

## Measuring and Correcting for Size Selection in Electrofishing Mark–Recapture Experiments

CHARLES S. ANDERSON

Minnesota Department of Natural Resources  
500 Lafayette Road, St. Paul, Minnesota 55155-4012, USA

**Abstract.**—In electrofishing, the form of size selectivity is unknown, so parametric models of selection have not been developed. Capture probability can be estimated as a continuous function of fish size for two-sample mark–recapture experiments such as those often made to describe stream fish populations. One approach is based on nonparametric regression by smoothing splines, in which general linear models are fitted to data by penalized maximum-likelihood methods. Capture probabilities modeled this way allow the data to reveal the form of size selectivity and can be used to estimate population size. A confidence band around the estimated capture probability function and confidence intervals for the population estimate are obtained by bootstrapping. A second approach is based on parametric regression models and requires some knowledge of the form of the selection function. A confidence interval for the population estimate is based on the variance estimated by the delta method. These approaches are illustrated by analysis of a mark–recapture experiment on brown trout *Salmo trutta* in a Minnesota stream, and results are compared with those found in application of Chapman–Petersen estimates to size-groups. The biases of estimates of population size  $N$  and of SE ( $\hat{N}$ ) by these approaches were examined by using Monte Carlo simulations of known populations. For the size-group in which capture probability changed rapidly with size and for the total population, spline-based population estimates typically had less bias and better coverage than Chapman–Petersen estimates, and they typically had a smaller SE ( $\hat{N}$ ) and root mean squared error than parametric model estimates.

Because electrofishing usually captures a disproportionately large number of larger fish, important information about younger, smaller, or slower-growing individuals is missed, and estimates of population size, length distributions, growth rates, and production may be biased (Reynolds 1983). Electrofishing is widely used to estimate fish populations and describe their size structure; however, the size selectivity of electrofishing has not been examined as extensively as that of gill nets (Hamley 1975; Boy and Crivelli 1988), trawls (Pope et al. 1975; Cadigan and Millar 1992), or hooks (Pope et al. 1975; Leclerc and Power 1980). Although gill-net selectivity is commonly described by normal or lognormal functions, trawl selectivity by a logistic function, and hook selectivity by a second-order polynomial, no such generalities exist for electrofishing selectivity. This disparity exists because the function relating electrofishing capture probability to fish size varies from sample to sample, and the most appropriate model is unknown. Electrofishing capture probability is the product of several conditional probabilities (i.e., that a fish is encountered; that once encountered, taxis or narcosis is induced; that once incapacitated, the fish is seen; and that once seen, the fish is netted). Each conditional probability may change between samples; therefore, an ex-

tremely flexible class of models is essential for identifying the form of size selection when estimates of demographic parameters are corrected.

Constant capture probability is a basic assumption of the Petersen estimate for two-sample mark–recapture experiments, and Seber (1982) described two tests of whether capture probability is independent of size (chi-square test and Mann–Whitney rank-sum test). When capture probability changes continuously with size, the inappropriate application of the Petersen estimate will yield biased results that underestimate population size,  $N$ . This has led several authors to suggest data be subdivided into length-groups to reduce the bias introduced by unequal capture probabilities (Sullivan 1956; Vincent 1971; Bohlin and Sundstrom 1977). Population estimates may be made separately for several size-groups within each of which capture probabilities are assumed constant, and the results may be summed to obtain an estimate of the total population. One weakness of this approach is that the estimates  $\hat{N}$  and  $\hat{V}(\hat{N})$  change erratically, depending on which size-groups are defined, especially when the numbers of individuals within size-groups are small. Thus a trade-off between reducing bias and increasing variance may be obscure. Moreover, the choice of groups is arbitrary. Therefore, additional methods of es-

timating capture probabilities and population size are needed. Restriction of analyses to a portion of the size range in which capture probabilities are approximately constant avoids bias, but information on other sizes is lost.

The nested hierarchy of maximum-likelihood mark-recapture models developed by Otis et al. (1978) allows tests of equal catchability over time, of a behavioral response, or of unequal catchability among individuals. The tests are rarely used in fisheries studies, however, because they typically require that more than two sampling efforts be made and that capture histories be identifiable. Miller (1988) extended this family of models to some situations in which capture probability could be described by a three-parameter logistic model. He derived analytic expressions for size-group estimates and their variances; however, no program is available to calculate covariance terms essential for estimating the variance of the total population estimate, and some grouping by size is necessary.

In this paper, I illustrate how, under the assumption that capture probability is a smoothly changing function of size, the relation between capture probability and size can be estimated by using smoothing splines. This flexible class of models provides a useful tool, even if it is only used to guide the choice of size-groups or parametric models. Then I calculate several population estimates based on smoothed capture probability functions. The simplest estimate (denoted  $M_z$ ) is a specific case of model  $M_h$  of Otis et al. (1978), in which capture probability is heterogeneous among individuals but determined by size,  $z$ . The smoothed model  $M_z$  is fitted to data from a two-sample mark-recapture experiment on brown trout *Salmo trutta*. I compare the output of this model with those of model  $M_o$  (constant capture probability), model  $M_{tz}$  (capture probability function differing among sample times), and model  $M_{bz}$  (depletion rate a function of size). Lastly, I compare the bias and root mean squared error of population estimates based on smoothing spline models  $M_z$  and  $M_{tz}$  with those obtained from a parametric regression model of capture probability or from Chapman-Petersen methods by means of Monte Carlo simulations of known populations with characteristics similar to those of brown trout in a Minnesota stream.

## Methods

### Nonparametric Regression with Smoothing Splines

Application of smoothing splines to binary data by penalized maximum-likelihood methods is the

nonparametric equivalent of logistic regression. The strength of the approach is its versatility in cases where the parametric form of the model is unknown. If one can assume the function is smooth, the data reveal the shape of the function and the amount of smoothing required (Yandell and Hogg 1988). The word "smooth" has a technical meaning of some importance, for the definition imposes restrictions on the class of admissible functions. A function is smooth if it and its first two derivatives are continuous; this definition of smooth matches well with common usage.

Schluter (1988) showed that the use of linear and quadratic models could lead to incorrect conclusions about whether natural selection was directional, stabilizing, or disruptive, and that nonparametric splines were better able to reveal local features of the fitness function. A smoothing spline behaves like a weighted moving average where the width of the window varies with the local density of the data points (Silverman 1984, 1985). Smoothing splines have been applied to describe the risk of heart disease, the spatial distribution of a potato virus in a field (O'Sullivan et al. 1986), individual variation of age-specific natality of leafhoppers (Yandell and Hogg 1988), and natural selection on phenotypic traits (Schluter 1988; Schluter and Nychka 1994). These studies and the book by Eubank (1988) provide a general introduction to smoothing splines.

An estimate of capture probability,  $P(z)$ , as a function of fish size,  $z$ , can be obtained by fitting the smoothing spline to data from two-sample mark-recapture experiments. The data are individual fish size,  $z_i$ , for each fish captured in a particular sample and a binary code,  $R_i$ , showing whether the fish was captured in the other sample (details on the coding for two study designs are presented below). Directly maximizing the log likelihood

$$L = \sum_{i=1}^S \{R_i \log_e [P(z_i)] + (1 - R_i) \log_e [1 - P(z_i)]\}, \quad (1)$$

$P(z_i)$  being the probability that the capture history code is  $R_i$  for a given  $z_i$ , would yield estimates of  $P(z_i)$  equal to the observed  $R_i$ . Maximizing the likelihood does not provide a very appealing estimator of the function  $P(z)$ ; indeed, it does not even provide a unique estimator of  $P(z)$ . Letting  $p(z)$  denote the observed recapture proportion among fish of size  $z$ , and letting  $z_1, z_2, \dots, z_s$  denote the distinct sizes of fish in the original sam-

ple, it is easily seen that any function passing through the points  $(z_i, p(z_i))$ ,  $i = 1, 2, \dots, s$ , is a maximum-likelihood estimator. If the sizes of different fish are distinct, then  $p(z_i)$  will either be zero or one; the estimator of  $P(z)$  will be very rough and have little predictive value.

Thus, one needs to choose an optimality criterion that considers not only the value of the likelihood but also the smoothness of the estimator. A prior idea of smoothness is incorporated by measuring roughness based on the curvature of the logit function  $f(z) = \log_e\{P(z)/[1 - P(z)]\}$ . The penalized maximum log likelihood is

$$L(f, \lambda) = L - s\lambda J(f); \quad (2)$$

here  $J(f)$  is the summed squared curvature of  $f$ ,

$$J(f) = \int [f''(z)]^2 dz, \quad (3)$$

which provides a measure of roughness;  $\lambda$  is a smoothing parameter that penalizes for roughness, and  $s$  is the number of distinct fish lengths.

Silverman (1985) noted the conceptual difficulty of understanding that the smoothed curve is defined implicitly as a solution to a minimization problem rather than as an explicit formula involving the data values. He showed that for large data sets, the smoothing spline is approximately a kernel method (a weighted moving average) in which the local bandwidth is inversely related to the local density of the data. Silverman (1984, 1985) argued that the relationship to local density is very close to the theoretically ideal relationship and illustrated the effective kernel function.

One useful property implied by the second-order derivative used to formulate roughness  $J(f)$  is that the smoothing spline function  $f(z)$  is a cubic polynomial in each interval between nodes. The four coefficients of the polynomial in each interval are not of particular interest in themselves, but they allow the spline to be estimated by solving a band-limited linear system of size  $s$  (Silverman 1985; Eubank 1988). The choice of a roughness penalty and the criterion of penalized maximum likelihood are restrictions on the class of admissible functions. When the penalized log likelihood is maximized for fixed  $\lambda$ , the resulting function is also smooth, owing to continuity of the first and second derivatives at the nodes. However, the choice of the penalty for curvature,  $\lambda$ , determines how sharply the function bends. Small  $\lambda$  allows the estimate to be rough, and large  $\lambda$  forces the estimate to approach a logistic model as determined by the prior logit function (O'Sullivan et al. 1986). Therefore, the choice of  $\lambda$  is important. To max-

imize the predictive value of  $\hat{P}(z)$ , a value for  $\lambda$  is chosen by generalized cross-validation (GCV; Craven and Wahba 1979), which compares each observation with the curve estimated by excluding that observation (Yandell and Hogg 1988). An iterative search finds the value of  $\lambda$  that minimizes the GCV score and approximately minimizes the sum of squared prediction errors. The GCV criterion is particularly appropriate when the smoothed curve is viewed as a predictive equation (Seber and Wild 1989). The smoothing spline function,  $f(z)$ , is then reestimated by maximizing the penalized log likelihood with the new value of  $\lambda$ , and  $P(z)$  is calculated from  $f(z)$ .

Theory suggests the "effective" number of parameters necessary to reproduce the curve is measured by the trace of the  $s \times s$  "hat" matrix that maps the  $R_i$  onto the predicted values (Schluter and Nychka 1994). If  $\lambda$  is large and the estimate has a nearly linear logit function, then the effective number of parameters will be two; if the estimate is more complex, the effective number of parameters will be larger. The degrees of freedom  $\nu$  of the function is  $s$  minus the effective number of parameters. The effective number of parameters and the degrees of freedom are usually not integers, and tests based on them must be interpreted with caution. The deviance statistic,  $D$ , a likelihood ratio of a saturated model and the model of interest, provides one measure of overall goodness of fit. It can seldom be tested directly because the number of individuals is usually not sufficiently greater than  $s$  that  $m$ -asymptotics can be used (McCullagh and Nelder 1983), but change-of-deviance statistics (also likelihood ratios) are useful to test a full model against a reduced model (Yandell and Green 1986; Yandell and Hogg 1988). Schluter and Nychka (1994) used a bootstrap comparison of log-likelihood values to test the fit of a full model against a reduced one. I present both tests because they are approximate and the sampling distribution of smoothing spline estimates is unknown. Goodness of fit may also be assessed by examination of deviance residuals and leverages (Yandell and Green 1986).

### Electrofishing

Mark-recapture experiments based on two electrofishing samples are done routinely to estimate stream fish populations. The data analyzed here were collected by Thorn (1992) as part of a study on West Indian Creek, a spring-fed stream in southeast Minnesota with a naturalized brown trout population. The study reach was 0.77 km

long and had a mean width of 5.2 m, a gradient of 2.8 m/km, a late-summer discharge of 0.21 m<sup>3</sup>/s, and an alkalinity (as CaCO<sub>3</sub>) of 220–250 mg/L. Habitat was improved in this reach in 1981 by installation of intensive overhead bank cover (Thorn 1988, 1992). In September 1988, the reach was electrofished by three people with pulsed DC power from a generator pulled in a small boat. The gear and procedures were like those described by Novotny and Priegel (1974). Brown trout captured in the first sampling effort were anesthetized with tricaine methanesulfonate (MS-222) and measured for total length (mm). The tops of their caudal fins were clipped to provide a recognizable batch mark, and the fish were released near the pool from which they had been captured. The same stream reach was similarly electrofished 3 d later, and the total length and presence or absence of a mark was recorded for each captured individual.

### Mark-Recapture Designs

If released individuals are uniquely marked, the data may be analyzed as a longitudinal study in which the recaptures from a known population provide information on capture probabilities for the second sampling effort. For this analysis, all individuals with a capture history of 11 (caught in both samples) are coded  $R_i = 1$ , and individuals with history 10 (caught in the first but not in the second sample) are coded  $R_i = 0$ . The capture probability function is fitted to individuals with these two histories, and an estimate of model  $M_z$  capture probability for the second sample is obtained.

Estimates of capture probabilities for the first sampling effort may be similarly obtained by examining the proportion of fish in the second sample that were previously marked. For this design, individuals with a capture history of 11 are coded  $R_i = 1$ , and those with history 01 (caught in the second but not in the first sample) are coded  $R_i = 0$ . The capture probability function is fitted to individuals with these two histories, and an estimate of model  $M_z$  capture probability for the first sample is obtained. This analysis, which treats the data as a single cross-sectional sample, is possible if marks are not uniquely identifiable, as in batch marking, or if marks are unique.

If fish are individually marked, the longitudinal and cross-sectional designs produce the separate functions of model  $M_{tz}$ . The log likelihood of this model is the sum of the likelihoods of the separate functions. The two coded data sets can be appended, and an estimate of model  $M_z$  with the same

capture function for each sample can be obtained. For each of the models described, a corresponding model  $M_o$  (capture probability independent of length) can be estimated directly from the same coded data set. In addition, log-likelihood and deviance values (McCullagh and Nelder 1983) can be calculated for each model.

Considering marked individuals as though removed from the population allows one to fit model  $M_{bz}$  as a depletion estimate. Model  $M_{bz}$  is appropriate only when there are concerns about a change in capture probability resulting from the first capture event; otherwise, fitting  $M_{bz}$  is an inefficient means of estimating the parameters of interest. It also requires an assumption that capture functions for unmarked fish were the same for each sample. Individuals with a capture history of 01 are coded  $R_i = 1$  and individuals with histories of 11 or 10 are coded  $R_i = 0$ . The smoothed function  $\hat{A}(z)$  is an estimate of the conditional probability that an individual of size  $z$  was first captured in the second sample, given that an individual that size was captured in either sample. If capture probabilities are high,  $\hat{A}(z)$  will be near zero. If capture probabilities are low, similar numbers will be first captured in each sample, and  $\hat{A}(z)$  will approach 0.5; larger values would indicate there was no depletion and that the additional assumption of equal sampling effort was violated. Log-likelihood and deviance values are calculated for this function. The corresponding model  $M_b$  (depletion independent of length) can be estimated directly from the coded data set. Capture probabilities of models  $M_{bz}$  and  $M_b$  are estimated from the conditional probabilities by  $\hat{P}(z) = 1 - \{\hat{A}(z)/[1 - \hat{A}(z)]\}$ .

### Demographic Estimates

For longitudinal and cross-sectional analyses of model  $M_z$ , and for Model  $M_{bz}$ , the size distribution of fish taken in one sample is corrected by weighting each fish by  $1/\hat{P}(z)$ , where  $\hat{P}(z)$  has been calculated for that sample by the smoothing spline or parametric regression model. The population estimate at each size,  $\hat{N}(z)$ , is therefore based on the number of fish of that size in one sample,  $n(z)$ , weighted by  $1/\hat{P}(z)$ :  $\hat{N}(z) = n(z)/\hat{P}(z)$ . The estimate of the total population is

$$\hat{N} = \sum \hat{N}(z) = \sum [n(z)/\hat{P}(z)]; \quad (4)$$

the summation is across the size range of interest. For model  $M_{tz}$ , the population estimate is the mean of the longitudinal and cross-sectional estimates. Population estimates and confidence intervals are rounded to integer values.

Capture probabilities,  $P(z)$ , based on the smoothing spline were estimated for the various models with a modified version of Schluter's (1988) program run on a PC-compatible computer. If some sizes are not represented in both samples, the smoothing spline algorithm does not directly provide capture probabilities for every size in the data. The missing estimates were obtained by interpolation from the logits of spline estimates at adjacent sizes. Estimates of  $P(z)$  were also made by fitting the logistic model  $P(z) = \beta_2 \cdot \exp(\beta_0 + \beta_1 z) / [1 + \exp(\beta_0 + \beta_1 z)]$  by using maximum-likelihood methods (SYSTAT; Wilkinson 1990).

#### Bootstrap Tests and Confidence Intervals

A test of whether model  $M_z$  provides a significantly better fit than model  $M_o$  is done by bootstrapping according to the procedure described by Schluter and Nychka (1994). The goodness of fit of model  $M_z$ ,  $L(\hat{P}, M_z)$ , is compared to the frequency distribution of values produced when model  $M_z$  is fitted to bootstrap samples generated with a constant capture probability,  $P_o$ . The procedure consists of three steps: (1)  $P_o$  is estimated from model  $M_o$ ; (2) a bootstrap sample is generated by assigning each individual in the coded data set a new capture history code, where the probability of assigning  $R_i = 1$  is  $\hat{P}_o$ ; and (3) model  $M_z$  and its fit  $L(\hat{P}, M_z)$  are estimated. Steps 2 and 3 are repeated 1,000 times. The simpler model  $M_o$  is rejected if the original log likelihood of model  $M_z$  is large relative to the distribution of bootstrap log-likelihood estimates (the percentile method; Efron and Tibshirani 1986). A similar procedure is followed to test whether model  $M_{bz}$  provides a significantly better fit than model  $M_b$ , except  $\hat{H}_b$  replaces  $\hat{P}_o$  in step 2.

A test of model  $M_{tz}$  against model  $M_z$  (same function) requires a bootstrap assigning one of three capture histories to each observed individual. In step 1, models  $M_{tz}$  and  $M_z$  are estimated. The two capture functions are denoted  $\hat{P}_1(z)$  and  $\hat{P}_2(z)$  (for model  $M_z$  they are the same). Let  $\theta(z)$  be the estimated probability that an individual was observed in either sample; that is,  $\theta(z) = 1 - [1 - \hat{P}_1(z)] \cdot [1 - \hat{P}_2(z)]$ . In step 2, a bootstrap sample is generated by assigning each individual in the coded data set a new capture history; each realized set of capture histories,  $\{x_{10}, x_{11}, x_{01}\}$ , is a random sample drawn from an underlying multinomial. Let  $\{x_{10}, x_{11}, x_{01}\}$  denote the set of random variables corresponding to the observable capture histories 10, 11, and 01 of fish of length  $z$ . Then these variables are multinomially distributed

$$\{X_{10}, X_{11}, X_{01}\} \sim \text{multinomial}[n(z), \pi_1, \pi_2, \pi_3],$$

in which

$$\pi_1 = P_1(z) \cdot [1 - P_2(z)] / \theta,$$

$$\pi_2 = P_1(z) \cdot P_2(z) / \theta,$$

and

$$\pi_3 = [1 - P_1(z)] \cdot P_2(z) / \theta,$$

because the number of observed fish of length  $z$ ,  $n(z)$ , is fixed and each fish is sampled independently with probabilities  $P_1(z)$  and  $P_2(z)$  (until it is observed at least once in the experiment). When models  $M_{tz}$  and  $M_z$  are compared the capture probabilities are those of model  $M_z$ ; thus,  $P_1(z) = P_2(z)$  and  $\pi_1 = \pi_3$ . In step 3, two coded data sets are constructed from the new capture histories as described for longitudinal and cross-sectional designs; these are analyzed and the bootstrap log likelihood of model  $M_{tz}$  is recorded. Steps 2 and 3 are then repeated. Again, the simpler model is rejected if the original log likelihood of model  $M_{tz}$  is large relative to the distribution of bootstrap estimates.

Confidence intervals about the spline-based capture function and  $\hat{N}$  cannot be calculated directly, so estimates are also obtained by bootstrapping. The distributions of spline-based estimates of  $N$  are likely to be asymmetrical because of the weighting by  $1/\hat{P}(z)$ ; consequently, percentile rather than pivot intervals are desirable. To construct confidence intervals,  $b = 1,000$  bootstrap estimates of capture probabilities and  $\hat{N}$  were made. Each bootstrap sample was treated in four steps. In step 1, a new capture history code (or capture history for  $M_{tz}$ ) was randomly assigned to each individual in the sample being fit; however, the probabilities were now based on the  $\hat{P}(z)$  or  $\hat{H}(z)$  of the full model rather than the reduced model. For model  $M_{tz}$ ,  $P_1(z)$  may not equal  $P_2(z)$ . In step 2, a full model was fitted to the new data in the same way, by a grid search for  $\lambda$ . In step 3, new  $\hat{P}_b(z_i)$  values were merged with the size distribution of fish, missing  $\hat{P}_b(z_i)$  were found by interpolation, and an estimate of  $N$  was calculated. In step 4, values for  $\hat{N}$  were recorded. Steps 1 to 4 were then repeated. The program calculated the mean and standard deviation of  $\hat{P}_b(z)$  and the weights ( $1/\hat{P}_b(z)$ ) directly during this cycling. The mean and variance of the bootstrap  $\hat{N}$  values were calculated, and the 95% confidence intervals were obtained directly from the tails of the distribution (SYSTAT; Wilkinson 1990).

Confidence intervals for the parametric-model-based estimate  $\hat{N}$  were pivot intervals based on a

variance obtained by the delta method and use of the estimated covariance matrix of the parameter estimates (Seber 1982; Hoenig et al. 1989).

Simulations

The accuracy and precision of the estimates of  $N$  obtained with smoothing splines can only be evaluated by applying the method when the population size is known. Monte Carlo simulations were therefore made of mark-recapture experiments on populations of known size, for which the fish length distributions and capture probabilities approximated those of brown trout in West Indian Creek. Specifically, the populations comprised equal numbers of individuals in two age-groups; age-1 fish were drawn from a normal length distribution with a mean of 140 mm and an SD of 20 mm, and age-2 fish were drawn from a normal distribution with a mean of 240 mm and an SD of 30 mm. Capture probabilities in each sample were described by a three-parameter logistic function  $P(z) = \beta_2 \exp(\beta_0 + \beta_1 z) / [1 + \exp(\beta_0 + \beta_1 z)]$ , where  $\beta_0 = -4.7958$ ,  $\beta_1 = 0.0326$ , and  $\beta_2 = 0.5$  (values were chosen because they described the smoothing spline curve for West Indian Creek after an adjustment for estimated bias). Four population sizes were examined (124, 250, 500, and 1,000) and 500 simulations were made at each population size. At each population size individuals were drawn at random from the specified size distribution, and all simulations were done on this population with a fixed length distribution (results were similar when the length distribution was not fixed). In each simulation capture histories were randomly assigned, dependent on individual length. The underlying multinomial distribution of the set of random variables corresponding to the observable capture histories 10, 11, 01, and 11 of fish of length  $z$  is

$$\{X_{10}, X_{11}, X_{01}, X_{11}\} \\ \sim \text{multinomial}[N(z), \pi_1, \pi_2, \pi_3, \pi_4],$$

in which

$$\pi_1 = \pi_3 = P(z) \cdot [1 - P(z)], \\ \pi_2 = P(z)^2,$$

and

$$\pi_4 = [1 - P(z)]^2,$$

because the number of fish of length  $z$ ,  $N(z)$ , is fixed for each simulation and each fish is sampled independently with probability  $P(z)$ , which depends on length  $z$  but not on sampling occasion. Population estimates were then made by the Chap-

man-Petersen, smoothing spline model  $M_z$  (batch marked), smoothing spline model  $M_{tz}$ , and parametric regression methods. All the estimates were discarded if the number of recaptures in the size-group was less than four (Ricker 1975) or if any capture probability estimates were less than 0.02. Preliminary inspections of  $\hat{N}$  distributions obtained by smoothing spline or parametric regression methods showed that large outliers appeared when there were few recaptures at the smallest sizes and capture probability estimates were less than 0.02. In real applications, one would first truncate the size range over which  $N$  is estimated rather than discard the experiment. However, such truncation is not done here because it would preclude comparison of the three methods on the same simulations.

Chapman-Petersen estimates of  $N$  and  $SE(\hat{N})$  were made for the total population (size ignored) for the two age-groups separately (individuals shorter than 180 mm were assigned to age-1), and for the total population as the sum of age-group estimates.

Each smoothing spline was fitted after a grid search for  $\lambda$ , results were merged with the size distribution of fish in the appropriate sample, missing  $\hat{P}(z)$  values were found by interpolation, and population estimates were calculated by equation (4). Estimates for each age-group and the total population were made by summing over the appropriate length interval. In addition, the data of each simulation were bootstrapped 20 times for each smoothed model, and the SD of the resulting population estimates was calculated as an estimate of the precision of the Monte Carlo  $\hat{N}$ . For each population estimate, the 95% pivot confidence interval was calculated. Obtaining percentile intervals by bootstrapping many times would make simulations prohibitively slow.

A three-parameter logistic model was fitted by maximum-likelihood methods with a Marquardt algorithm, and covariances of the parameter estimates were found by inverting the negative of the Hessian matrix. The program was written by following guidelines of Dennis and Schnabel (1983) and Thisted (1988). Capture probabilities were estimated from the model for each marked fish, and population estimates were calculated by equation (4). For each population estimate, a variance was estimated by the delta method, and the 95% confidence interval was calculated.

Coverage, bias, and root mean squared error (RMSE) were estimated for each method. Coverage is the percentage of simulations for which the

TABLE 1.—Brown trout population estimates made by Chapman–Petersen, nonparametric spline, and parametric regression methods. Various numbers of size-groups were used in calculating Chapman–Petersen estimates. Smoothing spline and parametric regression methods calculate a separate weighing variable for every size in one sample (158 weights when analyzed for batch-marked fish, 154 when analyzed for individually marked fish). Variances of the spline estimates were calculated by bootstrapping 1,000 times. Variances of the parametric regression estimates were found by the delta method.

Method and type of marking	Number of size- groups	Popula- tion ( $\hat{N}$ )	Confidence interval (95%)	Variance ( $\hat{V}(\hat{N})$ )
Chapman–Petersen	1	1,023	914–1,132	3,092
	2	1,030	918–1,141	3,212
	4	1,158	972–1,344	8,991
	6	1,144	966–1,322	8,273
	8	1,149	973–1,326	8,116
	10	1,147	964–1,330	8,734
	12	1,130	950–1,310	8,431
	14	1,207	955–1,459	16,564
Spline–bootstrap				
Batch mark, $M_z$	158	1,137	994–1,334	8,882
Individual mark, $M_z$	154	1,139	1,001–1,345	7,688
$M_{Iz}$	179	1,138	1,030–1,279	4,338
Parametric model				
Batch mark	158	1,236	1,109–1,363	4,187
Individual mark	154	1,188	985–1,391	10,747

estimated 95% confidence interval included the true population size. Coverage depends on the accuracy with which both  $N$  and  $SE(\hat{N})$  are estimated. Differences between the means of the ( $\sim 500$ ) Monte Carlo population estimates and the known population sizes were calculated to measure the bias (inaccuracy) of each estimation method. Differences between the mean of the ( $\sim 500$ ) estimates of  $SE(\hat{N})$  and the actual SD of the Monte Carlo population estimates  $\hat{N}$  were calculated to measure the bias in estimates of sampling precision. The SD of the Monte Carlo population estimates is a direct measure of  $SE(\hat{N})$ .

## Results

### Brown Trout

In West Indian Creek, 418 brown trout were marked and released ( $M$  in Ricker's 1975 notation), and 198 of them were recaptured ( $R$ ) in the second sample of 485 brown trout ( $C$ ). Total lengths ranged from 101 to 395 mm. Because the assumption of equal catchability was violated, the data was subdivided into (up to 14) length-groups. Population estimates were made for each group by the Chapman–Petersen method (Ricker 1975), and

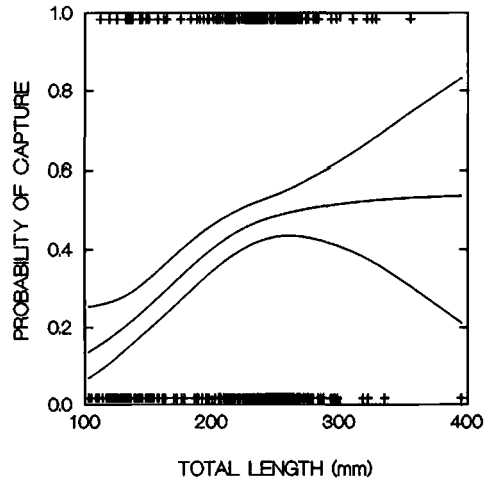


FIGURE 1.—Probability of capture in the first sample as a smoothed function of total length for brown trout sampled by electrofishing. Confidence bands show  $\pm 2$  SEs of predicted values from 1,000 bootstrap replicates of the selectivity function. Fish measured in the second sample were coded with a 1 if marked and 0 otherwise; each datum is shown by a plus sign (+) slightly offset from the border.

total population estimates were made by summing the group estimates (Seber 1982). As the number of size-groups was increased, estimates of both  $N$  and  $V(\hat{N})$  increased, though the changes became erratic because the number of fish in each group ( $M$ ,  $C$ , and  $R$ ) became small (Table 1).

Because the experimental design involved batch marking, capture probabilities for the first sampling effort were estimated by fitting the smoothing spline to data from the second sample. The estimated capture probabilities increased with fish size in a sigmoid relationship (Figure 1; deviance  $D = 191.8$ ,  $df = 151.18$ ). The null model was  $\hat{P}(z) = 0.41$  ( $D = 222.5$ ,  $df = 153$ ). Both the change in deviance (30.7,  $\sim 1.82$   $df$ ,  $P < 0.001$ ) and the bootstrap test comparing goodness-of-fit values ( $P < 0.001$ ) confirmed the size dependence of capture probability. The confidence band about the probability function was wider for the larger sizes, for which the data were sparse, than for the small sizes, for which capture probabilities were low. An analysis of a similar function by O'Sullivan et al. (1986) suggested that the bias of the capture probability estimates would be small, that the smoothed function would tend to be overly straight at sizes for which the curvature was greatest, and that the estimates would tend to be slightly high for the largest sizes. Any remaining bias was ignored in

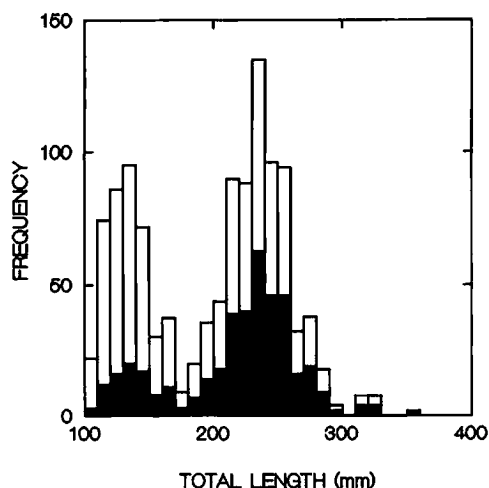


FIGURE 2.—Length distributions of brown trout in the original sample (dark bars) and after correction with the probability-of-capture function (open bars).

estimating the actual length frequency distribution from the length distribution in the first sample (Figure 2).

The coefficient of variation ( $CV = 100 \cdot SD / \text{mean}$ ) of the weighting factors ( $1/P(z)$ ) showed that a population estimate could be made with reasonable precision for all but the smallest sizes represented in the samples (Figure 3). Although inclusion of the smallest individuals would inflate the variance, population estimates were made over the entire size range. The smoothing spline yielded  $\hat{N} = 1,137$ . The variance of the bootstrap estimates was 8,882, a value considerably larger than that obtained without correcting for selectivity, yet within the wide range found by grouping by size (Table 1). The distribution of bootstrap estimates of  $N$  showed significant skewness ( $G1 = 3.311$ ,  $P < 0.001$ ) and kurtosis ( $G2 = 36.163$ ,  $P < 0.001$ ), thus the asymmetrical percentile confidence interval (CI) was appropriate ( $\hat{N} = 1,137$ ; 95% CI = 994–1,334; Figure 4). The Chapman–Petersen estimates approach those made by the spline–bootstrap method as the number of length-groups considered increased.

Although fish were similarly marked, close inspection of the data showed that measurement error was small enough that fish length could be used to identify individuals marked and released, allowing an approximate longitudinal analysis. Of the 198 recaptured fish, 158 had lengths that matched those of marked fish, 38 were within 2 mm, and 2 differed by 3 mm. After assigning each recapture to an individual in the first sample, the

data were analyzed as though fish had been individually marked. The smoothed capture probability function for the second sample had a slightly higher elevation than that for the first sample and did not appear to level off at the largest sizes ( $D = 308.32$ ,  $df = 155.68$ ; Figure 5). The null model was  $\hat{P}(z) = 0.50$  ( $D = 335.35$ ,  $df = 157$ ). Again, both the change in deviance (27.6,  $\sim 1.32$  df,  $P < 0.001$ ) and the bootstrap test of goodness of fit ( $P < 0.001$ ) confirmed size-dependent heterogeneity. The population estimate was similar (1,139 versus 1,137; Table 1), thus the model  $M_{tz}$  estimate was 1,138. Model  $M_{tz}$  provided a better fit than model  $M_z$  (same capture function for each sample) ( $P < 0.001$ ; 95% CI = 1,030–1,279).

When capture probabilities for the batch-marking design were fitted with a three-parameter logistic model, the function (Figure 5) resembled the spline-based function, and the differences were in the direction that might have resulted from bias of the smoothed function; thus, the logistic model is a reasonable choice. The parametric model yielded  $\hat{N} = 1,236$  and  $\hat{V}(\hat{N}) = 4,187$ . When the longitudinal design was similarly analyzed, the logistic capture probability function had a more sharply sigmoid shape and a smaller value for the upper asymptote than the corresponding smoothed function; thus, the logistic function appeared a poor choice. Estimates were  $\hat{N} = 1,188$  and  $\hat{V}(\hat{N}) = 10,747$  (Table 1).

Model  $M_{bz}$  is a depletion estimate of probability of capture in the first sample, because recaptured fish, which may exhibit a behavioral response, are omitted from the analysis. The function estimated by depletion was lower than the smoothed function for a batch-marking design (Figure 5). This suggests that marked fish had a higher capture probability in the second sample than unmarked fish,

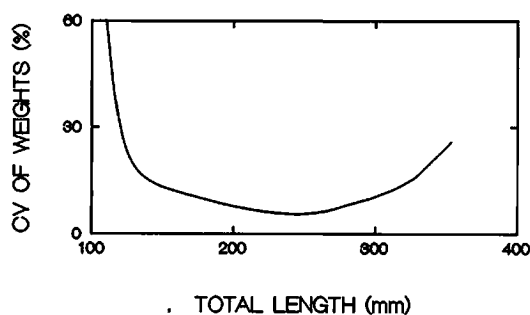


FIGURE 3.—Precision of the population estimate of fish near a specified length, as measured by the coefficient of variation ( $CV = 100 \cdot SD / \text{mean}$ ) of the weighting factors ( $1/\hat{P}(z)$ ) in 1,000 bootstrap estimates.



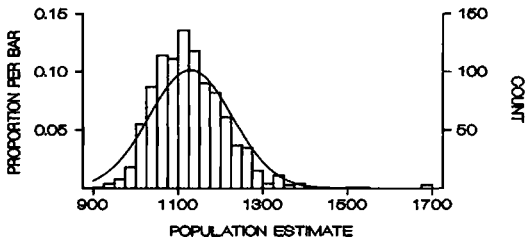


FIGURE 4.—Frequency distributions of 1,000 bootstrap estimates of  $N$ . Three outliers (1,768, 2,059, and 2,475) are grouped in the final bar.

which would have inflated the batch mark estimate. However, impossibly low depletion estimates for fish shorter than 130 mm suggest that the second sampling effort was greater or more efficient than the first and that a depletion model is inappropriate. The same low capture probability estimates preclude making an estimate of the total population.

#### Simulations

The criteria used for deciding whether a sample was adequate for analysis (four or more recaptures in the size-group; all  $\hat{P}(z) > 0.02$  in the size-group) ensured that the performances of the three methods were compared on the same simulations (Table 2). Most samples were adequate for estimating the number of age-2 fish by each method, but fewer were adequate for estimating the number of age-1 fish, especially when the true population was small. The simulations thus span the practical limits at which further increasing the number of size-groups would often simply truncate the size range over which estimates could be made. Only 226 samples had four or more recaptures of age-1 fish when the true population was 62, and only 467 samples when the population was 125. The true form of the parametric model was used in these simulations; however, in many simulations when  $N$  was 250 or less, the estimate had a sharply stair-stepped shape with many very low  $\hat{P}(z)$  values. The global fit to the data resulted in some inaccurate estimates when there were more than four recaptures and some accurate estimates even when there were few recaptures. In field experiments, model specification errors would also limit success in estimating  $\hat{P}(z)$  at extreme sizes. The behavior of the spline-based model as a local average made it less sensitive to the influence of distant points that would force very small  $\hat{P}(z)$  estimates, and fewer simulations resulted in absurd population estimates.

The bias of  $\hat{N}$ , bias of  $\widehat{SE}(\hat{N})$ , and resulting coverage and RMSE of each method are summarized for the same simulations in Table 2. The spline-based method performed better than Chapman–Petersen or parametric regression methods for estimating age-1 fish. Chapman–Petersen population estimates had a negative bias and the 95% confidence intervals often failed to include the true value. Smoothing spline population estimates had a smaller negative bias, a wider dispersion, better coverage, and a similar RMSE. Parametric model population estimates had the least bias (usually positive), but the population estimates had the widest dispersion and the dispersion was estimated with the greatest bias, resulting in poor coverage and high RMSE. For estimating age-2 fish, each method had a smaller bias and smaller RMSE. The Chapman–Petersen method underestimated  $SE(\hat{N})$  and had slightly low coverage. The smoothing spline model  $M_z$  and parametric regression methods overestimated  $SE(\hat{N})$  (the parametric model greatly so) and had high coverages. The smoothed model  $M_{tz}$  consistently yielded less-biased estimates of  $SE(\hat{N})$  and had lower coverage than the cross-sectional  $M_z$ .

For making total population estimates, the biases of each smoothing spline  $\hat{N}$  and  $SE(\hat{N})$  generally fell between those of the other methods.

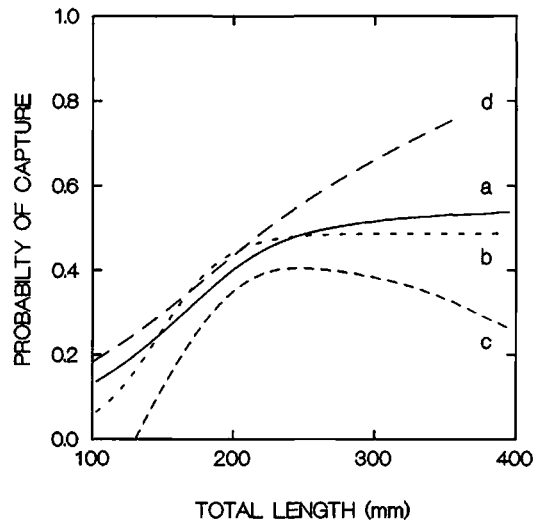


FIGURE 5.—Probability of capture in the first sample (a, b, and c) and in the second sample (d). Capture probability was estimated by smoothing spline and a batch-marked design (a), logistic model and a batch-marking design (b), smoothing spline and a depletion model (c), and smoothing spline and an individual marking design (d).

TABLE 2.—Bias of  $\hat{N}$ , bias of  $SE(\hat{N})$ , coverage, and root mean squared error (RMSE) in application of Chapman–Petersen (CP), smoothing spline batch-mark model  $M_z$  ( $SS_z$ ), smoothing spline model  $M_{Iz}$  ( $SS_{Iz}$ ), and parametric model (PM) methods to the same samples. Four levels of the true population size were examined. Populations were composed of equal numbers of age-1 (mean total length, 140 mm; SD, 20) and age-2 (mean total length, 240 mm; SD, 30) fish. Coverage is the percent of simulations for which the estimated 95% confidence interval covered the true population size.

N	Method	Number of samples fitted	Percent bias of			Cover-age	RMSE
			SE( $\hat{N}$ )	$\hat{N}$	$\widehat{SE}(\hat{N})$		
$\hat{N}$ total fish (age-groups ignored)							
124	CP	279	14.9	-13.8	-3.0	65.2	22.7
	SS <sub>z</sub>		32.0	1.3	41.9	98.9	32.0
	SS <sub>Iz</sub>		30.7	1.2	6.1	90.7	30.7
	PM		50.6	9.4	718.7	88.2	51.9
250	CP	343	23.0	-13.1	-5.5	59.2	40.0
	SS <sub>z</sub>		40.5	-0.9	44.3	96.2	40.5
	SS <sub>Iz</sub>		40.8	-0.9	4.3	88.0	40.8
	PM		53.9	5.5	459.3	91.0	55.6
500	CP	394	33.0	-14.3	-7.0	35.5	78.9
	SS <sub>z</sub>		55.7	-3.4	26.4	92.1	58.1
	SS <sub>Iz</sub>		55.7	-3.4	-0.4	82.0	58.1
	PM		78.5	3.2	72.1	83.2	80.1
1,000	CP	418	44.9	-13.8	-0.7	17.5	144.8
	SS <sub>z</sub>		76.6	-3.1	21.6	91.6	82.6
	SS <sub>Iz</sub>		76.3	-3.1	-5.8	83.5	82.2
	PM		108.3	2.6	20.4	85.1	111.3
$\hat{N}$ total fish = $\hat{N}$ age-1 + $\hat{N}$ age-2 fish							
124	CP	146	14.6	-14.0	-1.0	65.8	22.7
	SS <sub>z</sub>		17.4	-8.5	111.9	98.6	20.3
	SS <sub>Iz</sub>		17.4	-8.5	54.7	87.0	20.3
	PM		20.3	-5.4	2565.4	95.2	21.3
250	CP	329	29.9	-6.6	3.8	79.9	34.0
	SS <sub>z</sub>		37.6	-1.8	52.6	96.4	37.8
	SS <sub>Iz</sub>		38.5	-1.8	7.0	87.5	38.7
	PM		51.0	-4.3	507.6	91.8	52.2
500	CP	394	51.6	-5.5	-3.4	82.5	58.3
	SS <sub>z</sub>		55.7	-3.4	26.4	92.1	58.1
	SS <sub>Iz</sub>		55.7	-3.3	-0.4	82.0	58.1
	PM		78.5	3.2	72.1	83.2	80.1
1,000	CP	418	72.7	-4.4	1.3	83.0	85.1
	SS <sub>z</sub>		76.6	-3.1	21.6	91.6	82.6
	SS <sub>Iz</sub>		76.3	-3.1	-5.8	83.5	82.2
	PM		108.3	2.6	20.4	85.2	111.3
$\hat{N}$ age-1 fish							
62	CP	146	10.9	-28.6	2.6	50.7	20.8
	SS <sub>z</sub>		13.8	-19.7	125.1	90.4	18.4
	SS <sub>Iz</sub>		13.5	-20.1	63.5	80.8	18.4
	PM		18.2	-11.7	1116.1	72.6	19.5
125	CP	329	26.9	-12.0	4.3	77.5	30.7
	SS <sub>z</sub>		32.0	-6.3	51.9	91.5	33.0
	SS <sub>Iz</sub>		33.0	-6.1	8.3	83.3	33.9
	PM		48.9	8.0	164.1	67.8	49.8

TABLE 2.—Continued.

N	Method	Num- ber of sam- ples fit- ted	Percent bias of			Cover- age	RMSE
			$SE(\hat{N})$	$\hat{N}$	$SE(\hat{N})$		
250	CP	394	46.7	−10.2	−0.7	78.7	53.2
	$SS_z$		49.3	−9.1	23.3	87.6	54.2
	$SS_{Iz}$		48.8	−9.1	0.3	76.4	53.9
	PM		75.3	6.0	42.0	54.8	76.7
500	CP	418	68.3	−8.4	0.9	80.9	80.2
	$SS_z$		70.0	−8.4	15.4	85.4	81.6
	$SS_{Iz}$		69.5	−8.4	−8.6	76.3	81.1
	PM		105.2	4.7	−60.5	49.5	107.7
<b><math>\hat{N}</math> age-2 fish</b>							
62	CP	492	9.4	−0.2	−10.3	89.6	9.4
	$SS_z$		13.1	4.9	46.5	98.4	13.4
	$SS_{Iz}$		12.6	4.8	−3.9	90.2	12.9
	PM		9.8	1.7	1,096.9	98.4	9.9
125	CP	500	12.8	−0.8	−3.9	92.0	12.8
	$SS_z$		14.7	2.9	44.7	99.0	15.1
	$SS_{Iz}$		13.8	2.8	−2.9	91.8	14.2
	PM		13.7	1.2	827.2	99.0	13.8
250	CP	499	17.8	−0.4	−2.5	92.4	17.9
	$SS_z$		19.0	2.4	37.5	99.0	19.9
	$SS_{Iz}$		18.8	2.4	−12.0	88.6	19.7
	PM		18.5	0.9	320.6	99.8	18.6
500	CP	500	25.6	−0.4	−2.2	93.2	25.6
	$SS_z$		26.7	2.0	33.7	97.8	28.5
	$SS_{Iz}$		26.2	2.0	−13.1	90.2	28.0
	PM		26.3	0.6	204.8	98.8	26.4

The smoothing spline model  $M_z$  had the best coverage at all population sizes, and both spline models and relatively low RMSE values. The relative performance of the Chapman–Petersen method remained the same when the Chapman–Petersen total population estimate was calculated as the sum of the age-group estimates. This change reduced the magnitude of the negative bias of  $\hat{N}$ , improved coverage at populations of 250 or less, and reduced the RMSE, but many simulations at populations of 250 or less did not contain adequate recaptures to allow separate estimation of the age-1 group.

The similarity of the parametric model to the actual function used to generate the data suggests the parametric model should eventually yield a lower RMSE than the smoothed models when the population is increased. The parametric model produced a lower RMSE for age-2  $\hat{N}$  at total populations of 124 or more and for age-1 and total  $\hat{N}$  at total populations of 4,000 or more (additional simulations were run at  $N = 2,000, 4,000,$  and  $8,000$ ). Parametric model coverages remained near

the values presented for  $N = 1,000$  (Table 2), except the coverage for age-2 fish was 85–89%. Coverage did not approach 95% because the size range of each age-class was open at one end, allowing rare small and large individuals to influence the variance estimates, and because variance estimates are inaccurate at small  $P(z)$  values.

The difference between the mean of the smoothed bootstrap  $\hat{N}$  estimates and the original  $\hat{N}$  of each simulation had a small negative correlation with the actual error for age-1 and total population estimates. Therefore, using this difference to estimate or correct for remaining bias would make the error worse. According to Gu (1987), this additional correction by bootstrapping would be similar to measuring curvature on a squared fourth derivative. Although theoretically this should reduce bias for very large samples in simulations for which the function really is smooth on fourth derivatives, neither condition is certain in real applications.

### Discussion

The spline-based approach is well suited for identifying the form of size selectivity in simple mark–recapture experiments and can be used to reduce the associated bias in estimates of population size and length distributions. The assumptions necessary for estimating capture probabilities and population size by the spline-based approach are the same as those for the Chapman–Petersen method (Seber 1982), except that capture probability is assumed to be a smooth function rather than a constant. The assumptions of the spline-based approach and of the continuous parametric model are (1) the population is closed; (2) capture probability is a smoothly changing function of size in the sample for which the probability is estimated; (3) all fish of the same size have equal probability of capture in the first sample or in the second sample, or marked and unmarked fish of the same size mix randomly between samples; (4) marking does not affect subsequent probabilities of capture (except in models  $M_b$  and  $M_{bz}$ ); (5) marks are not lost between samples; (6) all marks of fish are reported; and (7) sizes of all fish are reported. Each estimator is based on an assumption that individual responses  $R_i$  are independent, and there may be sources of dependency that are not covered by assumptions 1–7.

In the electrofishing experiment with brown trout, the assumption of population closure could not be tested, but the short duration of the experiment (3d), the rapid recovery of fish upon release,

and the relatively long stream reach sampled suggest the assumption is reasonable. The assumption that capture probability is a smoothly changing function of size can be checked in several ways, the simplest being by inspection of the shape of the spline at several smaller values of  $\lambda$  and by inspection of deviance residuals and leverages. In addition, the bootstrap comparison of goodness-of-fit values and the change in deviance statistic allow approximate tests of whether the smoothed model is significantly better than the null model of constant probability. For parametric models, a family of models could be fitted by maximum-likelihood methods, and change-in-deviance tests could be made to identify the simplest adequate one. Violations of assumption (3) would mean there was more capture heterogeneity than could be attributed to a function of size and that individual capture probabilities may not be independent. Lack of independence is difficult to distinguish from a poor choice of capture function. Only more complex sampling designs allow detection of such heterogeneity (Otis et al. 1978); grouping by size will not remove it. The interval between samples was probably sufficient to minimize any stress-induced change in catchability (Peterson and Cederholm 1984). Assumptions (5–7) were reasonable because fish were carefully examined and measured.

Bias may arise in three ways: (1) by the inaccuracy of a method even when all the assumptions of the method are met, (2) as a consequence of unmet assumptions, and (3) through sampling error at small sample sizes. Monte Carlo simulations can be used to compute the total bias of an estimator of  $N$ . In my simulations, the bias of each regression estimator stemmed primarily from inaccurate methods, because the true source of capture probability heterogeneity was a continuous model and the parametric form was a three-parameter logistic model. The bias of the Chapman–Petersen  $\hat{N}$  was from violated assumptions (capture was heterogeneous); the method is otherwise nearly unbiased when applied to samples with at least four recaptures (Ricker 1975). The removal of samples with fewer than four recaptures or with  $\hat{P}(z) \leq 0.02$  was done to minimize small-sample bias. In application to real data, the criteria for deciding the range of sizes over which an estimator can be applied, or for removing bootstrap estimates, must be carefully examined. Plotting the smoothed functions produced by various  $\lambda$  will clarify the nature of the compromise made between bias and variance.

One can obtain confidence intervals for the estimated smoothing spline function  $\hat{P}(z)$  and the smoothing spline estimates  $\hat{N}(z)$  and  $\hat{N}$  only by bootstrapping, because analytic variances are not available. Many authors have shown that bootstrapping can provide estimates of variances or confidence intervals for some statistic under the assumptions of the model. Buckland and Garthwaite (1991) showed that bootstrapping would allow variance estimates for any mark-recapture estimate under any model for which parameter estimates can be calculated. When applied to the spline model, bootstrapping yields confidence intervals that stay within the permissible parameter range  $0 \leq \hat{P}(z) \leq 1$ . Fitting the spline by generalized cross-validation approximately minimizes the mean squared prediction error, which is the sum of variance and squared bias. A larger smoothing parameter reduces variance but increases bias, so some bias remains. Confidence intervals for estimates developed from parametric models are obtainable by bootstrapping or by using a variance estimated by the delta method. In the simulations, the estimates of  $SE(\hat{N})$  obtained by the delta method often had large biases, and the distribution of  $\hat{N}$  was not normal. Thus bootstrapping to obtain percentile confidence intervals may be preferable even for parametric models. For each method, the biases in estimates of  $N$  and  $V(\hat{N})$  may result in poor coverage; therefore, it is essential to examine the form of the capture function and to consider the likely directions of the biases.

Simulations do not incorporate all the decisions a biologist is likely to make about choice of models, goodness of fit, plausibility of model estimates, or the range of sizes that can be analyzed with acceptable precision. Carefully used, each approach may perform slightly better than these simulations suggest; however, the relative limitations should remain the same.

Calculation of spline-based capture probability functions and population estimates is simple, but the calculation of confidence bounds can be computationally intensive; therefore, alternative approaches may be of practical value. The smoothing spline is a value even if it is only used in exploratory analysis to help identify a parametric function with an appropriate form that can be fitted to data, or to guide choices of length-groups. For the limited range of parameter values used in my simulations, there appear to be few advantages of switching to a single parametric model of the capture function for a two-sample design.

With real data, one may choose a parametric form that fits the observed  $R_i$  better or worse than the unknown true function would. The parametric model may be superior for analyses when the probability of capture is low at all sizes and an appropriate parametric model has been identified. Smoothing spline functions tend to develop multiple peaks and valleys when capture probabilities are low, forming local maxima around the sparse recaptures. Maximum-likelihood estimation methods commonly perform poorly when capture probabilities are less than 0.1 (Otis et al. 1978). The computational simplicity of making separate estimates by size-groups is an advantage over other methods, but improvements can be made by using the spline to guide the choice of length-groups.

Extending the class of estimators developed by Otis et al. (1978) and Miller (1988) to readily incorporate various size-selective functions in multiple mark-recapture designs would be valuable. It appears that models relating to effects of time or behavior could be extended to include size selection so they can be tested by regression with smoothing splines. These models would require more complex coding and the application of multivariate splines.

When biologists fail to record lengths of all fish in each sample, analysis by size is not possible. Collection of size data should be recognized as essential in mark-recapture experiments. The West Indian Creek example demonstrates the value of correcting the size-selective sampling. Smoothed functions illustrated the dependence on size and time, and suggest that a depletion estimate would be inappropriate. Fisheries managers should note that it may be possible to obtain good estimates of abundance of small fish at sizes that are often excluded from analyses. In addition, it would behoove managers to document capture probabilities of large fish carefully, because capture probabilities for relatively rare large fish are most problematic and increasing the abundance of large fish is often a management goal. In my example with capture probability increasing with size, the grouping of large brown trout with any smaller size range would underestimate their capture probability, overestimate their abundance, and yield unrealistically narrow confidence intervals.

### Acknowledgments

This project was funded in part by the Federal Aid in Sport Fish Restoration (Dingell-Johnson) Program, and this paper is a completion report for Minnesota project F-26-R, study 690. I thank W.

Thorn for providing data on brown trout and D. Schluter for providing a copy of his FORTRAN program to fit a univariate spline. Copies of the original smoothing spline program or my modified versions of it are available from me on request. A. Bindman, R. Dorazio, M. Drake, K. Wilson, P. Wingate, and anonymous reviewers suggested useful improvements of the manuscript.

## References

- Bohlin, T., and B. Sundstrom. 1977. Influence of unequal catchability on population estimates using the Lincoln index and the removal method applied to electro-fishing. *Oikos* 28:123–129.
- Boy, V., and A. J. Crivelli. 1988. Simultaneous determination of gillnet selectivity and population age-class distribution for two cyprinids. *Fisheries Research* 6:337–345.
- Buckland, S. T., and P. H. Garthwaite. 1991. Quantifying precision of mark-recapture estimates using the bootstrap and related methods. *Biometrics* 47: 255–268.
- Cadigan, N. G., and R. B. Millar. 1992. Reliability of selection curves obtained from trawler trawl or alternate haul experiments. *Canadian Journal of Fisheries and Aquatic Sciences* 49:1624–1632.
- Craven, P., and G. Wahba. 1979. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik* 31:377–403.
- Dennis, J. E., Jr., and R. B. Schnabel. 1983. Numerical methods for unconstrained optimization and non-linear equations. Prentice-Hall, Englewood Cliffs, New Jersey.
- Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1:54–77.
- Eubank, R. L. 1988. Spline smoothing and nonparametric regression. Marcel Dekker, New York.
- Gu, C. 1987. What happens when bootstrapping the smoothing spline? *Communications in Statistics, Part A* 16:3275–3284.
- Hamley, J. M. 1975. Review of gillnet selectivity. *Journal of the Fisheries Research Board of Canada* 32: 1943–1969.
- Hoenig, J. M., F. B. Martin, and C. M. Heywood. 1989. Model-based sampling methods for effort and catch estimation. *American Fisheries Society Symposium* 6:181–189.
- Leclerc, J., and G. Power. 1980. Selectivity of fly-fishing, spinning and gillnet for brook char and ouaniche in a large northern Quebec river. *Environmental Biology of Fishes* 5:181–184.
- McCullagh, P., and J. A. Nelder. 1983. Generalized linear models. Chapman and Hall, New York.
- Miller, D. D. 1988. A model of the interactions between salmonid population dynamics and angling regulations. Master's thesis. Colorado State University, Fort Collins.
- Novotny, D. W., and G. R. Priegel. 1974. Electrofishing boats: improved designs and operational guidelines to increase the effectiveness of boom shockers. Wisconsin Department of Natural Resources Technical Bulletin 73.
- O'Sullivan, F., B. S. Yandell, and W. J. Raynor, Jr. 1986. Automatic smoothing of regression functions in generalized linear models. *JASA (Journal of the American Statistical Association)* 81:96–103.
- Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monographs* 62:1–135.
- Peterson, N. P., and C. J. Cederholm. 1984. A comparison of the removal and mark-recapture methods of population estimation for juvenile coho salmon in a small stream. *North American Journal of Fisheries Management* 4:99–102.
- Pope, J. A., A. R. Margetts, J. M. Hamley, and E. F. Akyuz. 1975. Manual of methods for fish stock assessment. Part III. Selectivity of fishing gear. FAO (Food and Agriculture Organization of the United Nations) Fisheries Technical Paper 41 (Revision 1).
- Reynolds, J. B. 1983. Electrofishing. Pages 147–164 in L. A. Nielsen and D. L. Johnson, editors. *Fisheries techniques*. American Fisheries Society, Bethesda, Maryland.
- Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. *Fisheries Research Board of Canada Bulletin* 191.
- Schluter, D. 1988. Estimating the form of natural selection on a quantitative trait. *Evolution* 42:849–861.
- Schluter, D., and D. Nychka. 1994. Exploring fitness surfaces. *American Naturalist* 143:597–616.
- Seber, G. A. F. 1982. The estimation of animal abundance and related parameters, 2nd edition. Griffin, London.
- Seber, G. A. F., and C. J. Wild. 1989. Nonlinear regression. Wiley, New York.
- Silverman, B. W. 1984. Spline smoothing: the equivalent variable kernel method. *Annals of Statistics* 12:898–916.
- Silverman, B. W. 1985. Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society B* 47:1–52.
- Sullivan, C. 1956. The importance of size grouping in population estimates employing electric shockers. *Progressive Fish-Culturist* 18:188–190.
- Thisted, R. A. 1988. Elements of statistical computing: numerical computation. Chapman and Hall, New York.
- Thorn, W. C. 1988. Brown trout habitat use in southeastern Minnesota and its relationship to habitat improvement. Minnesota Department of Natural Resources, Section of Fisheries, Investigational Report 395.
- Thorn, W. C. 1992. Validation of a trout habitat model

- for planning stream habitat improvement projects. Minnesota Department of Natural Resources, Section of Fisheries, Investigational Report 423.
- Vincent, R. 1971. River electrofishing and fish population estimates. *Progressive Fish-Culturist* 33:163–169.
- Wilkinson, L. 1990. SYSTAT: the system for statistics. SYSTAT, Evanston, Illinois.
- Yandell, B. S., and P. J. Green. 1986. Semi-parametric generalized linear model diagnostics. *Proceedings of the Statistical Computing Section, American Statistical Association* 1986:48–53.
- Yandell, B. S., and D. B. Hogg. 1988. Modelling insect natality using splines. *Biometrics* 44:385–395.

Received February 22, 1993

Accepted February 5, 1995