STUDY FINAL REPORT

State: Michigan

Project No.: F-81-R-5

Study No.: 230710

Title: Evaluation and development of quantitative methods for fishery surveys, assessments, and inventory programs in Michigan

Period Covered: October 1, 2003 to September 30, 2004

Study Objective: To develop appropriate statistical survey design to evaluate Michigan's inland stocking program as part of a Statewide Resource Inventory Program.

Findings: This final report summarizes the work of Jobs 1 (Advise on sampling, sample size and similar issues), 4 (Write annual performance report), and 7 (Write final report). Results of Job 2 are presented in:

Hayes, D., E. Baker, R. Bednarz, D. Borgeson, Jr., J. Braunscheidel, J. Breck, M. Bremigan, A. Harrington, R. Hay, R. Lockwood, A. Nuhfer, J. Schneider, P. Seelbach, J. Waybrant, and T. Zorn. 2003. Developing a standardized sampling program: the Michigan experience. Fisheries 28(7):18-25.

Job 3, serving on the Statewide Creel Survey Committee, will be continued through Study 230646; Jobs 5 and 6 will not be completed, because they have been moved to Study 230712.

Abstract.–Sample sizes for two inland lake habitat and three lakeshore development metrics were evaluated for eight southern Michigan lakes. Habitat metrics were index of vegetation cover and counts of submerged trees. Lakeshore development metrics were dwelling and dock counts, and percentage of shoreline armored. Metrics were sampled per 1,000 ft of shoreline. Sample sizes were estimated using three different methods. Bootstrapping techniques were used to evaluate accuracy of estimates, empirical means and SD of means were used to evaluate data precision, and power analysis of individual lake to cluster comparisons were made. Bootstrapping results indicated that a minimum sample of five 1,000 ft shoreline transects was required to provide reliable estimates of each metric mean. Sampling of 15 to 30 transects was necessary to attain maximum improvement in individual metric precision. Power analysis was fixed at α=0.05, and β=0.20 was considered acceptable protection against Type II Error. Acceptable levels of power were attained with <25 transects sampled at 50%-71% of lakes. Remaining lakes required >35 transects to attain acceptable levels of power. To further evaluate characteristics of variability, three large lakes and three impoundments were sampled and precision was evaluated for counts of submerged trees, dwellings, docks, and percentage of shoreline armored. Sampling of 10 to 25 transects was necessary to attain maximum improvement in individual metric precision for the three lakes and 15 to 30 transects for the three impoundments. Results and recommendations of this study are that a minimum of 30 randomly selected shoreline transects should be sampled, and that additional transects should be sampled whenever possible.

Background.–Human activities can negatively affect inland lake ecosystems through alterations in water quality and physical habitat. For example, increased nutrient loadings from septic seepage and lawn fertilizers can increase primary production, increase algae and aquatic vegetation to nuisance levels, and decrease concentrations of dissolved oxygen when excess algae and vegetation decompose. In addition, the quantity and quality of physical habitat available to fishes in the littoral zone can be altered by removal of coarse woody debris, by an increase or decrease (via chemical or mechanical removal) in aquatic macrophytes, and by homogenization of the shoreline through erosion control efforts (e.g., rip-rap and sheet piling). Such changes in water quality and habitat features have been shown to negatively impact fish growth (Schindler et al. 2000), limit natural reproduction of certain fish species (Rust et al. 2002), and reduce fish species richness and shift assemblage structure towards more tolerant species (Jennings et al. 1999). Consequently, monitoring and assessing the influence of human activities on the condition of inland lake systems is necessary for sound management of these resources.

A primary goal of the Michigan Department of Natural Resources' Lakes Status and Trends Program (Hayes et al. 2003) is to monitor and assess the impacts of human activities on inland lakes. Ultimately, this program will allow managers to take advantage of the large sample sizes associated with a given class of lakes. However, few guidelines are available for setting appropriate sample sizes for measuring human impacts, especially to assess status and detect trends for a large number of lakes distributed throughout the State. The level of sampling effort must balance the needs between collecting quality data and being able to rapidly assess conditions in a relatively large number of lakes. The ability to make statistically valid inferences of cause and effect are often limited by high variance typical of ecological data. Pooling data from multiple lakes may enhance the ability to detect trends by increasing sample sizes. Fundamental to such a program are three preliminary tasks. First, determination of sampling effort needed to accurately estimate lake habitat and human lakeshore development. Second, determination of sampling effort needed to efficiently estimate the conditions in individual lakes. Third, is determination of statistical power to detect differences between individual lakes and a cluster of lakes sharing similar characteristics. This third evaluation directly assesses the status and trends program objective of using random samples for Michigan lakes to characterize statewide conditions.

The objective of this study was to determine appropriate sample sizes for characterizing littoral zone habitat and human lakeshore development by determining the precision of visually estimated vegetation and coarse woody debris abundance, counts of dwellings and docks, and estimates of the percentage of shoreline armoring in inland lakes. The goal of this study was to determine sample sizes required to adequately characterize conditions in individual lakes and to determine statewide status and detect trends. This study determined the precision of these metrics for individual lakes and estimated the power when data from an individual lake was compared to a pooled set of data from similar lakes.

<u>Procedure</u>.–

Study Sites

Eight Southern Michigan Lakes

Eight southern Michigan lakes were selected for shoreline sampling (Table 1). Lakes varied in size from 29 to 796 acres. Shoreline development index (SDI) was measured for each lake using methods found in Orth (1983). Lakes were sampled during July and August during 2002 or 2003. These lakes were considered similar and would likely be found within a single status and trends class. This assumption was appropriate since these lakes are all found at approximately the same latitude (thus share similar annual temperature variation), have similar physical and limnological characteristics, and have similar fish species composition. This study was designed to provide an appropriate evaluation of the status and trends sampling program.

Three Large Lakes and Three River Impoundments

To further evaluate variability of the metrics measured in the eight southern Michigan lakes, all of which were <1,000 acres in size, three additional lakes (2 of which were >1,000 acres in size) and three large river impoundments (2 of which were >1,000 acres in size) were selected for sampling (Table 1). Shoreline development index (SDI) was measured using methods found in Orth (1983). Lakes and impoundments were sampled during July and August 2004. Characteristics of these three lakes and the three river impoundments are quite different. They are found at different latitudes, physical and limnological characteristics are not similar, and fish species composition is quite variable. Consequently, placement of these water bodies in a single class would not be appropriate. The purpose in evaluating each metric's variability was to determine if variability was associated with lake size and characteristics. Specifically, this study compared sample size estimates for smaller lakes with those for lakes and impoundments that were quite different.

Methods

Data collection

Lake habitat and lakeshore development metrics were visually assessed by boat during daylight hours. Data were recorded in 1,000 ft intervals and each transect was completed in ~2 minutes. Sampling crews consisted of 2-3 members and each person was responsible for collection of specific data. Hand held Global Positioning System (GPS) units (Garmin Map 76S) were used to measure transects within each lake. During data collection, the boat traveled approximately 100-200 ft from shore to limit the amount of movement in and out from shore to avoid structures such as docks. Transect length paralleled the shoreline and did not include deviations to avoid structures.

Two habitat and three development metrics were selected for shoreline sampling. The habitat metrics were index of vegetation cover and counts of submerged trees. The index of vegetation provided a measure of the percentage of the littoral zone that was occupied by aquatic macrophytes (emergent, submergent, and floating), and was estimated using abundance categories where 0% = 1, 1-25% = 2, 25-50% = 3, 50-75% = 4, and 75-100% = 5. The number of submerged trees large enough to provide fish habitat (main trunk averaging roughly 3in or greater) were counted as the boat moved along the shoreline. Any submerged trees visible between the boat and the lakeshore were counted. No effort was made to count trees that were too deep to be readily visible from the boat.

Dwellings having obvious lake frontage were counted within each transect. Criteria for obvious lake frontage included: contiguous lawn from near edge of a dwelling to the near edge of the lake, and no road separation between a dwelling and the lake. Only dwellings located immediately along the shoreline were counted, visible dwellings occurring behind other shoreline dwellings were not counted.

Docks extending from the shore into or upon the lake were counted. Docks piled on the shore and not in obvious use were not counted. Size of docks and number of hoists or mooring positions were not evaluated.

Percentage of shoreline, within a 1,000 ft transect, having armoring was visually estimated. Bank (shoreline) armoring included wood or steel sheet pilings, cement walls, or gabions positioned along the shoreline in a vertical or sloping position to prevent erosion. Loosely placed cobbles with no structural support were considered armoring if it appeared that these structures were more than just a decorative border.

Three analyses of sample size (i.e., number of 1,000 ft transects) were completed for each metric. For the eight southern Michigan lakes, the number of samples necessary to provide a reliable estimate of individual lake metric means was determined. For the eight southern Michigan lakes, three large lakes, and three impoundments, sample sizes were estimated to attain varying levels of relative precision (mean ±proportion of that mean). Finally, for the eight southern Michigan lakes, power analysis was done for each metric to ensure reliability in interpretation of status and trends data. For final sample size evaluations, general characterizations across lakes were made.

Data analysis

Evaluation of reliability of means

This analysis was done for the eight southern Michigan lakes only. Sample sizes necessary to provide reliable estimates for each lake-metric mean were calculated using bootstrapping techniques (Efron and Tibshirani 1993). Number of transects sampled varied from 1 to 30 per lake and metric. While only one of the sample lakes (Wamplers Lake) had circumference ≥ 30,000 ft (thus, allowing completion of 30 transects), the variation seen within this group of lakes was assumed to be representative of lakes found in Michigan. Following traditional bootstrapping techniques, sampling was done with replacement and up to 30 transects were considered for each lake.

The mean $\bar{x}_{l,m}$ for lake *l* and metric *m* was estimated as:

$$\bar{x}_{l,m} = n^{-1} \sum_{k=1}^{n} \left(\frac{\sum_{i=1}^{j} x_{l,m,i}}{j} \right), \tag{1}$$

with *j* transects sampled and *n*=2,000 iterations. The standard error $SE_{l,m}$ was estimated from the bootstrap standard deviation (Efron and Tibshirani 1993, Manly 1997) for mean $\bar{x}_{l,m}$ as:

$$SE_{l,m} = \left((n-1)^{-1} \cdot \sum_{k=1}^{n} \left[\theta_{l,m} - \left(\sum_{i=1}^{j} x_{l,m,i} \Big/ j \right) \right]^2 \right)^{-1}, \tag{2}$$

with expected data set mean $\theta_{l,m}$. Thus, $SE_{l,m}$ included variability within each data set and any bias due to sampling error. For each metric then, the coefficient of variation of the mean $CV_{\bar{x}l,m}$ was:

$$CV_{\bar{x}l,m} = \frac{SE_{l,m}}{\bar{\theta}_{l,m}}. \tag{3}$$

To provide relative measures of accuracy for each metric, $CV_{\bar{x}}$ for each metric were averaged across lakes for each sample size (transects sampled).

Evaluation of precision

This analysis was done for all eight southern Michigan lakes, three large lakes, and three impoundments. Sample sizes $N_{l,m}$ for each lake l and metric m were estimated as (Snedecor and Cochran 1989:52):

$$N_{l,m} = t^2 \frac{s_{l,m}^2}{L_{l,m}^2}, \quad (4)$$

with Student-t t of 1.96 (α=0.05) and standard deviation s. The relative error L for precision p was:

$$L_{l,m} = \theta_{l,m} p, \quad (5)$$

with levels of precision from 1% to 377%. Percent changes were based on individual means and were selected to represent an appreciable and detectable change from current status for a given lake and metric. These detectable changes (i.e., changes of a given mean) were estimated for 15 levels for each metric (Table 2). From the chosen maximum change, each of the subsequent 14 levels in change represented a 23.16% improvement in precision and consequently more sampling effort. Maximum change in vegetation index was set at 1. A change in index of 1 would mean that the index status shifted to a different category. Maximum change for remaining metrics corresponded to the approximate range in values observed in the eight southern Michigan Lakes. Changes similar to the range would represent either a dramatic increase or decrease in mean for a given metric. Consequently, the maximum count of: submerged trees was set at 4; dwellings was set at 10; docks was set at 6; and percentage of armored shoreline was set at 50%.

Underestimation of sample size from equation (4) was corrected following tables and methods given in Kupper and Hafner (1989). The finite population correction term was not included in individual lake variance equations.

Power

This analysis was done for the eight southern Michigan lakes only. Each lake was compared to the remaining group of lakes. Thus, a group of lakes was considered a single stratum, and strata means and variances were estimated by analysis of variance. Means and variances from strata were assumed to be representative of lakes with similar physical or chemical characteristics. Metrics for individual lakes were then compared with the appropriate stratum mean and variance. This method provides estimates of power for comparisons of some lake to a group of lakes within a stratum (Hoenig and Heisey 2001). This stratification methodology and comparisons follow the Michigan Department of Natural Resources' Lakes Status and Trends Program design.

Potential power (1-β) for each lake l and metric m was estimated using PASS (power analysis and sample size) software (NCSS, version 6.0). For the lake strata, the mean square error term from one way analysis of variance was used to estimate the standard deviation for each metric. The mean square error term provides a satisfactory estimate of the variance remaining after the fixed effects are removed. For individual lakes, standard deviation was estimated for each metric as for independent samples. For all power analyses, $\alpha = 0.05$ and $\beta = 0.20$ was considered adequate protection against Type II Error. Power evaluation for each metric was for comparison of an individual lake to a group of lakes within a stratum using a t-test.

Results–

Eight Southern Michigan Lakes

Shoreline development index (SDI) varied from 1.18 at Round Lake to 2.11 at Half-Moon Lake (Table 1). Mean vegetation index for individual lakes varied from 1.00 at Round Lake to 5.00 at Chenango Lake (Table 3). Vegetation index was not measured at Sand Lake. Five of seven lakes had mean vegetation index <2. Mean number of submerged trees varied from 0.33 at Joslin Lake to 4.00 at Crooked Lake. Submerged trees were not counted at Sand Lake. Three of seven lakes had less than 1 tree per 1,000 ft transect. Shoreline dwellings varied from 0.20 per 1,000 ft at Chenango Lake to 13.74 at Sand Lake. Four of eight lakes had more than 10 dwellings per 1,000 ft of shoreline. Mean number of docks per 1,000 ft of shoreline varied from 0.40 at Chenango Lake to 12.39 at Sand Lake. Six of eight lakes had mean dock counts ≥ 5. Percentage of armored shoreline varied from 0.00% at Chenango Lake to 63.46% at Wamplers Lake. Two of eight lakes had more than 50% of their shoreline armored.

Reliability of means

For all metrics, greatest improvement in reliability of means occurred between 1 and 5 transects sampled (Figure 1). N = five transects sampled occurred at the point of each line adjacent to the area of greatest slope. Only minimal improvements occurred when more than 10 transects were sampled.

Vegetation index required least sampling effort to provide reliable estimates of means. With 5 transects sampled mean CV was 15.4%, and by 30 transects sampled CV decreased to 6.1%. Submerged tree counts were the least reliable metric with small number of transects sampled. With 5 sections sampled, mean CV was 68.0% and decreased to 28.0% with 30 transects sampled.

Dwelling and dock counts, and percentage of shoreline armored produced near identical results. With 5 transects sampled mean CV values were 38.5% (dwellings), 38.6% (docks), and 43.3% (armoring). Similarly, with 30 transects sampled mean CV values were 15.4% (dwellings), 15.5% (docks), and 17.3% (armoring).

Precision

Index of vegetation was evaluated at 7 of 8 study lakes. Maximum change (±1) was attainable for all lakes with 9 sections counted (Figure 2). Greatest reduction in variability was achieved with 20 transects sampled. Recall that Round Lake had vegetation index of 1.00 for all transects and Chenango Lake had vegetation index of 5.00 for all transects. With 20 transects sampled, detectable change varied from <±0.27 to ±0.59 and <±0.20 to ±0.46 with 30 transects sampled. Greatest improvement in precision was achieved with 30 sections counted.

Submerged trees were counted at 7 of 8 lakes. Maximum change (±4) was attainable for all lakes with 5 sections counted (Figure 3). Detectable change varied from ±0.17 to ±2.05 with 20 sections counted and ±0.13 to ±1.70 with 30 sections counted. Greatest improvement in precision was achieved with 20 sections counted.

Dwelling counts were fairly consistent for seven of the eight lakes sampled (Figure 4). Maximum change (±10) was attainable for all eight lakes with 2 sections counted. Detectable change varied from ±0.25 to ±4.54 with 12 sections sampled, <±0.25 to ±3.80 with 20 sections sampled, and

<±0.25 to ±2.90 with 30 sections sampled. Greatest improvement in precision was achieved with 15 sections counted.

Dock counts were made at all eight lakes and were most precise for Chenango Lake (Figure 5). Precision of remaining seven lakes were quite similar to one another. Maximum change (±6) was attainable for all lakes with 5 sections counted. Detectable change varied from ±0.38 to ±3.25 with 20 sections counted and ±0.30 to ±2.65 with 30 sections sampled. Greatest improvement in precision was achieved with 20 sections counted.

Percentage of shoreline armoring was estimated for all eight lakes (Figure 6). However, no armoring was observed at Chenango Lake (all counts were 0.00%) and, consequently, the SD was 0.00%. Maximum change (±50%) was attained for all eight lakes with 3 sections counted. Detectable change varied from <±7.0% to ±20.0% with 20 sections counted and <±6.0% to ±16.1% with 30 sections counted. Greatest reduction in variation occurred with 20 sections counted.

Power

For index of vegetation, substantial increases in power with additional transects sampled were realized for three of eight lakes sampled (Figure 16). Power of 0.80 was achieved by sampling 15 transects at Joslin, Half-Moon, and Strawberry lakes. Two lakes, Round and Chenango, had estimated variances of 0.00 and consequently power was 1.00 for one or more transects sampled. The remaining two lakes, Wamplers and Crooked, had mean indices similar to the group means and had substantial variation. Consequently, power was quite low, ≤ 0.15, with 30 transects sampled.

Tree count power increased substantially with minimal increases in transects sampled for four of seven lakes (Figure 8). Power ≥ 0.80 was achieved for Round, Joslin, Crooked, and Wamplers lakes with 10 transects sampled. One lake, Chenango, required ~17 transects sampled for power of 0.80. A sixth lake, Half-Moon, required ≥ 35 transects sampled to achieve power of 0.80. Mean tree count for the seventh lake, Strawberry, was quite variable, CV was 207% (Table 3), and consequently power was quite low with a substantial (>80 transects) number of transects sampled.

Dwelling counts were made at all eight lakes (Figure 9). One lake, Chenango, had power of 1.00 with one or more transects counted. Three lakes, Crooked, Sand, and Half-Moon, had power ≥ 0.80 with 15 transects counted. Substantial sampling effort was required to attain power of 0.80 at three other lakes. At Round, Wamplers, and Strawberry lakes 50 transects were needed to attain power of 0.64 – 0.80. The remaining lake, Joslin, had mean dwelling count (9.44) almost identical to the cluster mean of 9.96 (Tables 3 and 4), and power was low even with 80 transects sampled.

Dock counts were made at all eight lakes (Figure 10). Chenango and Crooked lakes had power ≥ 0.80 with five transects sampled. Sand Lake had power >0.80 with 15 transects sampled. Half-Moon had power >0.80 with 25 transects sampled. Round and Wamplers lakes required 55 transects to achieve power ≥ 0.80. Both Joslin and Strawberry lakes had low power with >80 transects sampled.

Percentage of shoreline armoring was estimated at all eight lakes (Figure 11). Chenango had no shoreline armoring and consequently the variance was 0.0 and power was 1.0. Sand and Half-Moon lakes each required <10 transects sampled to attain power ≥ 0.80. Wamplers Lake required <20 transects sampled and Round Lake >30 transects sampled to attain power ≥ 0.80. Sand, Strawberry, and Joslin lakes each required >80 transects to attain power ≥ 0.80.

Three Large Lakes and Three River Impoundments

Shoreline development index (SDI) was more varied for this group. For the three lakes, SDI varied from 1.5 at Houghton Lake to 4.26 at Lobdell Lake (Table 1). Shoreline development index was greater for all impoundments and varied from 3.30 at Croton to 8.59 at Moores Park. Number of 1,000 ft transects sampled per lake varied from 45 at Missaukee Lake to 145 at Houghton Lake. Similarly, number of transects sampled at the impoundments varied from 65 at Croton to 91 at Moores Park.

Vegetation index was not collected for any of the 3 large lakes or the 3 impoundments (Table 3).

Mean number of submerged trees varied from 0.09 at Missaukee Lake to 1.09 at Lobdell Lake. Houghton Lake was similar to Missaukee Lake and had a mean tree count of 0.10. Submerged trees were not counted at Moores Park. Croton had 3.37 trees per 1,000 ft of shoreline and Kent had 4.70. Shoreline dwellings varied from 10.69 per 1,000 ft at Missaukee Lake to 15.47 at Houghton Lake. Lobdell Lake was similar to Missaukee Lake with 10.81 dwellings per 1,000 ft of shoreline. For the impoundments, Kent had the fewest number of dwellings (0.01) and Croton the greatest (6.80). Mean number of docks per 1,000 ft of shoreline varied from 6.98 at Missaukee Lake to 11.53 at Lobdell Lake. Houghton Lake was similar to Lobdell Lake with 10.38 docks per 1,000 ft of shoreline. Croton had greatest number of docks (7.03) for any of the impoundments, while Kent and Moores were similar at 0.21 and 1.50, respectively. Percentage of armored shoreline varied from 25.67% at Missaukee Lake to 86.69% at Houghton Lake. Lobdell Lake was similar to Houghton Lake with 72.23% of shoreline armored. For the impoundments, Kent had the least percentage of its shoreline armored (14.00%) and Croton the greatest (43.69%). Moores Park had 30.25% of its shoreline armored.

Precision

Index of vegetation was not sampled at any of the three large lakes or three impoundments.

Submerged tree counts were completed at all 3 large lakes and maximum change in submerged tree count (±4) was attained with 1 section counted (Figure 12). Detectable change varied from ±0.29 to ±1.10 with 10 sections counted, ±0.20 to ±0.72 with 20 sections counted, and ±0.15 to ±0.62 with 30 sections counted. Greatest improvement in precision was achieved with 10 sections counted. Submerged tree counts were completed at 2 of the impoundments and maximum change (±4) was attained with 26 sections counted (Figure 13). Detectable change varied from ±2.7 to >±4.0 with 20 sections counted and ±2.2 to ±3.8 with 30 sections counted. Greatest improvement in precision was achieved with 30 sections counted.

Dwelling counts were made at all 3 large lakes and maximum change (±10) in dwelling count was attained with 3 sections counted (Figure 14). Detectable change varied from ±2.25 to ±4.00 with 20 sections counted and ±1.80 to ±3.25 with 30 sections counted. Greatest improvement in precision was achieved with 20 sections counted. With the exception of Kent, precision of impoundments was similar to the 3 large lakes (Figure 15). Dwellings were exceptionally rare at Kent (mean=0.01) and precision estimates indicated only 1 section sampled would yield change in count of ±0.4 dwellings. For Croton and Moores Park, maximum change (±10) in dwelling counts was attained with 1 section counted. Detectable change of ±1.7 to ±2.6 with 20 sections counted and ±1.5 to ±2.1 with 30 sections counted. Greatest improvement in precision was achieved with 15 sections counted.

Dock counts were made at all 3 large lakes and maximum change (±6) was attained with 4 sections counted (Figure 16). Detectable changes of ±2.25 to ±3.00 were attained with 20 sections counted and ±1.80 to ±2.45 with 30 sections counted. Greatest improvement in precision was achieved with 25 sections counted. For the impoundments, Kent required the least sampling effort, while Moores Park and Croton required additional sampling effort (Figure 17). For all impoundments, maximum change (±6) was attained with 3 sections counted. Kent required 10 sections counted to attain precision of ±0.43 with minimal improvements in precision when additional sections are counted. For Croton and Moores Park, detectable changes varied from ±2.4 to ±2.6 with 20 sections counted and ±1.9 to ±2.1 with 30 sections counted. Greatest improvement in precision was achieved with 25 sections counted.

Percentage of shoreline armored was evaluated at all 3 large lakes and maximum change (±50%) was attained with 2 sections counted (Figure 18). Detectable changes of ±12% to ±17% were attained with 20 sections counted and ±9% to ±14% with 30 sections counted. Greatest improvement in precision was achieved with 20 sections counted. Precision of percentage of shoreline armored was similar at the 3 impoundments and maximum change (±50%) was attained with 2 sections counted (Figure 19). Detectable change varied from ±12% to ±18% with 20 sections counted and ±10% to ±15% with 30 sections counted. Greatest improvement in precision was achieved with 20 sections counted.

Analysis–The purpose of Fisheries Division's Status and Trends Program is to characterize metrics associated with classes of lakes. These characterizations will allow statewide measures of status and trends, and provide a management framework for comparisons of individual lakes to appropriate clusters sharing physical and chemical characteristics.

Essential to any sampling program is the accurate measure of the metrics used to create classification strata. Accurate measures rely on three key components. First, selected metrics must be appropriate measures of lake physical or chemical attributes. The metrics used in this study have been shown to be important in determining the ecological potential of inland lakes (Jennings et al. 1999, Schindler et al. 2000, Rust et al. 2002). For example, submerged trees and vegetation cover can influence fish survival and growth by providing critical feeding areas that offer refugia from predation. Because human development can adversely affect water quality and the quantity of nearshore habitat, measures of the intensity of human activity (e.g., number of dwellings and docks) should help explain biotic and abiotic differences observed among lakes.

Second, any sampling program must provide randomly selected samples of a given metric. Non-random samples, such as quota or convenience samples, are problematic in that: they give measures that are not representative of the population being estimated (Peterson et al. 1999); they do not provide measures that may be expanded to a larger area (Thompson et al. 1998); the extent of bias (that is, error) introduced is not known (Snedecor and Cochran 1989); and the use of basic statistical formulas, theory, and inferences are not applicable (Remington and Schork 1970, Ferguson 1976, Cochran 1977). Pollock et al. (1994) noted that:

> "A sampling procedure must be consistent with sound statistical principles or it will be impossible to establish the properties of the estimators obtained from the sample in terms of bias, precision, and accuracy. Samples drawn subjectively to cut costs or to be vaguely representative are useless."

Third, a sample must be large enough to accurately characterize the true population metric (Wiley et al. 1997). Results of bootstrapping estimates of coefficient of variation indicate that a minimum of 5 randomly selected 1,000 ft transects must be sampled to characterize each of the metrics sampled. Sampling fewer than 5 transects will provide unreliable estimates of each metric and would certainly produce unreliable estimates of the variance associated with each (Wiley et al. 1997).

Similar to the importance of accuracy of a sample, is the measure of variation associated with that sample. For any estimate to be useful, the proportional variation of that estimate must be efficient given some optimal level of sampling effort. Variability of an estimate may be due to clustering of individuals or rarity of individuals. Stratification of sampling will effectively reduce estimate variability but requires a priori data and stratification by one metric may not be appropriate for another (e.g., Lockwood 1999). Estimates of rarely occurring individuals typically require substantial sampling effort or specialized sampling techniques such as adaptive cluster sampling (Thompson 1998). However, even imprecise estimates of rare individuals indicate their rareness and slight (even doubling) changes may not be noteworthy. For most lake-metric estimates, greatest reduction in variability occurred with ~20 1,000 ft transects sampled.

Estimated precision of the three large lake and 3 impoundment metrics, and consequently sample size, followed that of the eight southern Michigan lakes. Sample size methods and assumptions for the eight southern Michigan lakes appear to be quite robust and applicable across Michigan inland lakes.

The Status and Trends Program utilizes classification of lakes to monitor statewide lake conditions, and to make comparisons of individual lakes to an appropriate cluster of lakes for localized management purposes. Thus for an individual lake, the estimates from one lake will be compared to a cluster of lakes sharing similar characteristics. Estimated power (for this study, α=0.05 and β=0.20) to say that no difference exists between a lake metric and that metric for a cluster of lakes relies on the variability of the lake and the variability of the cluster. Variability of each is a function of sample size. Typically, precision (estimate ±a proportion of the estimate) of an individual lake will be less than that of the cluster. Individual lake sample sizes will also be less than the cluster sample size. Estimated power $\geq$ 0.80 was attained for at least half of the individual lake to cluster comparisons with 20 1,000 ft transects sampled for all metrics except dock counts. Dock counts required 30 transects to attain power $\geq$ 0.80 for half of the individual lake to cluster comparisons.

Estimates of power to detect individual lake differences from strata were markedly different for each metric. Power was at or near 1.0 with minimal number of samples for some lake-to-strata comparisons, while others required sample sizes beyond the scope of the status and trends program.

For large lakes where only a portion of the lakeshore may be sampled, inclusion of finite population correction term (fpc) may reduce estimated variability of each metric and, consequently, sample size. The lakes reported in this study were intended to be representative of Michigan lakes in general, many of which are of substantial size; sampling of more than 10% of their shoreline may not be feasible, and inclusion of fpc would have negligible effect. Similarly, determination of sample sizes requires balancing time constraints with statistical utility of the estimates. Within appropriate time limitations then, common sense dictates conservative estimates of sample size. Ultimately, inclusion of fpc would enhance evaluations of the lake status and trends metrics included in this report.

For this study, 1,000 ft shoreline sections were sampled to estimate each metric and handheld GPS units were used to measure each section. Two sources of potential edge effect error were considered and the shoreline section length was scaled to minimize each error source. First was the error introduced by the GPS unit. Manufacturer's documentation indicated that the GPS units introduce on average 49 ft of measurement error. Since this GPS measurement error can occur on each end of the shoreline section, GPS measurement error is doubled. Thus for 1,000 ft of shoreline sampled, this would represent a 9.8% edge effect error rate. This edge effect error decreased as shoreline section length increases. Greatest reduction in error was realized at shoreline section length of 400 ft. Reduction in GPS introduced error was 1% per 100 ft increase in shoreline section at 1,000 ft and only minimal reductions in error were realized with additional shoreline section length. Consequently, GPS measurement error was considered inconsequential by 1,000 ft. Second was the judgment error associated with including or excluding a metric unit in a count section. For example, typical lake lot sizes on Michigan inland lakes are 60 to 100 ft in width. For 60 ft lots, only 12% of lots in a 1,000 ft shoreline section would occupy edges, and for 100 ft lots 20% would occupy edges. Similar to GPS-introduced error, greatest reduction in error was realized by shoreline section of 500 ft. Reduction in error per 100 ft of shoreline was 1.5% at 1,000 ft with only minimal decreases for additional 100 ft increments.

Recommendations–Estimated mean for each of the metrics proved reliable with 10 transects sampled. Evaluations of precision and power for each metric indicated variation of estimates stabilized with 20-30 transects sampled per lake. Shoreline sampling for most lakes takes a minimal amount of time. Wamplers Lake with 29 transects sampled required ~1 h. Therefore, the following recommendations are made:

- All shoreline transects should be 1,000 ft in length
- All metric information should be recorded for each 1,000 ft transect
- For lakes with circumference ≤ 30,000 ft the entire shoreline should be sampled
- For lakes with circumference >30,000 ft, sample 30 transects starting from a randomly selected shoreline position with necessary spacing between transects to end equidistance from the starting location.

Example

A lake is 53,200 ft in circumference. Randomly pick an integer between 1 and 53,200. Suppose that integer is 1,152. With 30,000 ft to be sampled, 23,200 ft will be unsampled. Dividing 23,200 ft by 30 gives 733.3 ft of unsampled shoreline between each 1,000 ft section. For this example then: randomly select direction of travel from the launch site (e.g., coin toss to determine clockwise or counter clockwise direction of travel); from the launch site travel 1,152 ft, sample 1,000 ft of shoreline, skip 733.3 ft of shoreline, sample 1,000 ft of shoreline, continue until all 30 transects have been completed.

Note: all 1,000 ft transects along a lake's shoreline may be sampled (even though circumference is >30,000 ft) if time is available and/or the crew determines that it is easier or more efficient to sample the entire shoreline rather than a portion of it. Length of the last section sampled will no doubt ≠ 1,000 ft. In this situation, record the length of the last section for each of the metrics.

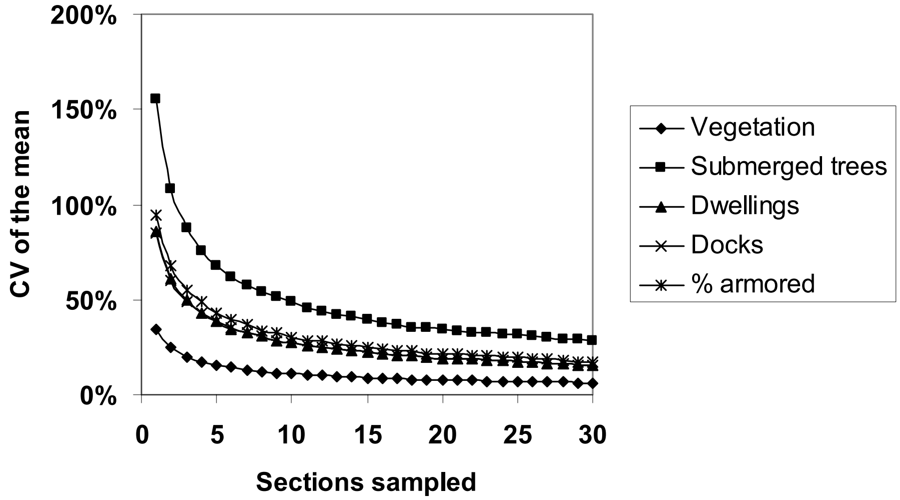

Figure 1.–Bootstrap estimates of coefficient of variation of the mean ($CV_{\bar{x}}$) for two habitat and three shoreline development metrics. $CV_{\bar{x}}$ for each metric and sections sampled were averaged across lakes to provide relative measures of accuracy and variability. Sections sampled represent number of 1,000 foot transects sampled.

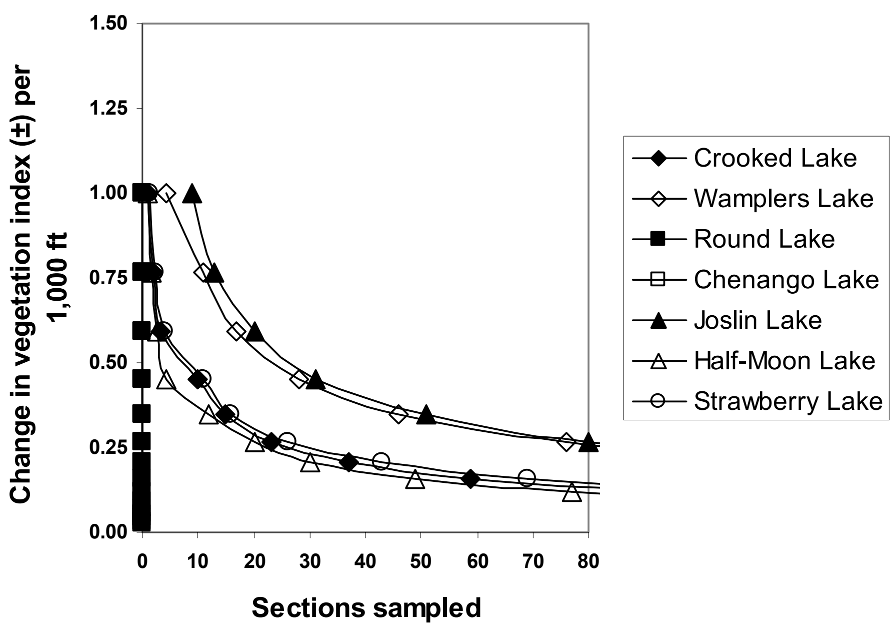

Figure 2.–Detectable change in mean vegetation index by sections sampled for each study lake. Estimates of precision are for α=0.05.

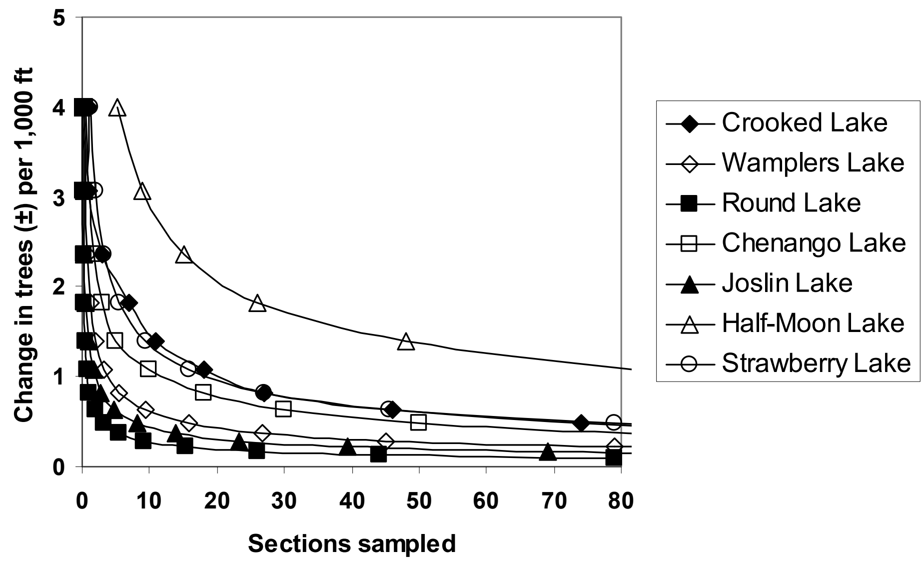

Figure 3.–Detectable change in mean submerged tree counts by sections sampled for each study lake. Estimates of precision are for α=0.05.

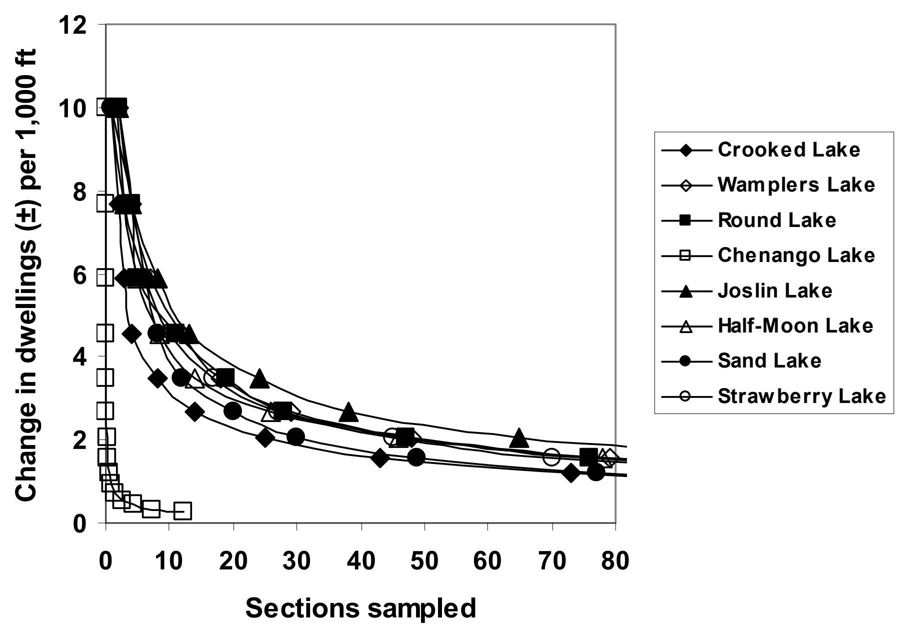

Figure 4.–Detectable change in mean dwelling counts by sections sampled for each study lake. Estimates of precision are for α=0.05.

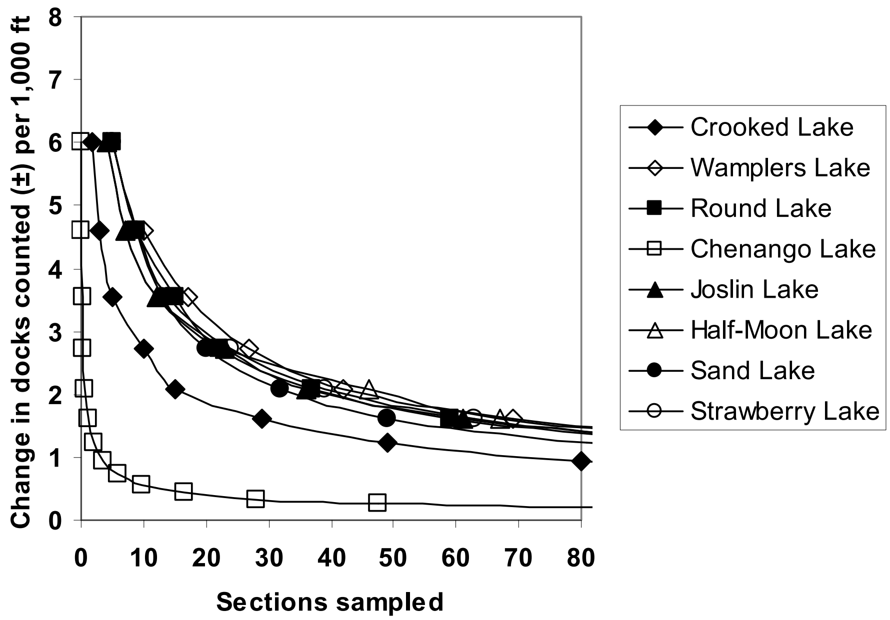

Figure 5.–Detectable change in mean dock counts by sections sampled for each study lake. Estimates of precision are for α=0.05.

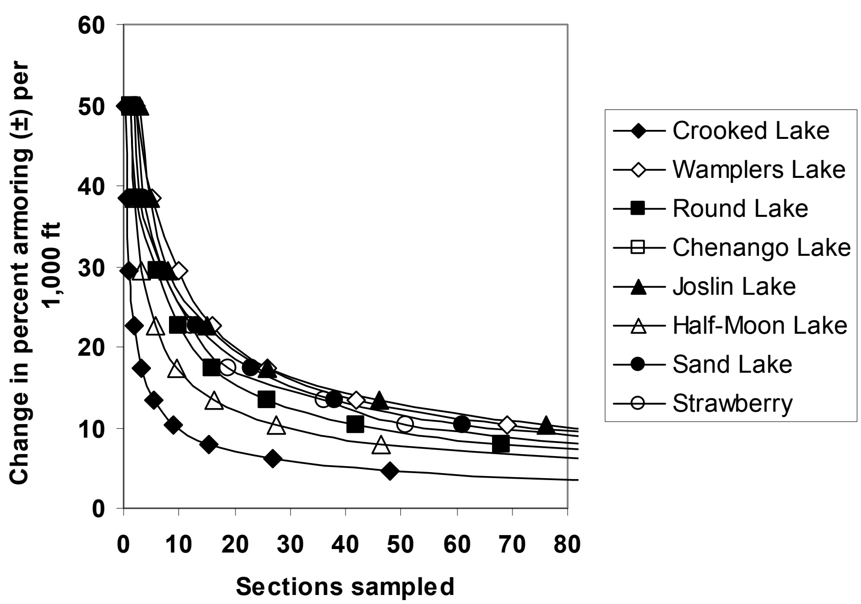

Figure 6.–Detectable change in mean shoreline armoring counts by sections sampled for each study lake. Estimates of precision are for α=0.05.

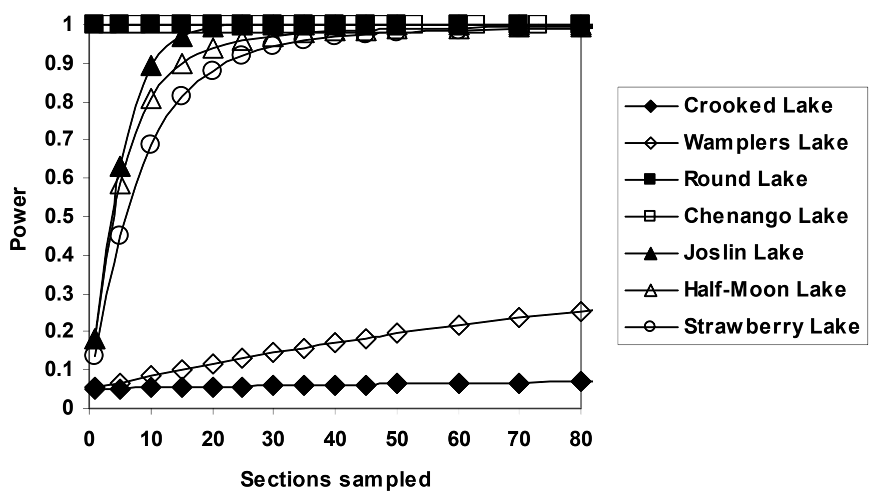

Figure 7.–Power of mean index of vegetation by sections sampled for each study lake relative to all other study lakes. Comparisons were made for a t-test with α=0.05.

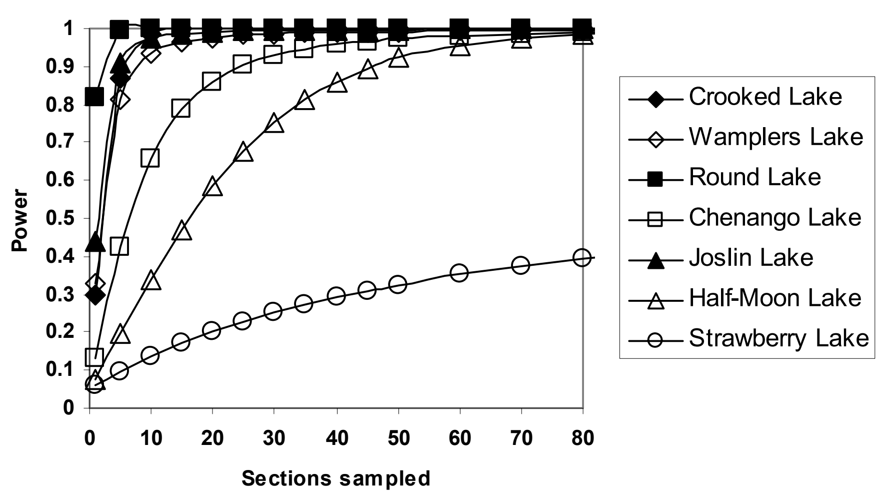

Figure 8.–Power of mean submerged tree count by sections sampled for each study lake relative to all other study lakes. Comparisons were made for a t-test with α=0.05.

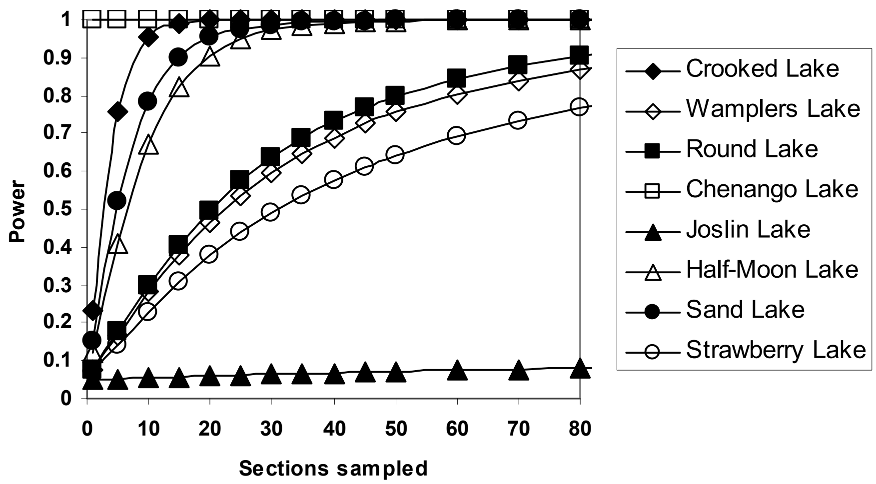

Figure 9.–Power of mean dwelling count by sections sampled for each study lake relative to all other study lakes. Comparisons were made for a t-test with α=0.05.

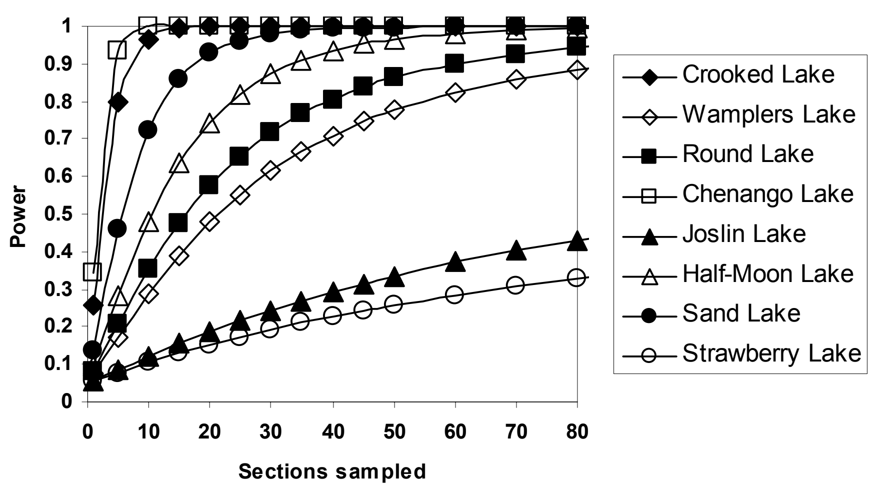

Figure 10.–Power of mean dock count by sections sampled for each study lake relative to all other study lakes. Comparisons were made for a t-test with α=0.05.

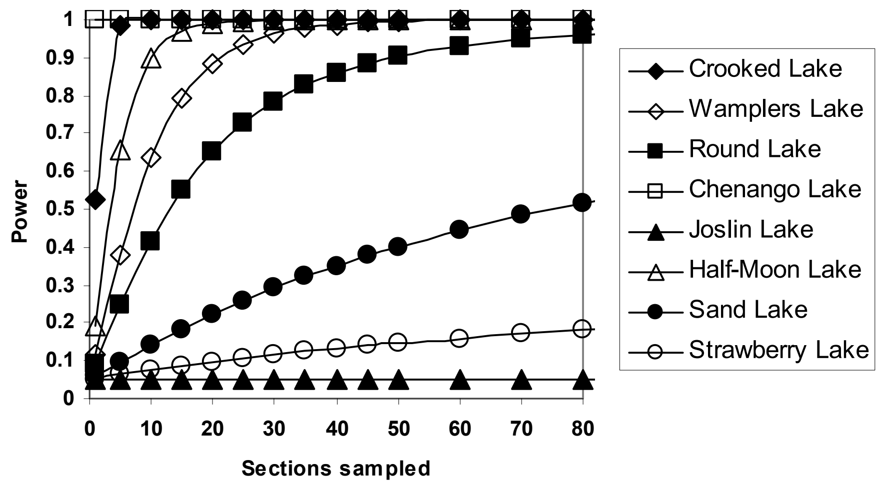

Figure 11.–Power of mean percent shoreline armoring by sections sampled for each study lake relative to all other study lakes. Comparisons were made for a t-test with α=0.05.

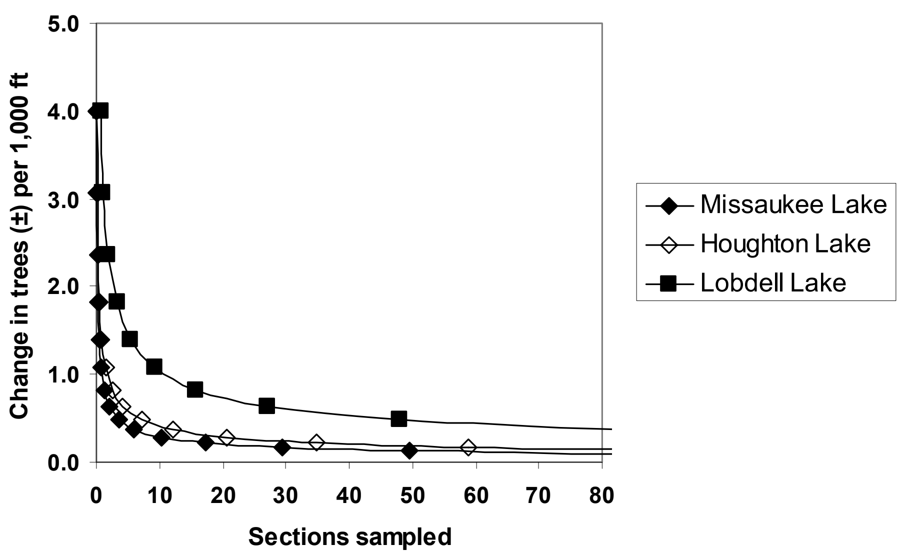

Figure 12.–Detectable change in mean submerged tree counts by sections sampled for each of the 3 large lakes. Estimates of precision are for α=0.05.

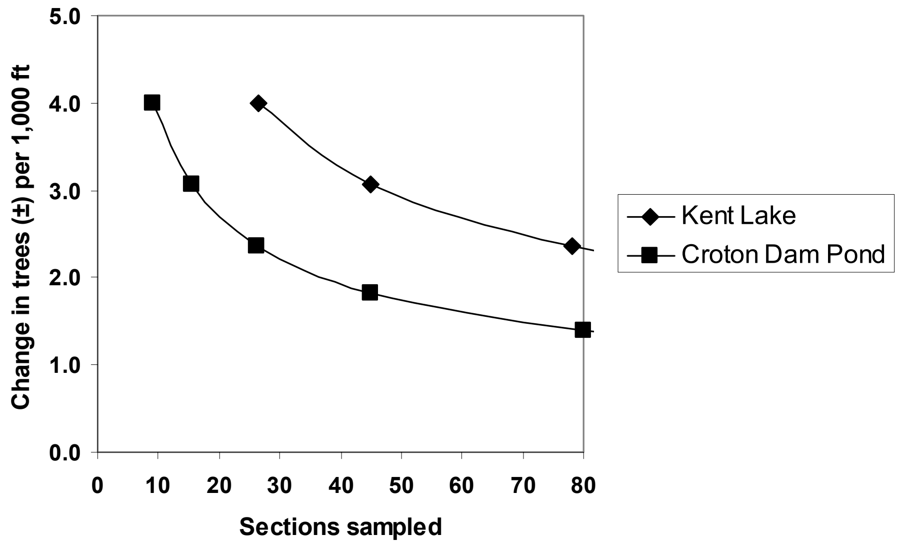

Figure 13.–Detectable change in mean submerged tree counts by sections sampled for 2 of the impoundments. Estimates of precision are for α=0.05.

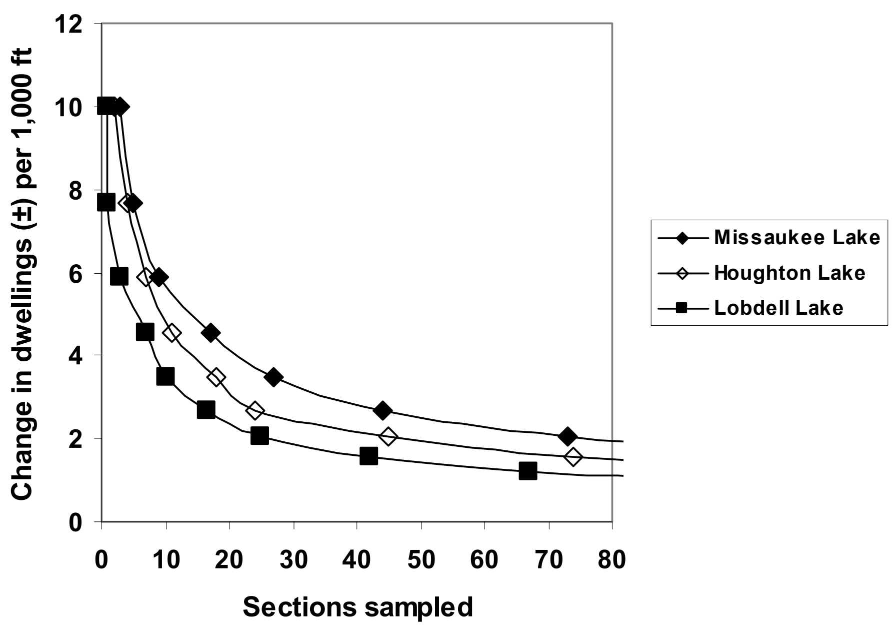

Figure 14.–Detectable change in mean dwelling counts by sections sampled for each of the 3 large lakes. Estimates of precision are for α=0.05.

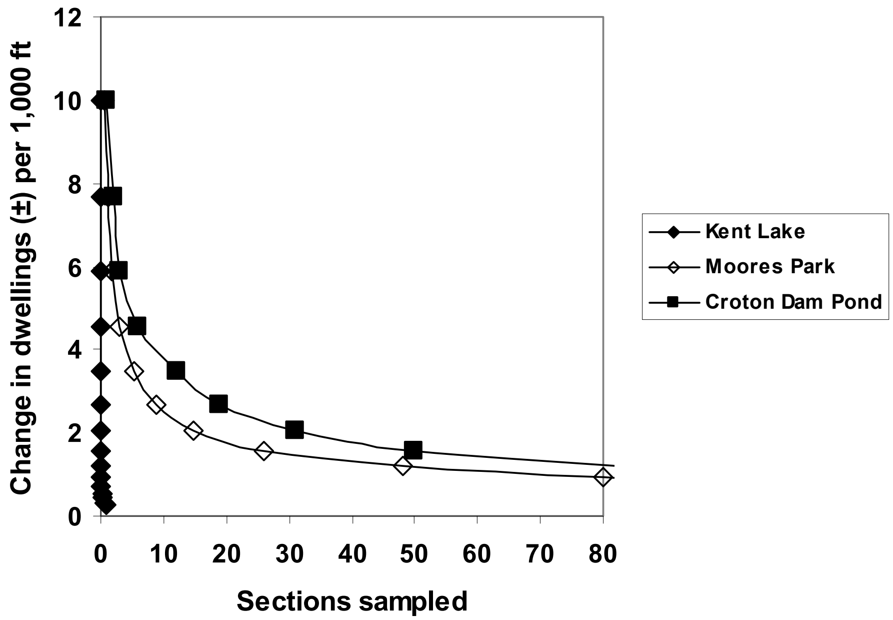

Figure 15.–Detectable change in mean dwelling counts by sections sampled for each of the 3 impoundments. Estimates of precision are for α=0.05.

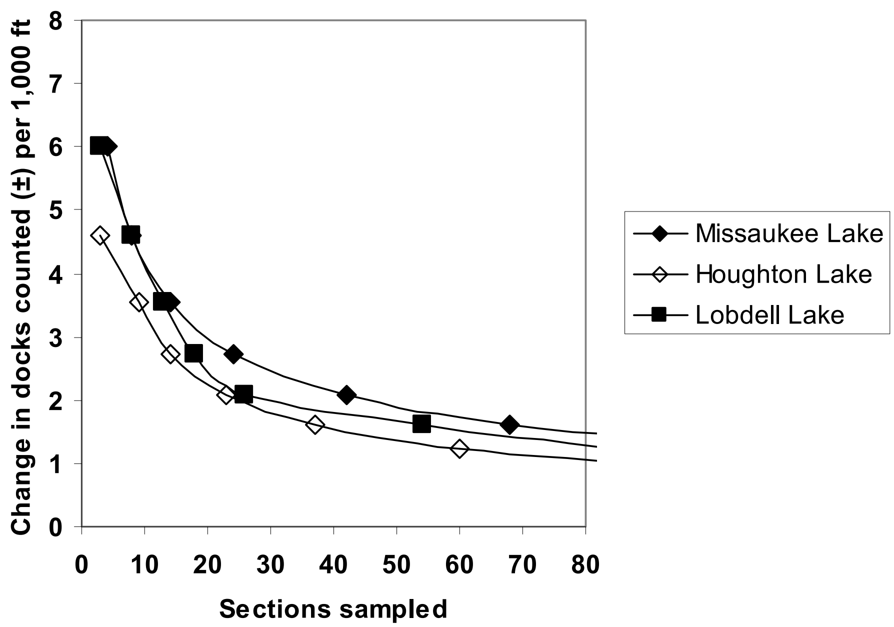

Figure 16.–Detectable change in mean dock counts by sections sampled for each of the 3 large lakes. Estimates of precision are for α=0.05.

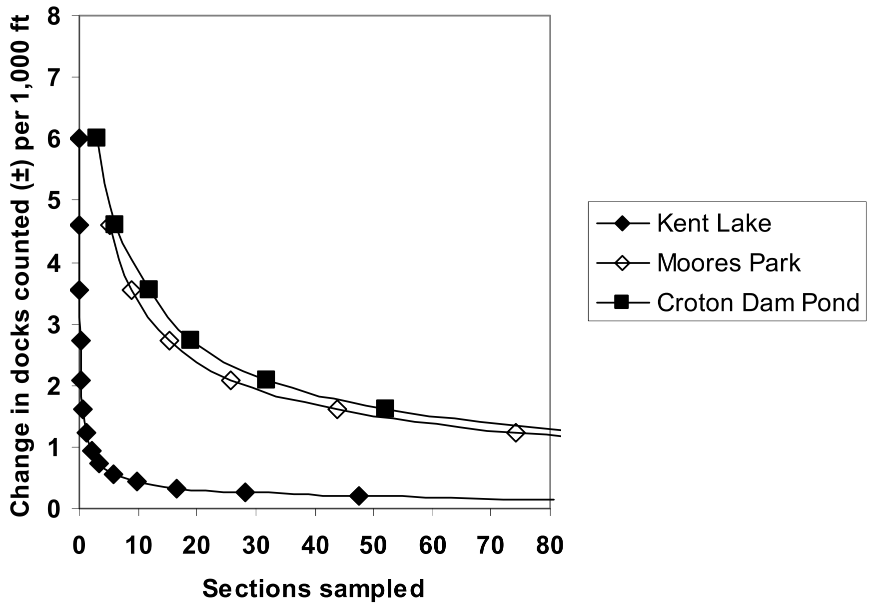

Figure 17.–Detectable change in mean dock counts by sections sampled for each of the 3 impoundments. Estimates of precision are for α=0.05.

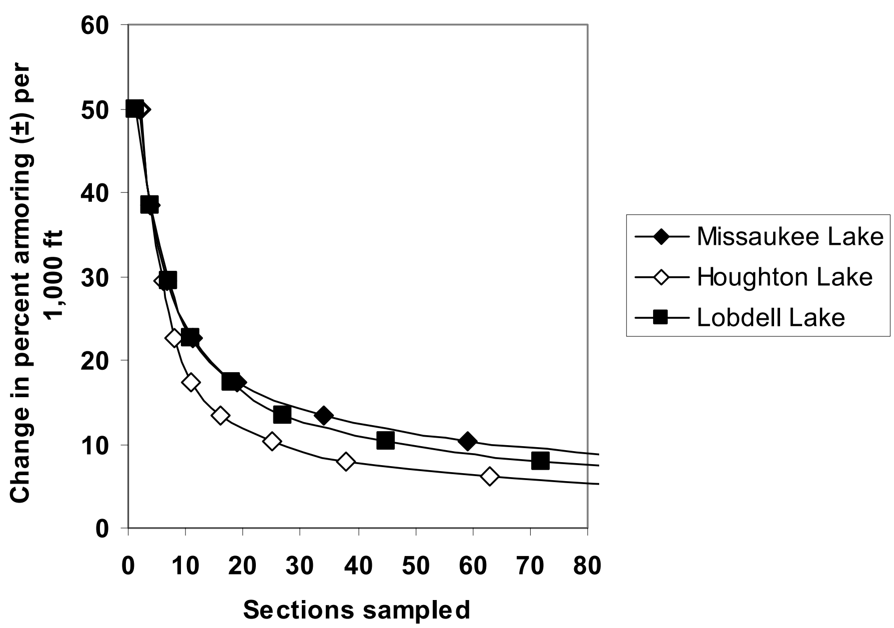

Figure 18.–Detectable change in mean shoreline armoring counts by sections sampled for each of the 3 large lakes. Estimates of precision are for α=0.05.

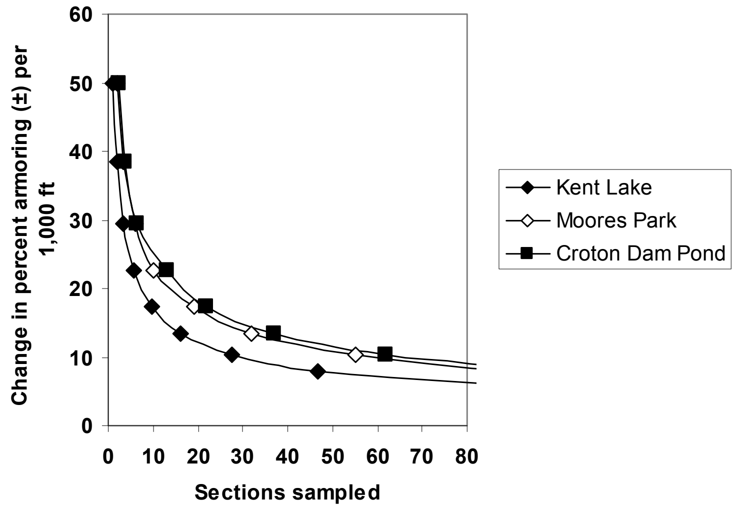

Figure 19.–Detectable change in mean shoreline armoring counts by sections sampled for each of the 3 impoundments. Estimates of precision are for α=0.05.

Table 1.–Lake name, location, size, circumference, shoreline development index (SDI), and number of 1,000 foot sample transects.

Lake name	Latitude	Longitude	Size (acres)	Circumference (feet)	SDI	Transects
Eight Southern Michigan Lakes						
Crooked	42.32491	84.11448	113	13,744	1.75	11
Wamplers	42.07237	84.14054	797	34,447	1.65	26
Round	42.08536	84.46997	151	10,735	1.18	9
Chenango	42.50591	83.89668	29	5,584	1.40	5
Joslin	41.41425	84.06902	194	12,648	1.23	9
Half-moon	42.41597	84.00462	241	24,276	2.11	18
Sand	42.04952	84.12685	546	29,247	1.69	19
Strawberry	42.45205	83.84898	261	23,567	1.97	18
Three Large Lakes						
Houghton	44.31977	84.76460	20,075	157,251	1.50	145
Lobdell	42.78504	83.83321	546	73,622	4.26	53
Missaukee	44.33978	85.22469	1,880	61,027	1.83	45
Three Impoundments						
Croton (Muskegon River)	43.43955	85.66593	1,129	82,014	3.30	65
Kent (Huron River)	42.52374	83.64515	1,015	128,097	5.43	80
Moores Park (Grand River)	42.72114	84.58642	350	118,901	8.59	91

Table 2.–Levels of detectable change for estimating sample sizes for two habitat and three shoreline development metrics. Each change, for individual metrics, represents a 23.16% improvement in precision from previous value.

	Vegetation index	Submerged trees	Dwellings	Docks	Percentage of shoreline armored
Maximum change	1.0000	4.0000	10.0000	6.0000	50.0000
	0.7684	3.0737	7.6842	4.6105	38.4209
	0.5905	2.3619	5.9047	3.5428	29.5233
	0.4537	1.8149	4.5372	2.7223	22.6862
	0.3486	1.3946	3.4865	2.0919	17.4325
	0.2679	1.0716	2.6791	1.6074	13.3954
	0.2059	0.8235	2.0587	1.2352	10.2933
	0.1582	0.6328	1.5819	0.9491	7.9095
	0.1216	0.4862	1.2156	0.7293	6.0778
	0.0934	0.3736	0.9341	0.5604	4.6703
	0.0718	0.2871	0.7177	0.4306	3.5887
	0.0552	0.2206	0.5515	0.3309	2.7576
	0.0424	0.1695	0.4238	0.2543	2.1190
	0.0326	0.1303	0.3257	0.1954	1.6283
Minimum change	0.0250	0.1001	0.2502	0.1501	1.2512

Table 3.–Empirical data for each habitat and development metric. Statistics for each include: mean, standard deviation (SD), and number of 1,000 foot transects sampled (N).

Sample year	Name	Statistics	Vegetation (%)	Submerged trees	Shoreline dwellings	Docks	Armored shore (%)
			Eight Southern Michigan Lakes				
2002	Crooked	Mean	1.70	4.00	4.45	3.91	10.91
		SD	1.06	1.90	4.91	4.06	15.78
		N	10	11	11	11	11
2003	Wamplers	Mean	1.61	0.42	12.19	10.92	63.46
		SD	1.08	0.99	6.43	6.15	39.29
		N	23	26	26	26	26
2002	Round	Mean	1.00	0.22	12.67	11.44	56.67
		SD	0.00	0.44	6.30	5.62	30.00
		N	9	9	9	9	9
2003	Chenango	Mean	5.00	2.00	0.20	0.400	0.00
		SD	0.00	1.58	0.45	0.89	0.00
		N	5	5	5	5	5
2003	Joslin	Mean	2.78	0.33	9.44	7.33	40.00
		SD	1.09	0.71	7.68	5.85	42.72
		N	9	9	9	9	9
2002/03	Half- Moon	Mean	1.33	3.44	5.44	5.39	15.00
		SD	0.48	4.67	6.73	6.30	27.49
		N	18	18	18	18	18
2003	Sand	Mean	-	-	13.74	12.39	49.47
		SD	-	-	4.86	5.07	37.49
		N	0	0	19	18	19
2003	Strawberry	Mean	1.33	1.06	12.00	9.83	45.00
		SD	0.59	2.18	6.10	5.82	34.17
		N	18	18	18	18	18
			Three Large Lakes				
2004	Houghton	Mean	-	0.10	15.47	10.38	86.69
		SD	-	0.66	6.12	4.26	20.98
		N	-	145	145	145	145
2004	Lobdell	Mean	-	1.09	10.81	11.53	72.23
		SD	-	1.67	4.47	5.29	30.44
		N	-	53	53	53	53
2004	Missaukee	Mean	-	0.09	10.69	6.98	25.67
		SD	-	0.47	8.25	6.29	38.75
		N	-	45	45	45	45

Table 3.–Continued.

Sample year	Name	Statistics	Vegetation (%)	Submerged trees	Shoreline dwellings	Docks	Armored shore (%)
			Three Impoundments				
2004	Croton	Mean	-	3.37	6.80	7.03	43.69
		SD	-	6.17	5.21	5.42	38.22
		N	-	65	65	65	65
2004	Kent	Mean	-	4.70	0.01	0.21	14.00
		SD	-	10.50	0.11	0.69	27.54
		N	-	80	80	80	80
2004	Moores Park	Mean	-	-	2.99	1.50	30.25
		SD	-	-	4.05	5.43	36.62
		N	-	-	91	90	91

Table 4.–Pooled estimates of each habitat and development metric for eight southern Michigan lakes. Analysis of variance was used to calculate means and standard deviations (using mean square error term). Data for the lake excluded is not included in the pooled estimate. Statistics for each include: mean, standard deviation (SD), and sample size (N).

Sample year	Lake excluded	Statistics	Vegetation index	Submerged trees	Shoreline dwellings	Docks	Armored shore (%)
2002	Crooked	Mean	1.76	1.26	10.50	9.24	43.65
		SD	0.77	2.49	6.14	5.74	34.93
		N	81	85	104	103	104
2003	Wamplers	Mean	1.80	2.00	9.26	8.08	33.82
		SD	0.69	2.80	5.91	5.42	31.68
		N	69	70	89	88	89
2002	Round	Mean	1.83	1.71	9.69	8.50	39.15
		SD	0.85	2.55	6.01	5.60	33.89
		N	83	87	106	105	106
2003	Chenango	Mean	1.56	1.55	10.36	9.11	42.36
		SD	0.83	2.47	6.15	5.71	34.26
		N	87	91	110	109	110
2003	Joslin	Mean	1.64	1.70	9.96	8.85	40.57
		SD	0.77	2.54	5.88	5.58	32.76
		N	83	87	106	105	106
2002/03	Half- Moon	Mean	1.85	1.14	10.75	9.35	45.26
		SD	0.87	1.48	5.89	5.46	34.64
		N	74	78	97	96	97
2003	Sand	Mean	-	-	9.17	8.04	38.75
		SD	-	-	6.24	5.70	32.77
		N	-	-	96	96	96
2003	Strawberry	Mean	1.85	1.69	9.54	8.52	39.69
		SD	0.85	2.49	6.02	5.56	33.50
		N	74	78	97	96	97

Literature Cited:

Cochran, W. G. 1977. Sampling techniques. John Wiley & Sons, New York, New York.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York, New York.

Ferguson, G. A. 1976. Statistical analysis in psychology and education. McGraw-Hill Book Company, New York, New York.

Hayes, D., E. Baker, R. Bednarz, D. Borgeson, Jr., J. Braunscheidel, J. Breck, M. Bremigan, A. Harrington, R. Hay, R. Lockwood, A. Nuhfer, J. Schneider, P. Seelbach, J. Waybrant, and T. Zorn. 2003. Developing a standardized sampling program: the Michigan experience. Fisheries 28(7):18-25.

Hoenig, J. M., and D. M. Heisey. 2001. The abuse of power: the pervasive fallacy of power calculations for data analysis. The American Statistician 55:19-24.

Jennings, M. J., M. A. Bozek, G. R. Hatzenbeler, E. E. Emmons, and M. D. Staggs. 1999. Cumulative effects of incremental shoreline habitat modification on fish assemblages in north temperate lakes. North American Journal of Fisheries Management 19:18-27.

Kupper, L. L., and K. B. Hafner. 1989. How appropriate are popular sample size formulas? The American Statistician 43:101-105.

Lockwood, R. N. 1999. Evaluation of sampling techniques for the Lake Michigan angler survey. Michigan Department of Natural Resources, Fisheries Research Report No. 2050, Ann Arbor.

Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall, New York, New York.

Orth, D. J. 1983. Aquatic habitat measurements. Pages 61-84 *in* L. A. Nielsen and D. L. Johnson, editors. Fisheries techniques. American Fisheries Society, Bethesda, Maryland.

Peterson, S. A., N. S. Urquhart, and E. B. Welch. 1999. Sample representativeness: a must for reliable regional lake condition estimates. Environmental Science & Technology 33:1559-1565.

Pollock, K. H., C. M. Jones, and T. L. Brown. 1994. Angler survey methods and their applications in fisheries management. American Fisheries Society, Special Publication 25.

Remington, R. D., and M. A. Schork. 1970. Statistics with applications to the biological and health sciences. Prentice-Hall, Inc., New Jersey.

Rust, A. J., J. S. Diana, T. L. Margenau, and C. J. Edwards. 2002. Lake characteristics influencing spawning success of muskellunge in northern Wisconsin lakes. North American Journal of Fisheries Management 22:834-841.

Schindler, D. E., S. I. Geib, and M. R. Williams. 2000. Patterns of fish growth along a residential development gradient in North Temperate lakes. Ecosystems 3:229-237.

Snedecor, G. W., and W. G. Cochran. 1989. Statistical methods. Iowa State University Press, Iowa.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic press, Inc., New York.

Wiley, M. J., S. L. Kohler, and P. W. Seelbach. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. Freshwater Biology 37:133-148.

Prepared by: Roger N. Lockwood
Date: September 30, 2004