

TREE REGRESSION ANALYSIS ON THE NESTING HABITAT OF SMALLMOUTH BASS

CYNTHIA REJWAN,^{1,4} NICHOLAS C. COLLINS,¹ L. JERRY BRUNNER,² BRIAN J. SHUTER,³ AND MARK S. RIDGWAY³

¹*Biology Group, Erindale College, University of Toronto, Mississauga, Ontario L5L 1C6 Canada*

²*Statistics Department, Erindale College, University of Toronto, Mississauga, Ontario L5L 1C6 Canada*

³*Harkness Laboratory of Fisheries Research, Aquatic Ecosystem Science Section, Ontario Ministry of Natural Resources, P.O. Box 5000, Maple, Ontario L6A 1S9 Canada*

Abstract. Habitat conditions may determine locations of patchily distributed smallmouth bass nests in lakes, since young smallmouth bass are fragile and are therefore vulnerable to suboptimal physical characteristics in their typically variable nesting area, the littoral zone. Knowing which habitat conditions are important to nest locations would ultimately be useful in protecting optimal nesting areas from anthropogenic disturbances in lakes. To evaluate factors related to the nest distributions of a smallmouth bass (*Micropterus dolomieu*) population, physical habitat conditions were measured at 36 1-km-long and 31 100-m-long sites in Lake Opeongo (Ontario, Canada). Both tree regression analysis (a recently devised and unique statistical tool) and standard multiple regression were used to determine the relationships between nest density and four habitat variables. Tree regression analysis does not require assumptions of linearity or homoscedasticity of variances, and it automatically identifies interactions among variables. Furthermore tree regression results were more accurate and more precise than standard multiple-regression results. Mechanisms underlying the significant relationships between nest densities and both littoral zone temperatures and shoreline reticulation in 1-km-long sites and the non-significant results at the 100-m-long scale, are discussed. Cross-validation results quantify the difficulty in extrapolating sample findings to whole populations in ecological research.

Key words: cross validation of results; Lake Opeongo, Ontario, Canada; littoral zone; *Micropterus dolomieu*; multiple regression analysis, precision; nesting habitat of fish; shoreline reticulation; smallmouth bass; temperature sensitivity; tree regression analysis.

INTRODUCTION

Smallmouth bass are among numerous species that reproduce in the nearshore environment of lakes, where physical and biological conditions are typically very heterogeneous. Goff (1985) and Shuter et al. (1980) have documented the fragility of this species during the first few weeks of life, when they are virtually immobile. These two circumstances suggest that the habitat requirements for smallmouth bass reproduction are probably not satisfied throughout the littoral zone. Therefore, identifying littoral conditions where nest aggregations are common could provide insight into the mechanisms influencing the production of young. Survival rates of smallmouth bass early in life are an important influence on year-class abundance in adulthood

(MacLean et al. 1981). Consequently, predicting high-density spawning locations would be prudent for their protection from angling activities during the spawning/rearing period (which are known to severely compromise smallmouth bass reproductive success (Phillipp et al. 1997)), and from anthropogenic development in and around littoral areas (Christensen et al. 1996).

Smallmouth bass reproduction involves the creation of a nest (a depression in the sediment excavated by the male, containing offspring). High nest-site fidelity of returning adults (Ridgway et al. 1991a), suspected phylopatriy (Gross et al. 1994), and stationary locations of nest aggregations from year to year (Rejwan et al. 1997) all suggest that habitat conditions influence the nest distributions of lake-dwelling smallmouth bass. Despite this evidence, no specific relationships between nest patch locations and habitat variables have been quantified (Rejwan et al. 1997). Of the numerous combinations of habitat conditions that may influence nest density, water temperature during the nest-guarding pe-

Manuscript received 7 February 1997; revised 3 December 1997; accepted 2 February 1998; final version received 3 March 1998.

⁴ Present address: Department of Biological Sciences, University of Calgary, Calgary, Alberta T2N 1N4 Canada.

riod, shoreline complexity, littoral-zone rugosity, and prevailing wind/wave exposure of study sites were measured in this study as a consequence of their importance to other aspects of smallmouth bass reproduction. Temperature, prevailing wind/wave exposure, and shoreline reticulation all affect the early survival (Shuter et al. 1980, MacLean et al. 1981, Goff 1985) and growth (Shuter et al. 1980, Ridgway et al. 1991b) of the young. Temperature is also an important influence on the year-class abundance of northern smallmouth bass populations (Shuter et al. 1980, MacLean et al. 1981, Serns 1982). Littoral benthic rugosity was measured since nest locations of riverine smallmouth bass populations are located in rugged river-bottom terrain (Sabo and Orth 1994).

Standard multiple-regression procedures have been commonly used to study relations between an organism and various aspects of its environment. However, non-linear relationships (Philippi 1993) and heteroscedasticity (Dutilleul and Legendre 1993), which are typical of ecological data, limit the effectiveness of standard multiple regression. Furthermore, interactions among environmental variables that are also typical of ecological data (Philippi 1993) are difficult to detect using this method of analysis. Tree regression analysis is a relatively new (Breiman et al. 1984) and useful alternative statistical approach (Efron and Tibshirani 1991) since it does not require the existence of linear relationships among the variables or homoscedasticity in variances, and because interactions among habitat variables in relation to the organism are detected automatically in the analysis. Tree regression results present the relations among physical habitat conditions and nest density in an easily interpretable way by expressing nest-site selection as a hierarchy of decision and/or selection processes, based on physical habitat criteria. Since both tree and standard multiple regression are statistical analyses that quantify the significance and the relative importance of independent (habitat) variables on a dependent variable (nest density), the usefulness of tree regression analysis was compared with the better known standard regression analysis.

FIELD SAMPLING METHODS AND PRELIMINARY ANALYSES

Detection of smallmouth bass nests

Sampling was conducted in Lake Opeongo, Ontario (5780 ha., 45°42' N, 78°22' W), where smallmouth bass nest depths range between 0.4 and 2.5 m (most frequently from 0.5 to 1.5 m). By snorkelling along the 1-m depth contour, exact nest locations were recorded throughout the 155-km perimeter of Lake Opeongo in four years (1984, 1992, 1993, 1994), and throughout Jones Bay (a 6.3-km particularly high-density nesting

area in Lake Opeongo) in eleven years (1977–1979, 1984, 1988–1994). These surveys of the smallmouth bass nesting population are among the most comprehensive records known for the breeding locations of a fish population (Rejwan et al. 1997).

Patchiness in nest distributions has been identified throughout a large range of spatial scales in Lake Opeongo (100-m to 10-km quadrat sizes) (Rejwan et al. 1997). Thirty-six 1-km-long littoral sites throughout Lake Opeongo and thirty-one 100-m-long littoral sites throughout Jones Bay, Lake Opeongo, were selected for measuring habitat conditions in such a way that they encompassed the entire ranges in nest densities in approximately uniform proportions.

Measurement of physical habitat variables

The wind/wave exposure of each study site was measured as the maximum fetch length along the prevailing wind directions at three equally spaced locations along the length of each 100-m and 1-km study site. Analysis (ANOVA: $F = 35.03$, $F = 2.14$, $df = 7, 72$, $P \ll 0.0001$, Tukey's multiple comparisons) of 10 years of continuously recording anemometer readings at the south tip of Lake Opeongo indicated that northwesterly and southwesterly wind velocities (in kilometers per hour) were significantly greater than winds from other directions over the duration of the average nest-guarding period (29 May to 30 June, Rejwan 1996). The maximum southwesterly and northwesterly fetch lengths were measured at each of three locations at 25%, 50%, and 75% of the distance along each 1-km and 100-m study site and were averaged to characterize exposure at each site. Southwest fetch lengths were multiplied by 1.37 since northwesterly winds were an average of 1.37 times slower than southwest winds ($P < 0.05$).

Water temperatures were recorded in a single central location within each 1-km and 100-m site at 1 m (approximate nest depth) using one of three types of temperature-recording instruments: continuous digital recorders (Ryan Temp-Mentors and DataSonde Hydro-labs), continuous paper recorders (Ryan Thermistors), and Geneq max-min thermometers. Midpoint temperatures were recorded from the minimum and maximum temperatures from each 2-d time interval at each site over the duration of the nesting period (25 May to 30 June 1993 and 1 June to 27 June 1994), regardless of the type of temperature recorder deployed (Rejwan 1996). Highly significant correlations between 1994 and 1993 temperatures at both the 1-km ($r^2 = 0.59$, $P < 0.0005$, $n = 27$ sites) and the 100-m ($r^2 = 0.31$, $P < 0.0025$, $n = 30$) spatial scales indicated that there was consistency between the years in the among-site temperature differences. Consequently, 1994 temperatures were used to represent the site-specific temper-

atures, since the temperature records for the 1993 nesting season were unavailable for 10 of 67 sample sites.

Reticulation of the shoreline along each 1-km and 100-m study site was calculated using a measurement protocol that facilitates inter-scale comparisons. This method takes into account the fractal properties of shorelines, whereby reticulated shoreline lengths increase rapidly with increasing measurement resolution relative to lengths of straighter shorelines (Kent and Wong 1982, Sander 1987). At the small and large spatial scales in this study, the shoreline reticulation of each site was measured using one high-resolution (with a Run-Mate curvimeter) and one low-resolution (with a ruler) measurement. On a 1:1793 scale map (100-m sites) and a 1:10 000 scale map (1-km sites) the real-image high resolutions were 1.8 m and 10 m, respectively, and low resolutions were 20 m and 100 m, respectively. Measurements of shoreline length with the low-resolution method were subtracted from the high-resolution site lengths of either 100 m or 1 km to generate a measure of shoreline complexity that increases with increasing complexity (Rejwan 1996).

The fourth variable, littoral-floor rugosity, was quantified by measuring the straight-line end-to-end length of an 18-m-long fine-link (1.0 cm) chain after lowering it over the profile of the littoral floor along the 1-m depth contour (approximate nest depth). The 18-m chain length was equal to the average territory range of a male smallmouth bass during the nest-guarding period (Scott 1993). Rugosity was measured at (and averaged over) ten and three equally spaced sampling sites within each 1-km and 100-m-long site, respectively. To produce a measure that was directly related to the littoral ruggedness of the terrain, these lengths were subtracted from the total chain length of 18 m.

METHODS OF STATISTICAL ANALYSIS

Standard multiple-regression analysis

A best-combinations method of standard multiple-regression analysis was employed whereby all combinations of physical variables were tested, and the smallest combination of physical variables that yielded the highest adjusted r^2 value was included in the standard multiple-regression equation. Triangular-shaped relationships that could not be linearized between nest density and both shoreline complexity and wave exposure at the 1-km scale, were expected to limit the effectiveness of the standard multiple-regression analysis.

The mechanisms involved in tree regression analysis

Tree regression analysis involves a recursive partitioning of the study sites into two groups (of high and a low nest density) that are as similar as possible in

nest density within each. The exact allocation of sites into the groups is determined by repeatedly dividing the data set (of thirty-six 1-km-long sites, or thirty-one 100-m-long sites) into two groups at every possible measurement within the entire range of measurements of each of the four habitat variables. After each of the binary splits, the variability (within-groups sum of squares) in nest density is calculated for the two groups of samples that are created, and these two values are summed. Ultimately, the habitat variable and split-point that results in the smallest amount of variance in nest density within the two created groups, combined, is used to divide the data set in the regression tree. Exactly the same assessment is then repeated with each of the two groups that were created, and each of the independent variables is assessed regardless of whether or not they were previously used in the tree. The data set is divided through this recursive binary-partitioning mechanism until there is no more than a single site in a group, or until there is no variation in nest density among sites within the group (Clark and Pregibon 1992, StatSci 1993).

The regression tree that is created is structured in a hierarchical fashion with the initial undivided data set at the top (the root) followed by binary splits, each of which are called "nodes," to final undivided groups of sites ("leaves") at the bottom of the tree. The proportion of variance in nest density that is explained by each split is indicated by the vertical length of the "branches" that extend from each node to each subsequent node or leaf. The size of the regression tree is measured by the number of leaves (final groups) in the tree. Thus, the undivided data set is considered to have a tree size of 1.

Cross-validation analysis: determining tree size

Near the top of a regression tree, the early partitions of study sites are relatively likely to reflect the relationships that actually exist between nest densities and the habitat conditions. However, as tree branching continues, involving the partitioning of smaller and smaller sample sizes, the precision of each split diminishes. Consequently, branches near the bottom of the regression tree are less generalizable outside of the sample to the rest of the population that has not been sampled. To estimate what part of the regression tree quantifies the relations between nest density and habitat conditions that exist outside of the samples, across the entire Lake Opeongo (1-km scale) and Jones Bay (100-m scale) populations, a procedure known as "cross-validation" was employed (Clark and Pregibon 1992, StatSci 1993).

Cross-validation analysis involved the random partitioning of the data set into 10 groups of equal or similar size and the creation of a "cross-validation re-

gression tree" with only 9 of the 10 groups. This cross-validation regression tree was used to predict nest density for each of the sites in the remaining 10th group (the Cross-Validation Group). Predicted and observed nest densities were compared by calculating variance from the predicted mean value within each leaf at all possible tree sizes, from the initial division of the data into a two-leaved tree, onwards. This procedure was repeated 10 times (10 cross-validation trees were created), so that each of the 10 groups of sites was used as the Cross-Validation Group once. The entire procedure was repeated 200 times, each time with a new random assortment of sites into 10 groups to ultimately produce 2000 cross-validation assessments (Manly 1991).

The precision of the regression-tree predictions increases with tree size as long as the total amount of explained variation in nest density (total variance subtracted by unexplained variance) among the cross-validated sites increases with increasing tree size. However, as tree size continues to increase, the total amount of explained variance in nest density will eventually decrease. This component of the regression tree is considered to be too imprecise to be generalizable beyond the sample. Therefore, all tree sizes for which no further increase in explained variance occurs are removed ("pruned") from the tree (Clark and Pregibon 1992).

Cross-validation analysis: precision tests

Significance tests (which are used in tree and standard multiple-regression analyses, and are described in the next section) do not indicate the extent to which the samples (in this case, 1-km and 100-m sites selected from Lake Opeongo and Jones Bay, respectively) are representative of the whole population. Precision tests are an important method of providing this information. The loss of model precision in extrapolating to the whole population depends on the proportion of the population that was sampled (which is controllable), the amount of variation in the population (which is uncontrollable), and the extent to which models over-fit the data (the ratio of the sample size to the number of explanatory variables used in the analysis, and the intricacy of the model, both of which are controllable). The assumption that tree and standard multiple-regression models can be extrapolated to whole populations without much loss of accuracy was tested.

In this study, sample sites encompassed large proportions (25% [$n = 36$ sites] and 50% [$n = 31$ sites]) of the total nesting populations at the 1-km and 100-m spatial scales in Lake Opeongo and Jones Bay, respectively. Precision was calculated using the same cross-validation analyses that were used to determine the appropriate size of the regression tree models, described in the previous section. In estimating model

precision, the important information from the cross-validation results is the actual amount of the total variance in nest density that can be explained in the Cross-Validation Group, rather than relative changes in the explained variance across different tree sizes. This cross-validation procedure was repeated (replicated exactly) in a cross-validation analysis of the standard multiple-regression results.

Tests of statistical significance

After the regression tree has been created and pruned using the cross-validation procedure, it is necessary to determine whether the pruned tree explains significantly more variance than a random regression tree of equal complexity. To determine this, the amount of variance (expressed as the r^2 value) explained by the True Regression Tree (created using the complete data set) was compared with r^2 values of regression trees (pruned to the same size as the True Tree) generated from 2000 random associations (Manly 1991) between nest density and the habitat variables. If the 'true' r^2 value was among the top 5% of the randomly generated r^2 values, the True Regression Tree represents significant (rather than chance) relations between nest density and the habitat variables used in the tree model.

Using this permutation method to determine the significance of the True Regression Tree is in many cases a more accurate way to test for significance than traditional significance tests used in standard multiple-regression analysis. This is because the traditional standard multiple-regression test for significance considers only the combination of physical variables that were used in the final best-fit regression line (or at most, the total number of independent variables used in the research), rather than adjusting for the a posteriori evaluation of often hundreds of possible combinations of independent variables that are initially considered in the search for the best-fit regression line. Thus, the calculated P value is many times smaller than it should be, resulting in an overestimate of the significance of the standard regression fit. This is a very important and common problem in standard multiple-regression analyses. In contrast, the P value that is calculated from the repeated sampling technique in tree regression correctly takes into account the approximate number of combinations of physical variables that are possible, since all combinations of the four physical habitat variables are considered in creating both the True Regression Tree, and the 2000 randomly generated trees.

RESULTS AND DISCUSSION

Smallmouth bass nest distributions at the 1-km spatial scale are not well described unilaterally by any of the four physical variables that were measured (temperature $r^2 = 0.17$, wind/wave exposure $r^2 = 0.09$,

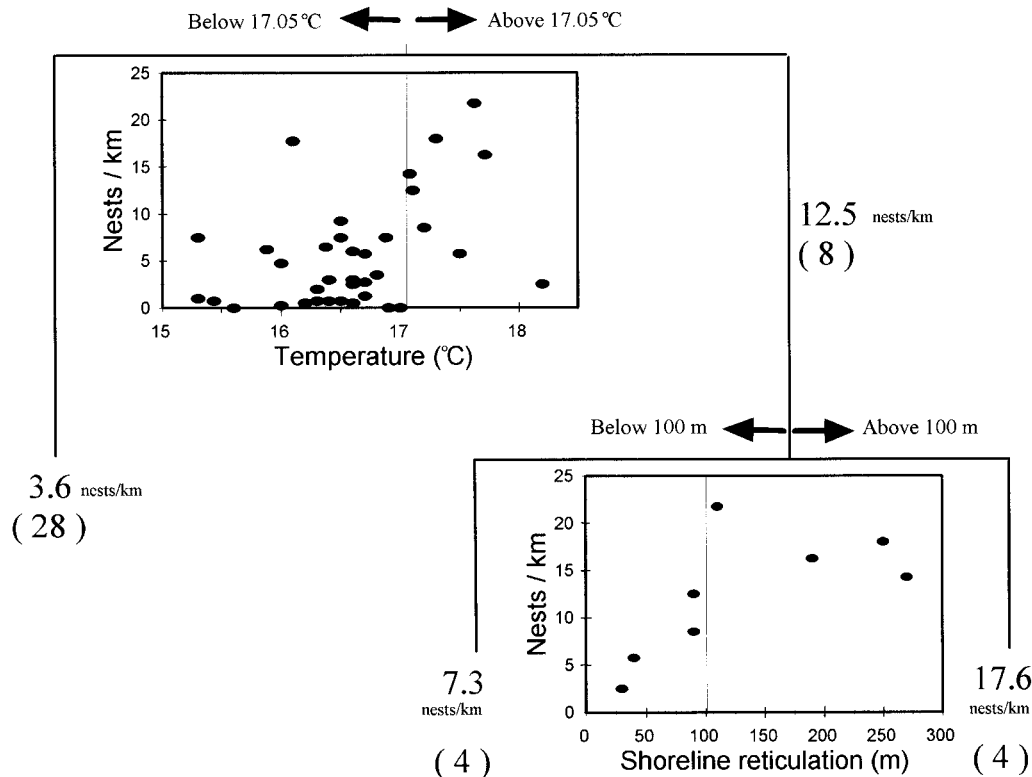


FIG. 1. This regression tree model ($r^2 = 0.58$, $P < 0.02$, $n = 36$ sites), generated from the 1-km-scale data, and pruned to a size of three terminal groups (leaves), distinguishes among high- and low-density nesting sites (solid black ovals). The mean number of nests per kilometer is listed for each group of sites at each tree branch and leaf, and immediately below are sample sizes in parentheses. The vertical line in each plot identifies the criterion used in the binary partitioning of sites into different sections of the tree according to their temperature ($^{\circ}\text{C}$) and degree of shoreline complexity (m). Shoreline complexities below 100 m indicate $<100\text{-m}$ difference in shoreline length between high (consistently 1 km) and low measurement-resolution methods.

shoreline reticulation $r^2 = 0.18$, benthic rugosity $r^2 = 0.00$). Standard multiple-regression results explained a significant amount of variance in nest density among the sites ($r^2 = 0.47$, adjusted $r^2 = 0.41$, $P < 0.0004$, $n = 36$ sites; nest density = $0.03r - 93.37t + 2.89t^2 + (8.1 \times 10^{-4})f + 751.36$, where r = shore reticulation, t = temperature, f = fetch exposure) despite the triangular (distinctly nonlinear) relations of fetch, and of littoral benthic rugosity, to nest density. Shoreline complexity, fetch, the temperature index, and a squared-transformation of the temperature index (to remove observed curvilinearity of residuals) were all significant components in this regression equation. However, on average, none of the variance in the cross-validated sites (median $r^2 = -0.07$, $n = 2000$ permutations) could be explained by standard regression equations created with the other 90% of the data. Thus, the standard multiple regression model would probably not be useful in explaining differences in nest density outside of the sample in the rest of the population.

The pruned regression tree ($r^2 = 0.58$, $P < 0.02$, $n = 36$ sites, Fig. 1) indicates a strong positive relationship between nest density and both temperature and shoreline complexity. The tree was pruned to a size of three leaves since subsequent sections of the tree model did not improve the predictions of nest densities among the cross-validated samples (Fig. 2). Cross-validation analysis also indicated that the tree regression model would likely maintain substantial precision in extrapolations outside the sample, to the population (median $r^2 = 0.20$, $n = 2000$ permutations, Fig. 2).

The regression tree results suggest that decision making by adult smallmouth bass is analogous to a hierarchical assessment of physical habitat variables, from primary considerations of temperature, to secondary considerations of shoreline complexity (Fig. 1). The large proportion of variance in nest density is explained by distinguishing the warmest 23% of the sites (above 17.05°C) from all other sites in the regression tree ($r^2 = 0.40$, $n = 8$ sites). This is probably largely generated

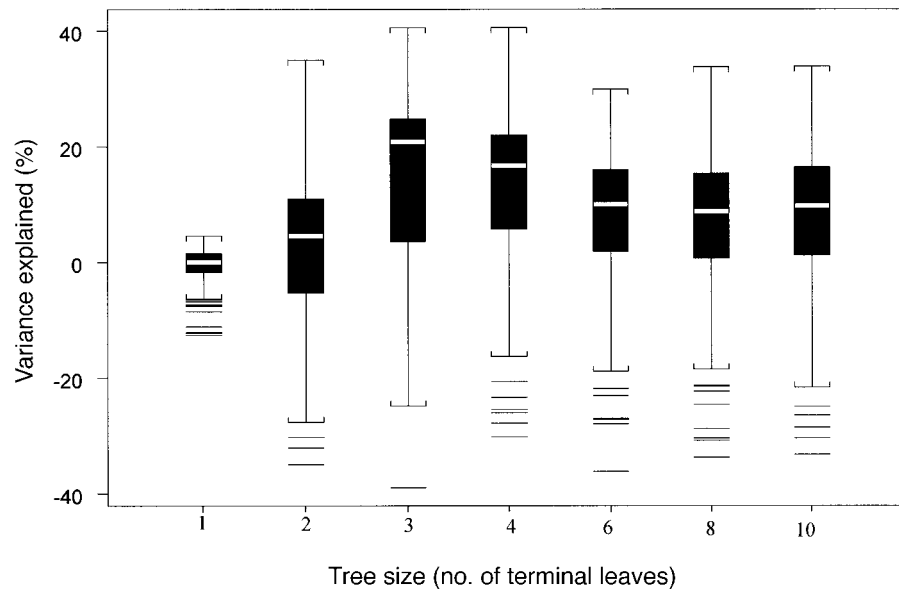


FIG. 2. The percentage of the total variation in nest density that is explained at each regression tree size (number of terminal nodes) was calculated from 2000 cross-validation repetitions. This plot is used to determine how much of the regression tree is imprecise and should be cut (pruned). Beyond a tree size of 3, no increase in explained variance in nest density is observed. The white line within each box plot indicates the median explained deviance, and the boxes encompass the interquartile distance (IQD). The bracketed vertical dotted lines encompass values that are within 1.5 IQD from the median, and all samples beyond this are represented by horizontal lines.

by one or both of two known effects of temperature on reproduction in smallmouth bass populations near the northern limits of their range. First, early growth and survival rates of broods are very closely tied to local thermal conditions (Shuter et al. 1980), which are in turn tied to the abundance of that cohort at adulthood (MacLean et al. 1981). At water temperatures below $\sim 15^{\circ}\text{C}$, survival rates of young in their nests decline dramatically (Shuter et al. 1980). Since phytopatry likely exists in the Lake Opeongo smallmouth bass population (Gross et al. 1994), high densities of nests would be less likely in areas where brood survival is consistently low. Second, in relatively warm environments (where the young develop more rapidly), guarding males can complete their nest-guarding duties sooner. Since adult food availability during the nest-guarding period likely limits the lifetime reproductive success of nest-guarding males (Ridgway and Shuter 1994), adults nesting in relatively warm environments are perhaps least likely to abandon their broods early, and/or are most likely to be able to return to nest in subsequent years.

If nest-patch locations in Lake Opeongo are restricted by minimum-temperature tolerances of smallmouth bass, as the above two mechanisms would suggest, nests in relatively warm areas may supply a larger proportion of their young to the cohort than do nests in colder regions. Since temperature patterns in Lake

Opeongo at the 1-km scale show significant consistency between years (Rejwan et al. 1997), temperature sensitivity may create the stable patterns in nest distributions that have been documented. If this association were common in lakes, then particularly warm regions should be protected from anthropogenic activities.

Despite significant consistency in nest-patch locations, nest-site fidelity is not perfect (Ridgway et al. 1991a); sparse distributions of nests that occur outside of nest patches (Rejwan et al. 1997) probably exist as a consequence of imperfect nest-site fidelity and environmental conditions that are adequate (at least in some years) for brood survival. It is also possible that a few nesters persist in suboptimal, low-density nesting environments as a consequence of some survival advantage over nesters from high-density areas at other stages in their life history.

An interaction was identified by the regression tree whereby nest densities are particularly high among the eight warmest sites that have relatively reticulated shorelines (but no such relationship existed among the cooler sites). This may be the result of various possible mechanisms. Shoreline complexity was included in this study with the expectation that complex shorelines would be preferred in typically high-exposure windward environments, where warm epilimnetic water accumulates. Consequently, complex shorelines were expected to provide protection from wind-driven waves,

known to be detrimental to brood survival (Goff 1985), while containing beneficial warm water temperatures. However, this mechanism does not appear to explain the positive relationship between shoreline complexity and high nest densities in Lake Opeongo. The three 1-km Jones Bay sites are among the four 1-km sites categorized in the regression tree by their particularly warm water conditions, complex shorelines, and high nest densities (Fig. 1), and Jones Bay is downwind from very large fetch lengths (up to 6 km) across the prevailing wind directions. Even though significant nest patchiness exists at a smaller (100-m) scale within Jones Bay (Rejwan et al. 1997), the nest patches are not confined to protected sections of the bay. Numerous alternative mechanisms may be responsible for the relation between complex shorelines and high nest density in warm areas, such as increased concealment of nest aggregations from brood predators.

At the 100-m scale (in Jones Bay) no significant relationship was detected between any of the four physical habitat variables and nest density from simple bivariate plots (r^2 range: 0.01–0.09, $n = 31$ sites for each of four plots), standard multiple regression ($r^2 = 0.07$, $P < 0.17$, $n = 31$ sites), or tree regression ($r^2 = 0.63$, $P < 0.36$, $n = 31$ sites) analyses. These results may be explained by one of two possibilities. First, statistical power may have been too low to detect relationships that indeed exist at the 100-m scale (relative to the 1-km scale) where less extreme (although significant) patchiness was observed (Rejwan et al. 1997). Alternatively, having satisfied the large-scale requirements of favorable thermal and shoreline reticulation conditions, small-scale nest patchiness in Jones Bay may be determined by other environmental factors such as temporally fine-scale differences in temperature (e.g., differences in short-term temperature fluctuations due to seiche activity, or in temperatures at the onset of the nesting period when among-site differences appear to be greatest).

The tree regression technique, based on the premise that bimodal relations among the variables exist, is better suited to the data in this study than the more commonly used standard multiple-regression analysis, which was extremely imprecise despite its statistical significance. Cross-validation results from both the standard and tree regression analyses demonstrate that caution should be taken in extrapolating sample-generated findings to processes underway within whole populations, even when very large proportions (23% of all 155 1-km-long sites) of the population have been sampled. Cross-validation analysis is an excellent method of evaluating the precision of new predictions, and should prove broadly useful for evaluating predictive models, rather than relying on significance tests alone.

ACKNOWLEDGMENTS

We are very grateful to Gene Wilde for his introduction to the subject of tree regression analysis. The work by Brenda Leach, Jenn Rowley, and numerous other field assistants from the Harkness Laboratory of Fisheries Research were very much appreciated. We are also grateful to Wayne Burchat from Ontario Hydro for the loans of Hydrolab temperature recorders. The advice and helpful suggestions of Philip Crowley and two anonymous reviewers were much appreciated.

This paper is a contribution of the Harkness Laboratory of Fisheries Research, Ontario Ministry of Natural Resources.

LITERATURE CITED

- Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and regression trees. Wadsworth, Belmont, California, USA.
- Christensen, D. L., B. R. Herwig, D. E. Schindler, and S. R. Carpenter. 1996. Impacts of lakeshore residential development on coarse woody debris in north temperate lakes. *Ecological Applications* 6:1143–1149.
- Clark, L. A., and D. Pregibon. 1992. Tree-based models. Pages 377–419 in J. M. Chambers and T. J. Hastie, editors. Statistical models in S. Wadsworth & Brooks, Pacific Grove, California, USA.
- Dutilleul, P., and P. Legendre. 1993. Spatial heterogeneity against heteroscedasticity: an ecological paradigm versus a statistical concept. *Oikos* 66:152–171.
- Efron, B., and R. Tibshirani. 1991. Statistical data analysis in the computer age. *Science* 253:390–395.
- Goff, G. P. 1985. Environmental influences on annual variation in nest success of smallmouth bass, *Micropterus dolomieu*, in Long Point Bay, Lake Erie. *Environmental Biology of Fishes* 14:303–307.
- Gross, M. L., A. R. Kapuscinski, and A. J. Faras. 1994. Nest-specific DNA fingerprints of smallmouth bass in Lake Opeongo, Ontario. *Transactions of the American Fisheries Society* 123:449–459.
- Kent, C., and J. Wong. 1982. An index of littoral zone complexity and its measurement. *Canadian Journal of Fisheries and Aquatic Sciences* 39:847–853.
- MacLean, J. A., B. J. Shuter, H. A. Regier, and J. C. MacLeod. 1981. Temperature and year class strength of smallmouth bass. *Proceedings of the symposium on early life history of fish. International Council for the Exploration of the Sea* 178:30–40.
- Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman & Hall, London, UK.
- Philipp, D. P., C. A. Toline, M. Kubacki, D. B. F. Philipp, and F. Phelan. 1997. The impact of pre-season catch and release angling on the reproductive success of smallmouth and largemouth bass. *North American Journal of Fish Management*, 17:557–567.
- Philippi, T. E. 1993. Multiple regression: herbivory. Pages 84–94 in S. M. Scheiner and J. Gurevitch, editors. The design and analysis of ecological experiments. Chapman & Hall, New York, New York, USA.
- Rejwan, C. 1996. The relations between smallmouth bass (*Micropterus dolomieu*) nest distributions and characteristics of their habitat in Lake Opeongo, Ontario. Thesis. University of Toronto, Mississauga, Ontario, Canada.
- Rejwan, C., B. J. Shuter, M. S. Ridgway, and N. C. Collins. 1997. Spatial and temporal distributions of smallmouth bass (*Micropterus dolomieu*) nests in Lake Opeongo, Ontario. *Canadian Journal of Fisheries and Aquatic Sciences* 54:2007–2013.
- Ridgway, M. S., J. A. MacLean, and J. C. MacLeod. 1991a. Nest-site fidelity in a centrarchid fish, the smallmouth bass

- (*Micropterus dolomieu*). Canadian Journal of Zoology **69**: 3103–3105.
- Ridgway, M. S., and B. J. Shuter. 1994. The effects of supplemental food on reproduction in parental male smallmouth bass. Environmental Biology of Fishes **39**:201–207.
- Ridgway, M. S., B. J. Shuter, and E. E. Post. 1991b. The relative influence of body size and territorial behaviour on nesting asynchrony in male smallmouth bass, *Micropterus dolomieu* (Pisces: Centrarchidae). Journal of Animal Ecology **60**:665–681.
- Sabo, M. J., and D. J. Orth. 1994. Temporal variation in microhabitat use by age-0 smallmouth bass in the North Anna River, Virginia. Transactions of the American Fisheries Society **123**:733–746.
- Sander, L. M. 1987. Fractal growth. Scientific American **256**: 94–100.
- Scott, R. J. 1993. The influence of parental care behaviour on space use by adult smallmouth bass, *Micropterus dolomieu*. Thesis. University of Guelph, Guelph, Ontario, Canada.
- Serns, S. L. 1982. Relation of temperature and population density to first-year recruitment and growth of smallmouth bass in a Wisconsin lake. Transactions of the American Fisheries Society **111**:570–574.
- Shuter, B. J., J. A. MacLean, F. E. J. Fry, and H. A. Regier. 1980. Stochastic simulation of temperature effects on first-year survival of smallmouth bass. Transactions of the American Fisheries Society. **109**:1–34.
- StatSci. 1993. Splus for Windows, version 3.1. Reference manual. Volume 1. Statistical Sciences, Inc., Seattle, Washington, USA.