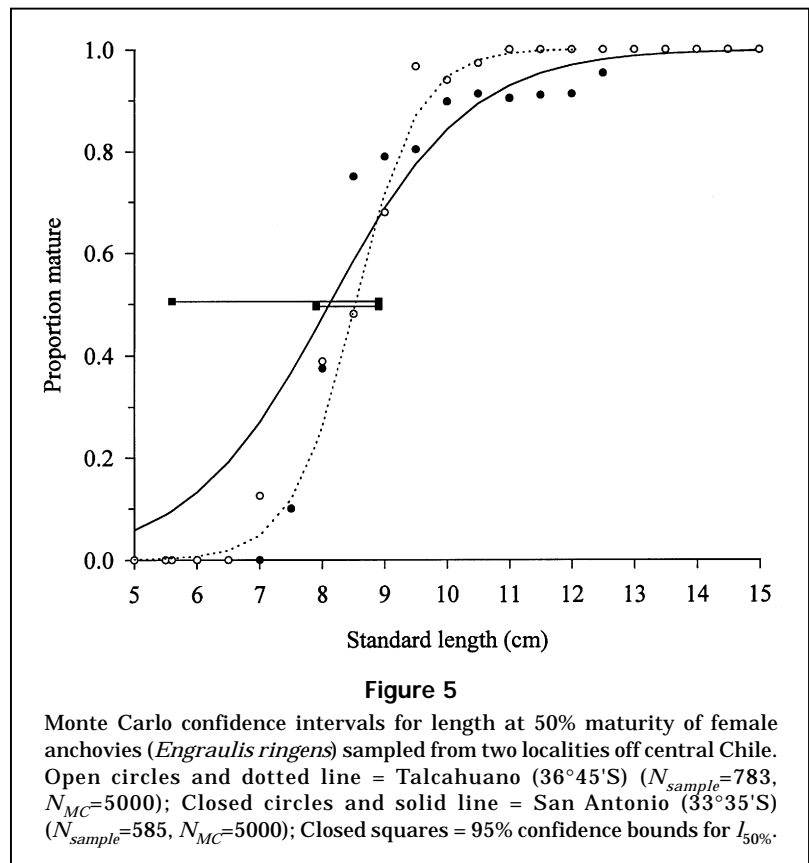
Bootstrap single assumption was that all observations from any given sample have the same probability to appear in a new sample. In contrast, the Monte Carlo method assumed a bivariate normal distribution of parameter estimates of the maturity model. Having simpler assumptions, it is unclear why the bootstrap method failed more than the nominal 5% of the times at low sample sizes. One reasonable explanation is that for every sample, there would be N_{boot} bootstrap samples, and therefore N_{boot} numerical solutions to the normal equations under the binomial likelihood model. We used here 5000 bootstrap samples and the SIMPLEX algorithm (Press et al., 1992). Small errors in the numerical algorithm coupled with a minimum bias, may add to the nominal 5%, accounting for the 1% to 2% increase in failure rate. In contrast, the Monte Carlo algorithm requires a single numerical solution so that it does not accumulate numerical errors. On the other hand, Fieller's analytical method requires a more complicated set of assumptions than the Monte Carlo approach. Fieller's method requires normality of a lin-



ear combination of parameter estimates and therefore assumes a symmetric interval estimate. Both bootstrap and Monte Carlo results indicate that the interval estimate can be quite asymmetric. This may account for the better performance of Monte Carlo, as compared with Fieller's, in proportion of success. These remarks, along with the facts that Monte Carlo's percentage success and bias are not affected by a natural mortality rate varying between 0.2 and 0.8, and that it is very fast on every computer platform, allows us to recommend the use of the Monte Carlo method to estimate size at $P\%$ maturity.

Acknowledgments

We would like to thank Bryan F. J. Manly for reviewing an earlier draft of this work and three anonymous reviewers who made several comments and criticisms which greatly improved the manuscript. This work was funded by FONDECYT grant 1950090 to R.R.



Literature cited

- Armstrong, M. P., J. A. Musick, and J. A. Convocoresses.**
1992. Age, growth, and reproduction of the goosefish *Lophius americanus* (Pisces: Lophiiformes). *Fish. Bull.* 90:217–230.
- Campbell, A., and M. D. Eagles.**
1983. Size at maturity and fecundity of rock crabs, *Cancer irroratus*, from the Bay of Fundy and southwestern Nova Scotia. *Fish. Bull.* 81(2):357–362.
- Campbell, A., and D. G. Robinson.**
1983. Reproductive potential of three American lobster (*Homarus americanus*) stocks in the Canadian maritimes. *Can. J. Fish. Aquat. Sci.* 40:1958–1967.
- Casella, G., and R. L. Berger.**
1990. Statistical inference. Brooks and Cole Publ. Co., Duxbury, California. 650 p.
- Chambers, J. M.**
1977. Computational methods for data analysis. John Wiley and Sons, New York, NY, 268 p.
- Comeau, M., and G. Y. Conan.**
1992. Morphometry and gonad maturity of male snow crab, *Chionoecetes opilio*. *Can. J. Fish. Aquat. Sci.*, 49:2460–2468.
- Draper, N. R., and H. Smith.**
1981. Applied regression analysis. John Wiley and Sons, New York, NY, 736 p.
- Efron, B., and R. J. Tibshirani.**
1993. An introduction to the bootstrap. Chapman & Hall, New York, NY, 436 p.
- Fieller, E. C.**
1944. A fundamental formula in the statistics of biological assay, and some applications. *Quart. J. Pharm.* 17:117–123.
- Finney, D. J.**
1977. Probit analysis, 3rd ed. Cambridge Univ. Press, New York, NY, 333 p.
- Gaertner, D., and F. Laloé.**
1986. Étude biométrique de la taille à première maturité sexuelle de *Geryon maritae* Manning et Holthuis, 1981 du Sénégal. *Oceanol. Acta* 9:479–487.
- González-Gurriarán, E., and J. Freire.**
1994. Sexual maturity in the velvet swimming crab *Necora puber* (Brachyura, Portunidae): morphometric and reproductive analyses. *ICES J. Mar. Sci.*, 51:133–145.
- Hosmer, D. W., and S. Lemeshow.**
1989. Applied logistic regression. John Wiley and Sons, New York, NY, 328 p.
- Johansen, S.**
1984. Functional relations, random coefficients, and non-linear regression with application to kinetic data. Springer-Verlag, New York, NY, 126 p.
- Leslie, P. H., J. S. Perry, and J. S. Watson.**
1945. The determination of the median body-weight at which female rats reach maturity. *Proc. Zool. Soc. Lond.* 115:473–488.
- Lovrich, G. A., and J. H. Vinuesa.**
1993. Reproductive biology of the false southern king crab (*Paralomis granulosa*, Lithodidae) in the Beagle Channel, Argentina. *Fish. Bull.* 91:664–675.
- Manly, B. F. J.**
1997. Randomization, bootstrap and Monte Carlo methods in biology, 2nd ed. Chapman & Hall, London, 399 p.
- Microsoft Corp.**
1995. FORTRAN PowerStation programmer's guide. Printed in the USA, 752 p.
- McCullagh, P., and J. A. Nelder.**
1989. Generalized linear models. Chapman & Hall, London. 511 p.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery.**
1992. Numerical recipes in FORTRAN, 2nd ed. Cambridge Univ. Press, Cambridge, 963 p.
- Roa, R.**
1993a. Annual growth and maturity function of the squat lobster *Pleuroncodes monodon* in central Chile. *Mar. Ecol. Prog. Ser.* 97:157–166.
1993b. Análisis metodológico pesquería de langostino colorado. Tech. Rep., Instituto de Fomento Pesquero (IFOP), Chile. 86 p. [Available from Subsecretaría de Pesca, Bellavista 168, piso 19, Valparaíso, Chile.]
- Shanubhogue, A., and P. A. Gore.**
1987. Using logistic regression in ecology. *Curr. Sci.* 56: 933–936.
- Smith, S. J.**
1990. Use of statistical models for the estimation of abundance from groundfish trawl survey data. *Can. J. Fish. Aquat. Sci.* 47:894–903.
- Somerton, D. A.**
1980. A computer technique for estimating the size of sexual maturity in crabs. *Can. J. Fish. Aquat. Sci.*, 37:1488–1494.
- Somerton, D. A., and R. A. MacIntosh.**
1983. The size at sexual maturity of blue king crab, *Paralithodes platypus*, in Alaska. *Fish. Bull.* 81(3):621–628.
- Somerton, D. A., and R. S. Otto.**
1986. Distribution and reproductive biology of the golden king crab, *Lithodes aequispina*, in the eastern Bering Sea. *Fish. Bull.* 84(3):571–584.
- Thompson, S. K.**
1992. Sampling. John Wiley & Sons, New York, NY, 343 p.
- Welch, D. W., and R. P. Foucher.**
1988. A maximum likelihood methodology for estimating length-at-maturity with application to Pacific cod (*Gadus macrocephalus*) population dynamics. *Can. J. Fish. Aquat. Sci.* 45:333–343.