AIM

Association for Information and Image Management

1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
301/587-8202

Centimeter

Inches

MANUFACTURED TO AIM STANDARDS
BY APPLIED IMAGE, INC.

WSRC-MS-94-0192

D. B. Moore-Shedrow, Manager
Authorized Derivative Classifier

INTER-REPLICATE VARIANCE AND STATISTICAL POWER OF
ELECTROFISHING DATA FROM LOW GRADIENT STREAMS

by

M. H. Paller
WESTINGHOUSE SAVANNAH RIVER COMPANY

APPROVED BY:

D. B. Moore-Shedrow, Manager
Environmental Sciences Section

MARCH 1994

Westinghouse Savannah River Company
Savannah River Site
Aiken, SC 29808

THIS DOCUMENT WAS PREPARED IN CONNECTION WITH WORK UNDER THE U.S. DEPARTMENT OF ENERGY (CONTRACT DE-AC09-89SR18035). By acceptance of this paper, the publisher and /or recipient acknowledges the U. S. Government's right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper, along with the right to reproduce and to authorize others to reproduce all or part of the copyrighted paper.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P. O. Box 62, Oak Ridge, TN 37831; prices available from (615) 576-8401.

Available to the public from the National Technical Information Service, U. S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161

**Inter-replicate Variance and Statistical Power of
Electrofishing Data from Low Gradient Streams**

Michael H. Paller

Environmental Sciences Section
Westinghouse Savannah River Company
Savannah River Site
Aiken, South Carolina 29808

Abstract: Electrofishing data from streams on the coastal plain of South Carolina were used to assess the relationship between inter-replicate variance and the mean for catch per unit effort expressed as number of fish per 100 m² of stream surface area. Variance and mean were strongly and positively related. This relationship was unaffected by type and quantity of instream structure and was consistent among schooling and nonschooling species. Computations of statistical power indicated that large numbers of samples (>25) were required to detect small (<20%) differences among means and that the required number of samples was greater when catch per unit effort was low. Inter-replicate variance increased as the reach length of individual replicates decreased, with the greatest increases occurring at replicate reach lengths under 60 m. Inter-replicate variance can be decreased by increasing sampling effort (i.e., multiple passes) and by employing replicate reaches that are long. However, the most cost-effective approach may be to employ single passes and replicates of moderate size.

Fish assemblages are often sampled when assessing environmental impacts in streams. In these studies, catch per unit effort is usually compared among sites and sample replicates often consist of individual stream reaches sampled by electrofishing. The reliability of such comparisons is strongly dependent upon the variance among replicate samples. High inter-replicate variability makes it impossible to discriminate small differences among sites unless the number of replicate samples within each site is large. An understanding of the influence of inter-replicate variability on the precision of statistics estimated from electrofishing samples is necessary for the rational design of environmental impact surveys and other studies based on electrofishing data.

Most types of aquatic organisms including benthos (Downing 1979, Morin 1985), periphyton (Morin and Cattaneo 1992), zooplankton (Pace et al. 1991), and ichthyoplankton (Cyr et al. 1992) exhibit aggregated distributions that result in a positive correlation between mean abundance and the variance of abundance. Taylor's Power Law (Taylor 1961) states that the relationship between variance and mean can be described by the regression of log variance on log mean abundance. The slope of this relationship putatively reflects the degree of aggregation (hence is sometimes referred to as the "index of aggregation") and can be used to calculate the type of data transformation that decorrelates the variance from the mean as required for ANOVA and other parametric analyses. Taylor's Power Law is often employed by biologists (Green 1979) and has been a useful tool in studies of sampling variability among aquatic organisms.

While electrofishing data are often analyzed statistically and the results used for decision making, the statistical power of such analyses is seldom computed (Peterman 1990) and little discussion is provided on the

effects of inter-replicate variability on sampling design. Studies addressing the latter problem have dealt primarily with the use of removal method estimators to estimate population sizes of single species in salmonid streams (Hankin 1984, Bohlin et al. 1989), although Harding and Connor (1992) recently described the precision of estimates of the mean for electrofishing samples from Florida lakes. In this paper, I describe sampling variability associated with electrofishing samples collected from diverse fish assemblages in first to fourth order streams on the coastal plain of South Carolina. These low gradient streams lack well defined riffle/pool/run sequences which may result in quantitatively different patterns of variation among spatial replicates than observed in high gradient streams. As is typical of many environmental studies in such habitats, my emphasis was on catch per unit effort rather than removal method population estimates for single species.

Methods

Study area and field methods. This study was conducted on the Savannah River Site (SRS), a 780 km² Department of Energy nuclear materials production facility located on the Atlantic coastal plain near Aiken, SC. Electrofishing samples were collected during the summer and fall of 1990 and 1992 from first through fourth order sites in streams in and near the SRS. These streams had little or no development of the prominent mesohabitats (riffles, pools and runs) characteristic of high gradient streams. Instream habitat consisted of snags, undercut banks, logs, submerged brush, woody debris, and macrophyte beds.

Samples were collected by DC electrofishing using a backpack electrofisher (Smith-Root Model 15-A Electrofisher), a somewhat larger unit (Smith-Root Model 1.5 KVA Electrofisher, 400 watt/120 VAC generator) mounted in a 2 m

long barge) or a boat mounted shocker (Smith-Root Model VI-A Electrofisher, 4500 watt/230 VAC generator). To maximize sampling efficiency, larger and more powerful gear was used in larger streams as explained more fully in Paller (1994).

Two types of sample stations were employed in this study. The first consisted of three replicate stream segments separated by approximately 10 m; each replicate was 100 m long and blocked with 6.35 mm mesh nets at both ends prior to sampling. The second consisted of numerous contiguous sample segments, usually approximately 20 m long, summing to a total length of 50 - 400 m. Individual segments were not blocked, although block nets were set at the downstream and upstream ends of the lowermost and uppermost segments, respectively. At both types of sample stations, the entire habitat area (bank to bank) was electrofished while moving upstream. All microhabitats (brush piles, logs, undercut banks, etc.) were carefully sampled, and all fish retained by the 6.35 mm mesh dip nets were collected. Seven passes were made through each sample station in an effort to sample the fish assemblages as thoroughly as possible. Fish were identified to species and measured after each pass, then released downstream of the lowermost block net. Habitat measurements made in conjunction with electrofishing included stream width, stream depth, current velocity, and the percentages of stream bottom covered or overhung by submerged debris (leaves, twigs, and woody debris), submerged brush/snags, submerged root masses, aquatic macrophytes, and low growing riparian vegetation. Further details on habitat measurements can be found in Paller (1994).

Data analysis. Estimates of fish abundance produced by electrofishing are typically expressed in terms of catch per unit effort (CPUE), with effort quantified in terms of stream surface area sampled, stream reach length sampled, or time spent sampling. I expressed fish abundance as number per 100

m² of stream surface area. For the sample stations with separate 100 m replicates, I calculated CPUE for each replicate, then calculated the mean and variance from the three replicate samples at each sample station. For the sample stations with small contiguous sample segments, I divided the segments into three groups of equal size, combined the segments within each group to form a single replicate, and calculated the mean and variance for the three resulting replicates. I skipped segments between groups so the replicates were separated by approximately 20 m. I performed this combining procedure to ensure that all sample stations included in the analysis were comparable in terms of general sampling methodology. The remaining difference between the two types of sample stations, presence/absence of block nets, did not significantly affect the relationship between variance and mean as explained more fully later.

I next log₁₀ transformed the means and variances and regressed the variance on the mean:

$$\log_{10} s^2 = a + b \log_{10} M \quad 1$$

where s^2 is the variance, M is the mean, and a and b are regression coefficients. I added habitat variables (e.g., stream width) to the log variance/log mean models (i.e., multiple regression analysis) to determine if they accounted for residual variance not explained by the mean. I also constructed separate log variance/log mean models (simple regression) for multiple pass and single pass data and compared their regression statistics.

Statistical power can be defined as the probability of rejecting the null hypothesis (H_0) when it is really false (see Peterman 1990 for an extensive discussion of the importance of power in fisheries research). Power is a function of alpha (the probability below which H_0 is

rejected), effect size (the magnitude of the difference that is being tested for), n (number of samples), and s^2 (the sample variance). The number of samples (n) required to detect a specified difference (effect size) between two population means with unequal population variances by iteratively solving for n using the following formula (Zar 1984, Morin and Cattaneo 1992):

$$n > \frac{s^2_1 + s^2_2}{D^2} (t_{\alpha(2),v} + t_{\beta(1),v})^2 \quad 2$$

where s^2_1 and s^2_2 are the inter-replicate variances for the two groups, D is the absolute value of the difference between the group means, $t_{\alpha(2),v}$ is the critical value for a two-tailed test with v degrees of freedom and an error probability, α , of false positive conclusions and $t_{\beta(1),v}$ is the critical value for a one-tailed test with v degrees of freedom and an error probability, β , of false negative conclusions. The degrees of freedom for this test were adjusted for unequal variances as follows:

$$v = \frac{\left(\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2} \right)^2}{\frac{\left(\frac{s^2_1}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{s^2_2}{n_2} \right)^2}{n_2 - 1}} \quad 3$$

where n_1 and n_2 are the sample sizes for the two groups that are being compared.

To obtain the estimates of variance required in equation 2, I used equation 1 with a correction for back transformation from the logarithmic to the arithmetic scale (Newman 1993). Correcting equation 1 for backtransformation yielded:

$$s^2 = 10^a M^b 10^{mse/2}$$

4

where mse is the mean square error from the regression.

For comparison with other studies and to illustrate some points more simply than I could with statistical power, I calculated the coefficient of variation of the mean (CV_M), which is a useful and simple estimate of the precision of measurement (Cyr et al. 1992). CV_M is a function of the sample mean, inter-replicate variance, and number of samples:

$$CV_M = s M^{-1} n^{-0.5}$$

5

The lower the CV_M , the greater the potential statistical power. CV_M should be distinguished from the coefficient of variation of samples (CV) which is independent of n:

Results and Discussion

Variance/mean regressions. CPUE (all species summed) was characterized by a strong positive relationship between the variance and mean ($R^2=0.77$) (Figure 1). The predictive power of the regression model was not significantly ($P \leq 0.05$) improved by the inclusion of higher order (i.e., quadratic and cubic) terms in the regression model. A slope of 2.0 for a log variance/log mean relationship indicates that the coefficient of variation among samples does not change with the mean. The slope of 1.57 for CPUE indicated that the coefficient of variation among samples decreased with the mean (Figure 2). This trend had important consequences for statistical power as I discuss later.

I individually added habitat variables (width, depth, each of the percent cover variables, and cumulative cover) to

the mean-variance regression model for CPUE to see if they accounted for any variability that was unexplained by the mean. Cumulative cover consisted of the sum of percent coverage of submerged debris, submerged brush/snags, submerged root masses, aquatic macrophytes, and low growing riparian vegetation. All independent variables were appropriately transformed following preliminary examination of the data. None of the habitat variables added significant explanatory power to the relationship between log variance and log mean.

I was also interested in determining whether the type of electrofishing gear and the use of block nets affected sampling variability. To test gear type, I conducted an analysis of covariance (ANCOVA) with log variance as the dependent variable, log mean as the continuous independent variable and electrofishing gear type (backpack, barge or boat) as the group independent variable. Gear type did not significantly influence sampling variability ($P=0.931$). To test the effects of block nets, I conducted a second ANCOVA with presence/absence of block nets as the group independent variable (in place of gear type). This factor, too, did not significantly influence sampling variability ($P=0.883$).

Variability among replicate electrofishing samples presumably stems from variations in the spatial distribution of fish and from sampling error. Sampling error would be expected to decline as sampling effort increases because high effort would be expected to represent actual spatial distribution with greater fidelity than low effort. Thus, an inverse correlation might be expected between sampling effort and variability among replicate samples. I tested this hypothesis by comparing the variance-mean relationship for data collected with one electrofishing pass and seven electrofishing passes. The y-intercept and slope values for single pass and seven pass electrofishing were similar (as

indicated by strong overlap between the 95% confidence intervals for both statistics, Table 1) indicating that the relationship between variance and mean was not affected by the number of electrofishing passes. However, mean CPUE will be higher when seven pass sampling is used (assuming that data from different passes are combined) resulting in a lower coefficient of variation (Figures 1 and 2).

Mean/variance relationships were examined for four of the most numerous species: redbreast sunfish Lepomis auritis, spotted sunfish Lepomis punctatus, yellowfin shiner Notropis lutrensis, and dusky shiner Notropis cummingsae. Species were categorized as schooling (yellowfin shiner and dusky shiner) or nonschooling (redbreast sunfish and spotted sunfish) based on field observations and literature accounts of species specific behaviors. Differences in the degree of aggregation among schooling and nonschooling species could conceivably affect mean/variance relationships. However, the regression statistics for schooling and nonschooling taxa overlapped considerably and were generally similar to the regression statistics for all species summed (Table 1). These results suggest that the basic model for all species summed adequately represented mean-variance relationships for most individual species as well.

Effects of replicate area on inter-replicate variability. I added replicate reach length to the mean-variance regression model for NCPUE to determine if it accounted for any variability that was unexplained by the mean. Its contribution was not statistically significant ($P=0.634$) after the effects of the mean were accounted for. However, most of the replicates in the regression analysis were fairly large (median reach length=100 m) which limited my ability to investigate the effects of very small replicate reach lengths through the use of regression analysis.

As a means of overcoming the preceding problem, I used a simulation approach to investigate the relationship between replicate length and inter-replicate variability. I selected the five longest sample stations that were composed of small contiguous replicates (16-18 in number), divided the segments from each station into two equal groups, calculated the number of fish/100 m² (CPUE) for each group, and calculated the inter-replicate variance for the two groups (which served as two sample replicates). Then I reduced the total reach length of each group by deleting progressively greater numbers of segments from it (an equal number was deleted from each group to keep the group areas equal). The end result was a curve illustrating how the variance changed as the length of the two replicates changed. Because the shape of this curve was affected by the segment chosen as a starting point, I randomly rearranged the segments before performing the preceding procedure, and repeated the procedure 100 times for each sample station. From the 100 resulting curves, I generated a single final curve for each sample station by averaging. This technique permitted me to investigate mean/variance relationships among replicates over a greater range of lengths (approximately 20-200 m) and allowed me to view the influence of replicate area on a station by station basis.

Note that it would be invalid to randomly rearrange segments if the abundance of fish within each segment was related to the abundance of fish in adjoining segments. To test this, I performed a runs test for trend data (i.e., runs up and down test, Sokal and Rohlf (1981)) on fish numbers in the segments (in their original order) from each of the 5 sample stations. In none of the 5 stations did the fish numbers in the consecutive segments exhibit significant ($P \leq 0.05$) departure from a random arrangement.

Figure 3 illustrates the results for one of the five sample stations; the others are not shown since they were very similar. The simulations demonstrated that variance increased as the replicate length decreased, with the sharpest increases occurring at replicate lengths under approximately 60 m. It is known that error variance tends to be maximized when the sample units size coincides with the scale of greatest patchiness (Green 1979). The scale of greatest patchiness among stream fishes in southeastern coastal plain streams is unknown, but might coincide with the stream surface areas or volumes occupied by snags, brushpiles, and log jams, all of which are important sources of habitat complexity and structure. Small replicates may more closely approach the scale of these sources of patchiness in fish distribution, thus increasing the variability among replicate samples. It is noteworthy, however, that while the average trend was for variance to decrease as the reach length increased, some of the individual simulations for each stream showed the opposite pattern. This occurred when segments with unusually high or low numbers of fish were added after the reach was already fairly long. Such simulations mimic the real world possibility of encountering increased environmental heterogeneity as sample reaches are enlarged.

Data transformation. Data with unequal variances must be transformed to achieve equality of variances before being analyzed with parametric methods unless techniques for unequal variances are employed (Zar 1984). The variance stabilizing transformation for a specific data set can be calculated once the slope of the log mean/log variance is known (Taylor 1961):

$$X' = X^{1-(b/2)}$$

where X is CPUE and X' is transformed CPUE. However, for simplicity biologists often prefer a more commonly used transformation such as the square root or log. The former is theoretically appropriate when the slope is one (i.e., the distribution is Poisson) and the latter when the slope is two (i.e., the distribution is lognormal). Because many aquatic organisms exhibit intermediate levels of aggregation, Downing (1979 and 1981) suggested that the fourth root transformation will often be appropriate. France (1987) indicated that the best transformation is the one that results in the smallest correlation between the variance and mean and found that the transformation that best decorrelates the variance from the mean is not always the one predicted by Taylor's Power Law.

Neither the square root nor the fourth root transformation removed the significant ($p < 0.05$) positive correlation between transformed mean and variance in the case of CPUE. While both the log and Taylor's Power Law transformation successfully decorrelated the mean from the variance, the slope of the original log mean/log variance relationship differed from 2.0 (Table 1), the value at which the log transformation is theoretically appropriate. Over-transformation is undesirable because it alters the relationship between variance and mean to a greater extent than necessary and in extreme cases can result in a significant negative correlation between variance and mean.

Statistical power. Statistical power was expressed as the number of replicates (per group) required to detect various percent differences between two group means. This approach resembles the question often posed by biologists studying impacts to a fish assemblage; e.g., how many samples are needed to detect a 50% difference between the mean density of fish upstream from and downstream from a sewage outfall. I set alpha (probability of rejecting the null when it is not really false) at the commonly employed level of 0.05 and beta

(probability of failing to reject the null when the null is really false) at 0.05 because I assumed that both types of error can have equally adverse consequences in environmental impact studies. I also evaluated a less rigorous beta level of 0.20.

The size of the mean strongly affected the required number of samples (Figure 4). For example, over 30 samples were required to detect a 50% reduction from a mean CPUE of 1.3 ($\alpha=\beta=0.05$) while only seven samples were required to detect a 50% reduction from a mean CPUE of 89.2. This pattern was not unexpected considering the effects of the mean on the CV of CPUE (Figure 2). The required sample number increased geometrically as the difference between group means became smaller making it necessary to collect large numbers of samples to detect small differences (e.g. 10%). This pattern persisted even at high values of the mean and when beta was increased to the relatively liberal level of 0.2. Note that the preceding results will be affected by replicate length since replicate length affects variance. The sample size requirements shown in Figure 4 apply for replicates with a length of approximately 60-130 m, which composed most of the data from which the regression of variance on the mean was computed.

Sample size requirements for fish captured by electrofishing can be compared with the sample size requirements for other types of aquatic organisms. Cyr et al. (1992) calculated that 33 samples were needed to detect a 50% change in larval fish density at an intermediate density of 10 fish per sample ($\alpha=\beta=0.05$). Morin and Cattaneo (1992) calculated that approximately 5-7 samples were needed to detect a 50% reduction ($\alpha=\beta=0.05$) in periphyton cell density at intermediate values of the mean. I estimated that 15 samples were needed to detect a 50% reduction in electrofishing CPUE at intermediate values of the mean

($\alpha=\beta=0.05$). Results have also been expressed in terms of precision. Morin (1985) calculated that approximately ten samples were needed to achieve a precision of 20% at an intermediate density of stream benthos ($10^3/\text{m}$). I calculated that approximately four samples would be required to attain a precision of approximately 20% at the median CPUE (approximately $13/100 \text{ m}^2$). While the relative magnitude of the preceding differences will not be maintained for all values of the mean nor for all data sets, these results suggest that the variability of electrofishing samples from South Carolina coastal plain streams is within the range of the variability exhibited by other types of aquatic organisms.

Implications for sampling design. Because power decreases as mean density decreases (Figure 4), high power will be difficult to obtain where the catch is low because fish are scarce (e.g., polluted areas) or difficult to collect (e.g., deep or swiftly flowing stream segments). In such cases the researcher has several options including increasing the number of replicates, increasing the sampling effort (e.g., performing multiple passes) so the mean catch is higher, or using larger replicates. All of these changes will result in an increase in precision, although at the cost of increased effort. Below, I compare the relative costs (in terms of sampling time or effort) of these changes in sampling design.

By repeatedly electrofishing a replicate stream segment, CPUE (summed over passes) can be increased resulting in potential increases in precision without increasing the number of replicates. I compared four pass sampling with single pass sampling in terms of the number of replicates and time required to obtain a precision (CV_M) value of 0.2. I chose four pass sampling because four passes collected most of the catchable fish at my study sites (greater numbers of passes generally produced relatively few additional fish).

Four pass sampling collected an average of approximately three times more fish than single pass sampling and took approximately three times as long per replicate to complete. The latter factor indicates, that for four pass sampling to be competitive with one pass sampling in terms of time, it must require one third or less replicates to reach the same level of precision. To perform the required computations, I rearranged the formula for computing precision to solve for n :

$$n = s^2 M^{-2} CV_M^{-2}$$

6

I found that the number of replicates needed for a CV_M of 0.2 with four pass sampling was approximately one-half the number needed with one pass sampling regardless of the mean CPUE. While these calculations indicate the greater efficiency of single pass sampling, multiple passes remain an option for increasing the precision of density estimates where the maximum potential sample area is comparatively small or when multiple passes are needed to meet other objectives.

Increasing the replicate length is another way to increase the precision of density estimates, although large sample areas will result in greater total effort unless the increase in precision resulting from the use of greater sample areas permits the use of fewer replicates such that the total sample area (sum of all replicates) stays the same or decreases. To evaluate the influence of replicate area on total effort, I determined the number of replicates required to obtain a CV_M of 0.1 for replicates of different length (Figure 3). Because of the relationship between replicate length and inter-replicate variance, the required number of replicates increased markedly as replicate length decreased. However, the total reach length that needed to be sampled (the product of the number of replicates and replicate length) changed little, although for all five sample stations

I analyzed in this manner there appeared to be slight increases in total reach length at individual replicate lengths over approximately 80-100 m. This analysis suggests that there is no marked advantage to using either very large or very small replicates, although this might not be the case in other streams that are characterized by a different relationship between replicate length and inter-replicate variance. In my streams, replicate lengths in the range of 60-100 m may be best since replicates of this size seem to require somewhat less total area than larger replicates and may represent a smaller investment in sample station layout and sample work-up than replicates of smaller size.

Summary and conclusions

- 1) CPUE data from stream fish communities on the South Carolina coastal plain were characterized by a strong correlation between inter-replicate variance and the mean.
- 2) The mathematical relationship between variance and mean was not directly affected by habitat, fish behavior (schooling vs nonschooling), or sampling effort, although the latter indirectly influenced inter-replicate variability by affecting the number of fish collected. However, inter-replicate variability was inversely related to replicate length, with the most marked increases occurring at replicate lengths under 60 m.
- 3) The number of samples required to detect a specified percentage difference between means was inversely related to the mean CPUE. However, even when the mean CPUE was high, large numbers of samples were required to detect small differences.
- 4) Statistical power can be increased by increasing the number of replicates, increasing the mean number of fish in each replicate through greater sampling effort (i.e., multiple passes) or by increasing the size of the replicates. In South Carolina coastal plain streams, the most cost-

effective strategy may be to use single passes and replicates 60-100 m in length.

Acknowledgements

Appreciation is extended to Henry Mealing, Alan Stuart, Hugh Westbury, and Bruce Saul for conducting field work associated with this project. I thank Cary Tuckfield and John Gladden for reviewing an earlier draft of this paper. This document was prepared in connection with work under the U.S. Department of Energy (Contract DE-AC09-89SR18035).

References

- Bohlin, T., S. Hamrin, T.G. Heggberget, G. Rasmussen, and S.J. Saltveit. 1989. Electrofishing - theory and practice with special emphasis on salmonids 173: 9-43.
- Cyr, H., J.A. Downing, S. LaLonde, S.B. Baines, and M.L. Pace. 1992. Sampling larval fish populations: choice of sample number and size. Transactions of the American Fisheries society 121:356-368.
- Downing, J.A. 1979. Aggregation, transformation, and the design of benthos sampling programs. Canadian Journal of Fisheries and Aquatic Science 36: 1454-1463.
- France, R.L. 1987. Aggregation in littoral amphipod populations: transformation revisited. Canadian Journal of Fisheries and Aquatic Science 44: 1510-1515.
- Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley & Sons, New York, New York.

Hankin, D.G. 1984. Multistage sampling designs in fisheries research: applications in small streams. *Canadian Journal of Fisheries and Aquatic Science* 41: 1575-1591.

Hardin, S., and L.L. Connor. 1992. Variability of electrofishing crew efficiency, and sampling requirements for estimating reliable catch rates. *North American Journal of Fisheries Management* 12:612-617.

Morin, A. 1985. Variability of density estimates and the optimization of sampling programs for stream benthos. *Canadian Journal of Fisheries and Aquatic Science* 42: 1530-1543.

Morin, A., and A. Cattaneo. 1992. Factors affecting sampling variability of freshwater periphyton and the power of periphyton studies. *Canadian Journal of Fisheries and Aquatic Science* 49: 1695-1703.

Newman, M.C. 1993. Regression analysis of log-transformed data: statistical bias and its correction. *Environmental Toxicology and Chemistry* 12: 1129-1133.

Pace, M.L., S.E.G. Findlay, and D. Lints. 1991. Variance in zooplankton samples: evaluation of a predictive model. *Canadian Journal of Fisheries and Aquatic Science* 48:146-151.

Paller, M.H. 1994. Relationship between fish assemblage structure and stream order in South Carolina coastal plain streams. *Transactions of the American Fisheries Society* 123:150-161.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. *Canadian Journal of Fisheries and Aquatic Science* 47:2-15.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. W. H. Freeman and Company, San Francisco, CA.

Taylor, L.R. 1961. Nature 189: 732-735.

Zar, J. H. 1984. Biostatistical analysis, 2nd edition. Prentice Hall, Englewood Cliffs, NJ.

Table 1. Statistics for simple regression models predicting $\log_{10}$ variance from the $\log_{10}$ mean electrofishing CPUE (number of fish per 100 m²) in South Carolina coastal plain streams. Values in parentheses are 95% confidence intervals.

Dependent variable	R ²	n	slope	y-intercept
All species summed (7 electrofishing passes)	0.77	56	1.57 (1.33 - 1.81)	-0.54 (-0.85 - -0.23)
All species summed (1 electrofishing pass)	0.74	56	1.42 (1.19 - 1.64)	-0.40 (-0.61 - -0.19)
Redbreast sunfish (7 electrofishing passes)	0.81	41	1.51 (1.30 - 1.72)	-0.42 (-0.59 - -0.25)
Yellowfin shiner (7 electrofishing passes)	0.79	32	1.46 (1.29 - 1.63)	-0.35 (-0.51 - -0.19)
Dusky shiner (7 electrofishing passes)	0.89	23	1.64 (1.38 - 1.90)	-0.17 (-0.33 - -0.01)
Spotted sunfish (7 electrofishing passes)	0.74	40	1.49 (1.20 - 1.78)	-0.55 (-0.75 - -0.35)

FIGURES

Figure 1. Relationship between the variance and the mean for electrofishing catch per unit effort (number/100 m²) in South Carolina coastal plain streams.

Figure 2. Relationship between the coefficient of variation (Standard deviation/mean x 100) and the mean for electrofishing catch per unit effort (number/100 m²) in South Carolina coastal plain streams.

Figure 3. The relationship between inter-replicate variance and the length of individual sample replicates (lower plot) as determined by a simulation program (described in text). Also shown is the relationship between replicate length and the number of samples needed to reach a CV_M of 0.1 (see text for definition of CV_M). Upper plot depicts the relationship between total sample site length (number of replicates times length of replicates) needed to reach a level of precision of 0.1 and the length of individual replicates.

Figure 4. Number of replicate samples required to detect a significant difference between two means with a one sample t-test for means with unequal variances. Means represent catch per unit effort expressed as number per 100 m². Data were not transformed for these calculations.

ALPHA=0.05, BETA=0.05

ALPHA=0.05, BETA=0.20

**DATE
FILMED**

6 / 24 / 94

END