Transactions of the American Fisheries Society 131:1033–1043, 2002

Relationships between Environmental Characteristics and the Density of Age-0 Eurasian Perch *Perca fluviatilis* in the Littoral Zone of a Lake: A Nonlinear Approach

SEBASTIEN BROSSE* AND SOVAN LEK

UMR 5576 Center for the Ecology of Aquatic Ecosystems, Université Paul Sabatier,
118 Route de Narbonne, 31062 Toulouse Cedex, France

Abstract.—We studied the spatial distribution of age-0 Eurasian perch *Perca fluviatilis* in the littoral area of a large lake (Lake Pareloup, France) using eight environmental variables as habitat descriptors. Nonparametric locally weighted scatterplot smother (Lowess) functions were used to visualize the relationships between spatial distribution and the habitat descriptors. The highest abundance was observed in the transition area between shallow water with dense vegetation cover and unvegetated open water. Habitat use depended on a combination of environmental variables, such as depth, distance from the bank, vegetation cover, and slope of the bank, with abundance exhibiting nonlinear responses to each variable. We hypothesized that these complex responses resulted from a trade-off between searching for food and avoiding predators. We then attempted to build a predictive model of age-0 perch abundance based on the environmental descriptors using an artificial neural network (ANN). The predictive quality of the model was high ($r^2 = 0.78$ between the observed and estimated perch densities) compared with that of the more classical linear modeling technique (i.e., multiple linear regression; $r^2 = 0.20$) and another nonlinear modeling technique (a generalized additive model; $r^2 = 0.33$). Finally, ANN sensitivity analyses of the environmental variables in the models confirmed the results obtained with the Lowess approach, which considered the influence of each variable on perch habitat use. In light of these results, ANN and Lowess methods have considerable potential in the prediction and explanation of ecological relationships.

There has been increasing interest in the study of habitat as it relates to the spatial distribution of fish populations (Beecher et al. 1993; Rossier 1995; Fischer and Eckmann 1997). Within lake ecosystems, the littoral zone is important to fish, with abundance there often being greater than in other areas (Fischer and Eckmann 1997; Brosse et al. 1999a). Age-0 fish use the littoral zone during summer because it provides food and shelter (Savino and Stein 1989; Persson and Eklöv 1995). However, most European studies of age-0 fish habitat use deal with cyprinids (e.g., Grossman et al. 1987; Copp 1992; Mastrorillo et al. 1996). The habitat use of Eurasian perch *Perca fluviatilis* has rarely been investigated even though this species is one of the most common in European lowland rivers and lakes (Persson and Eklöv 1995; Machacek and Matena 1997). Moreover, most studies of fish habitat use in lakes have been conducted at a large spatial scale (Rossier 1995; Fischer and Eckmann 1997; Brosse et al. 1999a), whereas small-scale habitat studies have been limited to streams and rivers (Copp 1992; Beecher et al. 1993; Baran et al. 1996; Mastrorillo et al. 1996).

Our lack of knowledge in this area is probably due to the complexity of habitat responses, which are governed by complex, often nonlinear interactions among various environmental variables (Brosse et al. 1999b). To deal with such complexity, the transformation of nonlinear variables by logarithmic, power, or exponential functions is often used, but the transformed variables sometimes fail to fit experimental data (Lek et al. 1996). On the other hand, to minimize the complexity of the natural system, environmental variables are usually grouped into classes or fish abundances are transformed into presence/absence data (e.g., Copp 1992; Beecher et al. 1993). However, these transformations are not always representative of biological reality and can cause biases in the analyses. Several alternatives to using linear methods and presence/absence data are available, such as nonparametric smoothing methods (i.e., the locally weighted scatterplot smoother [Lowess] method; Cleveland 1979). Such smoothing methods can accurately fit nonlinear data (Hastie and Tibshirani 1990; Trexler and Travis 1993; Brosse et al. 1999c) and thus are able to portray nonlinear relationships between two variables. From a predictive point of view, artificial neural networks (ANNs) perform the same task as regression analysis and are par-

---

* Corresponding author: brosse@cict.fr

Received April 3, 2001; accepted March 18, 2002

ticularly suited to nonlinear data (Rumelhart et al. 1986; Lek and Guégan 1999). We have previously used these modeling tools to predict fish species richness (Guégan et al. 1998) and the density and biomass of fish populations in various streams (Baran et al. 1996; Mastrorillo et al. 1997) and lakes (Brosse et al. 1999b, 1999c).

In this paper we first describe the relationships between age-0 Eurasian perch density and habitat using the Lowess method (Cleveland 1979). Then, a predictive model of density is formulated using an ANN. The predictive performance of the ANN is evaluated by comparing it with that of classical modeling techniques (i.e., multiple linear regression [MLR] and generalized additive models [GAMs]), and the reliability of the model's representation of ecological reality is compared with that of Lowess curves.

Methods

Study site and sampling.—The study was undertaken during summer 1997 in Lake Pareloup. Located in southwestern France, this reservoir has an area of 1,250 ha and a volume of approximately 168×10^6 m^3. The maximum depth is 37 m and the average depth 12.5 m. Lake Pareloup is a warm, monomictic lake that stratifies thermally at 10 m from early June to mid-September and develops low oxygen content in the hypolimnion. This prevents fish from colonizing deep water in summer. The reservoir is subject to large water level fluctuations (3–10 m) between summer and winter due to pumping by a power station. Therefore, the aquatic vegetation is limited to terrestrial vegetation located in the littoral areas that are flooded during the summer high-water period (April–October).

Sampling of age-0 Eurasian perch was done once per week in the littoral zone of the lake from the end of the fish's pelagic period (late June) to the end of summer (August), when the fish migrated to deeper areas. Samples were collected between 0900 and 1800 hours in order to avoid biases due to different fish occupation patterns during the day and at night (Hasler and Villemonte 1953; Imbrock et al. 1996). Point sampling abundance by electrofishing (Nelva et al. 1979) modified for young fish was employed to evaluate age-0 perch habitat use. Electrofishing was performed using a backpack electrofishing unit with a 10-cm ring anode. Such equipment can be used in a large range of situations and is efficient for the entire range of age-0 fish sizes (Copp 1989). For each sampling point, the anode was swiftly immersed into the water (generally about 50 cm but less at shallower points), and stunned fish were collected with a fine-mesh (1-mm) dip net. Each week, 30–40 sampling points (separated 5–10 m from each other to avoid biases due to fish escapement from one sample to the next) were haphazardly selected and investigated in the same area of the lake.

For each of the 306 resulting sampling points, nine habitat variables were measured: distance from the bank (m); depth (m); the local slope of the bottom (measured according to a scale ranging from 0 [none] to 3 [sheer]); the percentage of flooded-vegetation cover, which was visually estimated as the percentage of bottom area covered; and the percentages of bottom area composed of boulders, pebbles, gravel, sand, and mud, which were determined by means of the Cailleux (1954) methodology. Variables were measured in a 1-m^2 bottom area corresponding to each sample. The Pearson correlation matrix, which was derived from simple linear regressions, was used to test for collinearity and showed a strong negative correlation between the percentages of sand and mud ($r^2 = -0.97$, $P < 0.01$). As a result, the sand variable was removed from the data matrix. The correlations among the remaining eight variables were not significant ($P > 0.01$) for 21 of the 28 combinations. For the other 7 combinations, the coefficients of determination (r^2) were all less than 0.30 (Table 1), indicating low collinearity.

In addition to gathering data on the environmental variables, we collected age-0 Eurasian perch (when present) in each sample. These were preserved in a 4% solution of formaldehyde and later identified and counted in the laboratory. Prior to statistical analyses, density data were $\log_{10}(x + 1)$ transformed. Because age-0 fish generally aggregate in dense shoals, which leads to a large range of densities in similar habitats, statistical methods are typically unable to reliably fit these data. Previous researchers have attempted to solve this problem in two ways. The first was to aggregate the samples. However, the larger spatial scale inherent in that approach tends to obscure the association between environmental patterns and fish density. The second was to transform the data into presence/absence form, an approach that leads to the loss of a large amount of information. By contrast, logarithmic transformation reduces data heterogeneity, avoiding the undue influence of outliers (ter Braak and Looman 1995) while incorporating truly quantitative information. Moreover, logarithmic transformation of the dependent variable has been commonly and efficiently applied

TABLE 1.—Bivariate correlation matrix (r^2 values) between the original nine environmental variables. Asterisks indicate significant correlations ($P < 0.01$). Values associated with the variable San are given in bold (this variable was not used in analyses; see text for details). Abbreviations are as follows: Dep = depth, Dis = distance from the bank, Slo = slope, Bou = percent boulders, Peb = percent pebbles, Gra = percent gravel, San = percent sand, Mud = percent mud, Veg = percent flooded vegetation.

Variable	Variable							
	Dep	Dis	Slo	Bou	Peb	Gra	San	Mud
Dis	0.02							
Slo	0.02	0.18*						
Bou	0.00	0.00	0.00					
Peb	0.00	0.00	0.01	0.00				
Gra	0.00	0.00	0.01	0.00	0.09*			
San	**0.00**	**−0.12***	**−0.18***	**0.00**	**−0.01**	**−0.01**		
Mud	0.00	0.14*	0.17*	0.00	0.00	0.00	**−0.97***	
Veg	−0.02	−0.28*	−0.25*	0.00	−0.01	0.00	**0.13***	−0.12*

prior to Lowess smoothing and ANN modeling (Brosse et al. 1999b, 1999c; Lek and Guégan 2000).

Relationships between age-0 perch density and habitat.—Habitat use by age-0 Eurasian perch was estimated by representing the fish abundance in each sample as a function of the environmental variables. Then, to vizualize the relationship of abundance to each variable, the data were fitted using the locally weighted scatterplot smoother usually called Lowess (Cleveland 1979). In this procedure, each sample is smoothed using a defined proportion of the neighbors nearest to the target point. Optimal fitting is obtained iteratively by minimizing the residuals between the observed and estimated values. The proportion of the samples perfectly fitted by Lowess is indicated by the f value of the fitted distribution. The parameter f varies between 0 and 1 according to the sensitivity of the analysis and is determined empirically by testing various possibilities and selecting the one that provides the best ability to visualize general data tendencies (Trexler and Travis 1993). Two of the major advantages of this method are that it can accurately fit both linear and nonlinear data and it automatically shows the degree of dependence of the response to the predictor. However, no equation is associated with the Lowess curve due to its nonparametric nature, so only graphical results are obtained. These results could be validated by a nonparametric statistical test (e.g., Mann–Whitney or Kruskal–Wallis) after sorting the environmental values into different classes, but we did not do this because sorting data into classes obscures its continuous nature. Moreover, Lowess fitting is commonly used in ecology and provides reliable descriptive information without further analysis (Trexler and Travis 1993; James et al. 1996; Brosse et al. 1999c).

Prediction of age-0 perch density.—The whole data matrix (i.e., 306 records × 8 environmental variables) was divided into two submatrices. First, the 62 records with nonzero values for perch density were isolated from the 244 with zero values (i.e., no perch). Null values are often removed from analyses because they make the data noisy, greatly affect the statistical analysis, and can induce bias in the predictions of abundance and spatial distribution (Pennington 1996). Nevertheless, null values account for an ecological reality and, according to ter Braak and Looman (1995), should not be entirely discarded. As we could not include all of the null values (if most samples are null, the model will consider nonnull values as outliers), we randomly selected 25% of the records without fish (i.e., 61 records) and added them to the first submatrix, leading to a final matrix containing 123 records. The remaining 183 null records were treated as a validation data set that was used to test model predictions of the absence of perch. These matrices were used to calculate the GAM and the MLR and ANN models.

The predictive model of age-0 Eurasian perch habitat use was developed using a multilayer, feed-forward ANN (Rumelhart et al. 1986). The processing elements in the network, called neurons, are arranged in layers. The first layer, called the input layer, connects with the input variables. In our case, it comprised eight neurons corresponding to the eight environmental variables (Figure 1). The last layer, called the output layer, connects to the output variable(s). In our case, there was a single neuron corresponding to the value of the dependent variable (perch density). The layer be-

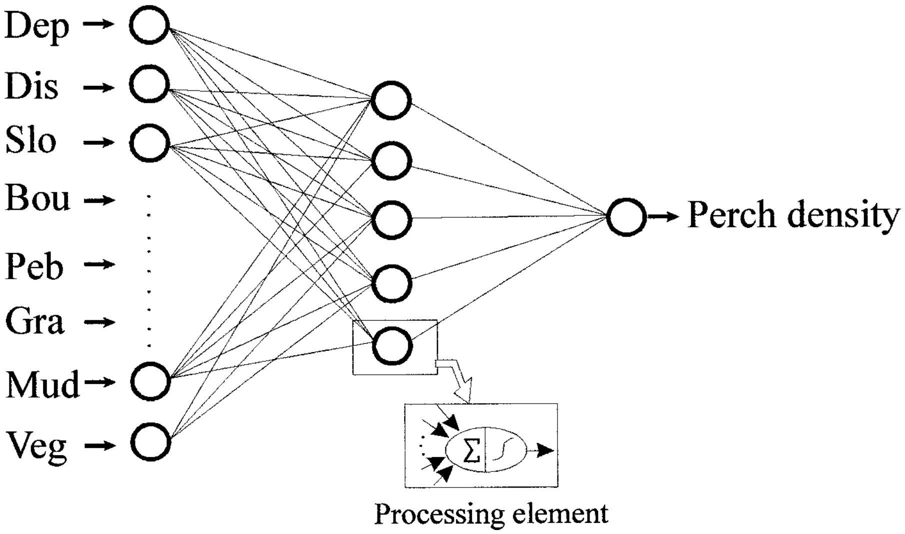

FIGURE 1.—Schematic of the three-layer, feed-forward artificial neural network used to estimate the density of age-0 Eurasian perch. The network consists of (1) eight input neurons corresponding to the eight independent environmental variables (Dep = depth, Dis = distance from the bank, Slo = slope, Bou = percent boulders, Peb = percent pebbles, Gra = percent gravel, Mud = percent mud, and Veg = percent flooded vegetation), (2) five hidden-layer neurons, and (3) one output neuron representing the dependent variable. The connections between neurons are shown by solid gray lines.

tween the input and output layers is called the hidden layer and was composed of five neurons (networks with more hidden neurons or two hidden layers did not do significantly better). The network configuration is determined empirically by testing various possibilities and selecting the one that provides the best compromise between bias and variance, that is, the best prediction (see Geman et al. 1992 and Kohavi 1995 for more details). Each neuron is connected to all of the neurons of adjacent layers and receives and sends signals through these connections. Signals are transmitted in only one direction: from the input layer to the output layer through the hidden layer. The connections are given weights that modulate the intensity of the signals they transmit. To extract biological information from the model, we used Garson's algorithm (1991) as modified by Goh (1995) to determine the relative importance of the environmental variables to perch habitat use based on the weights between the input and hidden and hidden and output layers that were calculated by the network.

As a check, the ability of the ANN to predict age-0 Eurasian perch density was compared both with that of classical multiple linear regression and with that of a more advanced nonlinear regression, the generalized additive model (Hastie and Tibshirani 1990). The GAM, a generalization of the MLR and generalized linear models, is a nonparametric regression method that models the dependent variable as the additive sum of unspecified functions of covariates. The least-squares and maximum likelihood methods used in the MLR and generalized linear models are replaced by quasilikelihood methods that rely on a nonparametric scatterplot smoother. The three modeling procedures (i.e., ANN, MLR, and GAM) were applied to the same data matrices after $\log_{10}(x + 1)$ transformation of age-0 perch density data (123 records × 8 environmental variables to derive the models and 183 null records × 8 variables to test their predictions as to the absence of perch). The predictive ability of the models was estimated using the leave-one-out cross-validation test applied to the first data set (123 records), where each sample is left out of the model formulation in turn and predicted once to determine model performance. This procedure is appropriate when the data set is quite small or each sample is likely to have ''unique information'' (Kohavi 1995). This method constitutes a reliable alternative to the hold-out procedure commonly used for larger data sets (Efron 1983; Kohavi 1995). Model reliability was assessed in terms of two criteria: the coefficient of determination (r^2) between the observed and estimated values, which provides information on the significance of the model; and SD_{PE}/SD_T, the ratio of the standard deviation of the prediction error to that of the training data. One minus this ratio describes the variance explained by the model, and a value below 1 indicates good regression performance (Statistica 2000).

Results

Environmental Variables and Age-0 Perch Density

The scatterplots of the individual environmental variables versus age-0 Eurasian perch density (Figure 2) illustrate the heterogeneity of the data set. A large range of responses was obtained for most of the environmental variables, and any a priori interpretation of the plots was not possible. Data fitting by parametric functions did not reveal any significant ($P > 0.05$) tendencies, leading us to use the nonparametric Lowess smoothing method. With that method, the influence of each of the eight environmental variables on age-0 perch abundance was clearly visualized. Four of these variables (depth, distance from the bank, slope of the bottom, and percentage of flooded vegetation) produced large variations in the Lowess curve and thus influenced perch habitat use (Figure 2). The four variables pertaining to the bottom substratum were less important, as the Lowess curve varied little. For the four most important variables, the Lowess fitting was a bell-shaped curve, indicating that age-0 perch density was maximal for medium values of the variables and lower for extreme values. As a result, density peaked in medium depths (0.25–0.45 m; Figure 2a) and distances from the bank (5–15 m; Figure 2b). In the same way, areas with gentle slopes were more densely inhabited than those with either sheer and no slopes (Figure 2c). Moreover, age-0 perch abundance was higher in areas with medium flooded-vegetation cover (40–60%) than in areas with low or high vegetation density (Figure 2d). With respect to the less influential variables, perch seemed indifferent to the percentage of fine substratum (mud; Figure 2e), but we noticed sharply lower abundance in larger substrata: perch were never found in areas with gravel and very seldom in areas with pebbles and boulders. As a consequence, age-0 perch were mainly located in intermediate areas that could be identified as an edge between flooded vegetation and open water.

Age-0 Perch Density Models

The MLR analysis produced the following equation ($r^2 = 0.20$):

$$\begin{aligned}\log_{10}(\text{perch density} + 1) = {} & \\ & 0.729 - 0.909 \cdot \text{Dep} + 0.008 \cdot \text{Dis} \\ & + 0.266 \cdot \text{Slo} - 0.008 \cdot \text{Bou} - 0.030 \cdot \text{Peb} \\ & - 0.013 \cdot \text{Gra} - 0.002 \cdot \text{Mud} - 0.001 \cdot \text{Veg},\end{aligned}$$

where Dep = depth, Dis = distance from the bank, Slo = slope, Bou = percent boulders, Peb = percent pebbles, Gra = percent gravel, Mud = percent mud, and Veg = percent flooded-vegetation cover.

Although the overall model was significant ($P < 0.01$), it did not perform better than a simple mean estimator, as the ratio of SD_{PE} to SD_T was 1.015, indicating that no variance was explained by the model. The predicted values showed that the medium and high values of fish abundance were always underestimated and most of the null values were either overestimated or given aberrant values (i.e., negative fish densities; Figure 3a). The points were not well distributed along the line of perfect prediction (the 45° line).

The GAM produced a slight improvement over the MLR model ($r^2 = 0.33$, $P < 0.01$) but also systematically underestimated high values of fish density and either overestimated or gave aberrant values for a large proportion of the null values (Figure 3b). Thus, even if the performance of the GAM was slightly better than that of the MLR model, with a $SD_{PE}-SD_T$ ratio of 0.83 it was still not acceptable, as the variance explained by the model was only 17%.

Predictions by the ANN models using the leave-one-out procedure (the ANN calibration process, in which correlation coefficients between the observed and predicted values of the training and testing data sets are plotted against the number of hidden neurons and the number of iterations, is illustrated by Lek et al. [1996]) yielded an r^2 of 0.78 ($P < 0.01$). Although some samples were under- or overestimated, a large proportion were well distributed along the 45° line (Figure 3c). Moreover, the $SD_{PE}-SD_T$ ratio was 0.49, a good regression performance with more than 50% of the variance being explained by the model.

The validation tests on the data matrix with 183 null samples produced different results with the MLR, GAM, and ANN approaches. A paired comparison of the predicted values with the Mann−Whitney nonparametric test showed significant ($P < 0.01$) differences between the MLR model and the GAM ($Z = -7.88$), the MLR and ANN models ($Z = -9.33$), and the GAM and the ANN model ($Z = -6.04$). Although the GAM per-

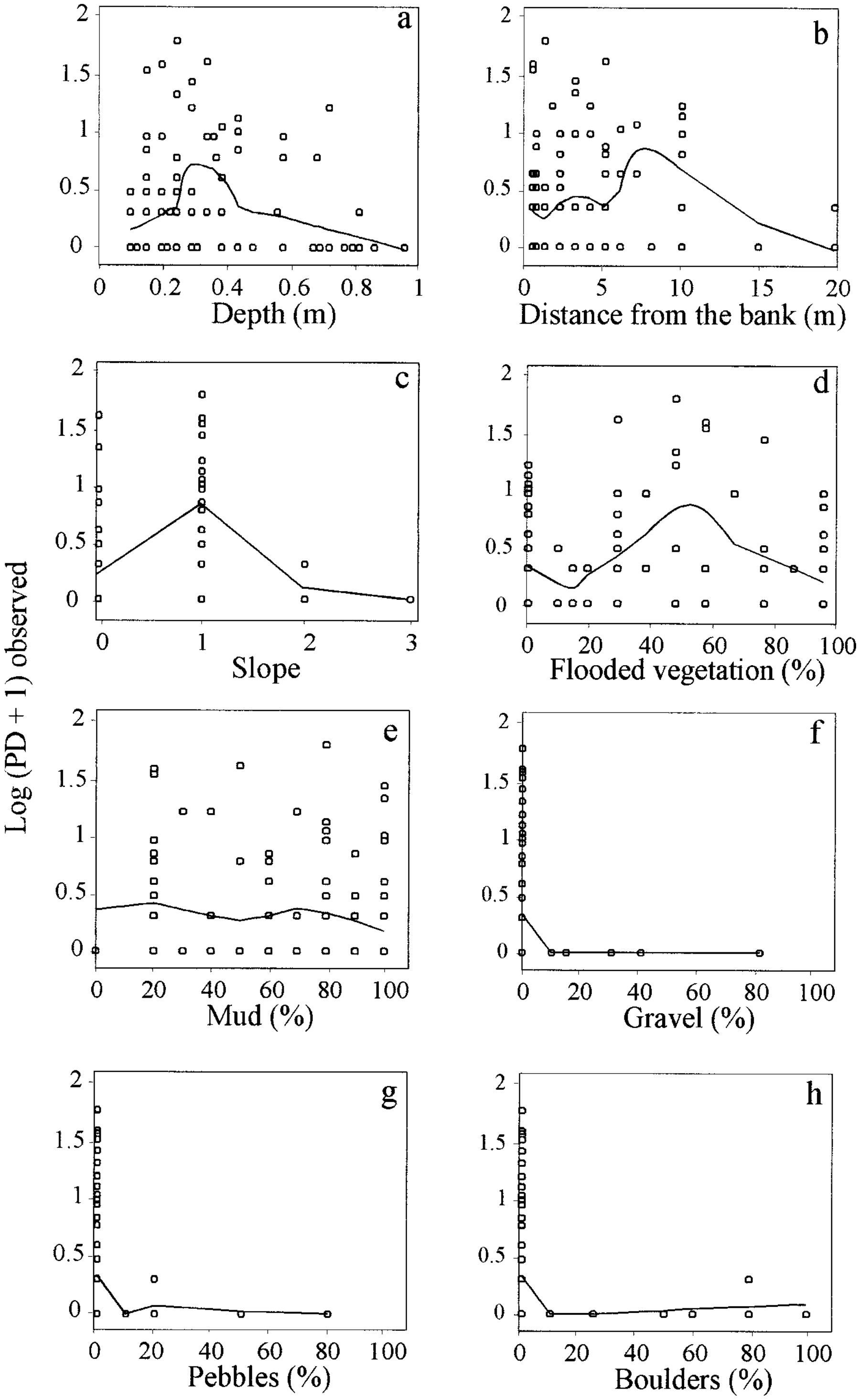

FIGURE 2.—Bivariate plots of the eight environmental variables against $\log_{10}$ transformed age-0 perch density (PD). Lowess curves (solid lines) were used to fit the data. The variables and their *f* values, which indicate the proportion of each sample that was perfectly fitted by the Lowess procedure, are as follows: **(a)** depth, 0.45; **(b)** distance from the bank, 0.35; **(c)** slope, 0.35; **(d)** percent flooded vegetation, 0.50; **(e)** percent mud, 0.50; **(f)** percent

formed significantly better than the MLR model (predicted values closer to null), the two regression methods systematically predicted medium perch densities. With the ANN model, most predicted values were closer to null than those predicted by the GAM or the MLR model (Figure 4). However, the ANN model did predict a high perch density for some samples (i.e., the outlier and extreme values in Figure 4).

In terms of the ecological information extracted from the ANN model, the results of the Garson−Goh algorithm stress the importance of four environmental variables (Figure 5), with contributions of roughly 22% for distance from the bank, 19% for depth and percent flooded vegetation, and 14% for slope. The four remaining variables each contribute less than 10% and therefore have limited influence on perch habitat selection. These results are in accord with those of the Lowess analysis, which also stressed the importance of the first four variables.

Discussion

Age-0 Perch Habitat Use

The information provided by the Lowess curves showed that age-0 Eurasian perch mainly occupied shallow, gently sloping, littoral areas with moderate densities of macrophytes. This is quite common for age-0 fish, for which vegetation and shallow water provide safety against predation by piscivorous fish, such as adult perch and northern pike *Esox lucius*, which are abundant in the open areas without vegetation in Lake Pareloup (Brosse 1999; Laffaille et al. 2001). Colonization of refuge areas by age-0 perch has already been demonstrated in controlled environments (Persson and Eklöv 1995; Eklöv and Persson 1996; Jacobsen and Berg 1998). However, these studies only considered vegetation density, without giving any attention to the other environmental characteristics that are likely to influence habitat choice. We obtained additional information on habitat use by age-0 perch by means of the accurate, precise representation of the environment afforded by quantitative measurements of a variety of environmental variables. Within littoral areas, the shallowest, most gently sloping areas closest to the banks were avoided. There are two possible reasons for this. First, age-0 perch may not have colonized these areas to avoid predation by terrestrial predators such as piscivorous birds (Winfield 1990). Although piscivorous birds are not particularly abundant at Lake Pareloup, the avoidance of shallower waters may be attributed to an acquired behavior called ‘‘the ghost of past predation’’ (Gliwicz and Warsaw 1992). Second, this behavior may have been related to the avoidance of competition with age-0 rudd *Scardinius erythrophthalmus*, a phenomenon that Eklöv and Hamrin (1989) have demonstrated in controlled environments. The fact that in Lake Pareloup rudd inhabit shallow littoral areas in great numbers during summer (Brosse 1999) supports this second hypothesis.

Age-0 Eurasian perch were primarily found in areas with between 30% and 70% vegetative cover, a finding that agrees with the results obtained in controlled and artificial environments (Christensen and Persson 1993; Diehl 1993; Persson and Eklöv 1995; Jacobsen and Berg 1998; Jacobsen and Perrow 1998). These studies have shown that age-0 perch are generally found in flooded vegetation, which affords protection from predators and provides a rich foraging habitat (Diehl 1993; Persson and Eklöv 1995). However, areas with more than 70% vegetative cover were avoided, as such vegetation density could reduce feeding ability (Diehl 1993). In the same way, the age-0 perch’s colonization of edge areas between vegetated and open waters and indifference toward substrate type supports the hypothesis of opportunistic feeding behavior by a fish that can use various food items (Hammer 1985; Dubois et al. 1994). However, age-0 perch were rarely found in gravel, pebbles, or boulders because these substrata are usually far from vegetation and are dangerous because of the higher abundance (Brosse 1999; Laffaille et al. 2001) and enhanced efficiency of predators such as northern pike and adult perch (Brabrand and Faafeng 1994; Eklöv 1997).

Most of the ecological knowledge of perch habitat use has been derived from experiments conducted in artificial and controlled environments under predation and competition pressures that were very different from those in nature (Eklöv and Hamrin 1989; Diehl 1993; Eklöv and Persson 1996; Jacobsen and Berg 1998). As a consequence,

---

←

gravel, 0.50; **(g)** percent pebbles, 0.50; and **(h)** percent boulders, 0.50. The value of *f* ranges from 0 and 1 according to the sensitivity of the analysis and is determined empirically by testing various possibilities and selecting the one that provides the best visualization of the general data tendencies.

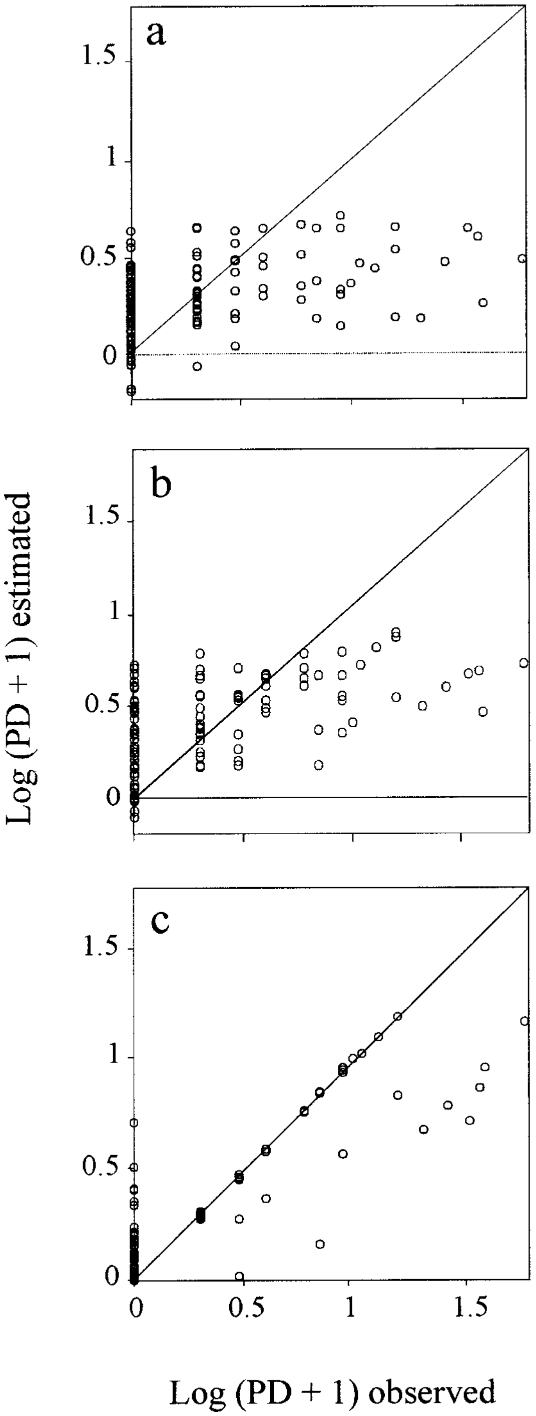

FIGURE 3.—Model predictions of age-0 perch densities (PDs) versus observed densities. Panel **(a)** shows the results for the multiple linear regression (MLR) model, panel **(b)** the results for the generalized additive model (GAM), and panel **(c)** the results for the artificial neural network (ANN) model. The 45° line indicates a perfect fit.

the results of studies that used cages or experimental ponds need to be interpreted with caution (Englund 1997) and should be verified in natural environments (Townsend et al. 2000). Although habitat use was more complex in our study than in these experiments, the perch tended to colonize areas with considerable structural heterogeneity. Such habitat preferences suggest a trade-off between the search for food and the requirement for shelter from predation. This trade-off between costs and benefits is well known in ecology (Persson and Eklöv 1995), but it differs among organisms and environments. As a consequence, our results—which were obtained in a natural, undisturbed environment—provide a quantitative description of habitat use by age-0 perch as well as support for the experimental and theoretical results in the literature that emphasize the importance of both structural characteristics and biotic interactions to habitat use by age-0 fish.

Prediction Capability of the Models

The MLR results showed that the main processes determining the abundance and habitat use of age-0 Eurasian perch can be approximated by linear functions to only a limited extent. Even with complex nonlinear transformation of the variables (i.e., the GAM), regression models were not able to faithfully reproduce the behavior of real systems when very low or very high values of the variables were considered. However, the improvement in the predictions produced by going from the MLR model to the GAM testifies to the nonlinearity of the variables and justifies the use of ANNs, which enable modeling of nonlinear systems (Goh 1995; Lek et al. 1996; Brosse et al. 1999b). In the present study, the network was able to reliably predict age-0 perch abundance on the basis of the eight environmental variables taken into account. Nevertheless, in some samples, medium fish densities were predicted where no perch were actually found (i.e., some areas that should have been suitable for age-0 perch were not being occupied). To explain this, we can hypothesize that these data points reflect a bias in the sampling methodology owing either to fish escapement or to underestimation by the electrofishing technique (Bain et al. 1985; Dewey 1992).

In the same way, the underestimation of some of the medium and high age-0 perch densities was due to the scarcity of samples with such densities. Because the network's predictive capabilities are limited by the information in the training data set, the limited number of points where fish were found

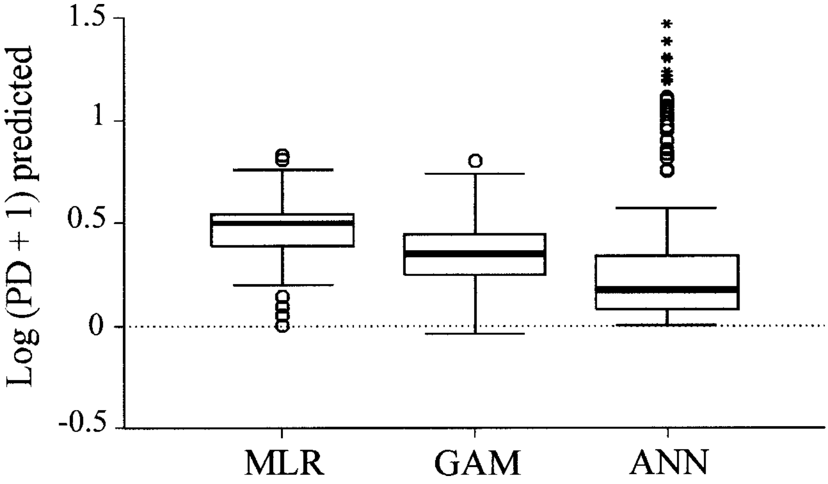

FIGURE 4.—Results of model validations using 183 samples in which no perch were found. The dashed line represents the actual null values. The box plots represent the predictions by the multiple linear regression (MLR) model, the generalized additive model (GAM), and the artificial neural network (ANN) model. The bold line within each box shows the 50th percentile; the top and bottom lines show the 75th and 25th percentiles, respectively. The whiskers represent the 10th and 90th percentiles, the open circles samples in which perch density was more than 1.5 box lengths from the 75th percentile (''outliers''), and asterisks samples in which perch density was more than 3 box lengths from the 75th percentile (''extremes'').

and unmeasured but potentially important biotic or abiotic variables (such as water temperature and species interactions) did not allow the network to deal with all the factors that might influence a fish distribution. Nevertheless, a large proportion of the points in the scatterplot are distributed along the diagonal representing the best predictions, and nearly all of the samples with medium and high observed values of age-0 perch density were predicted to be suitable habitats (even if some densities were underestimated). We can thus conclude that the ANN model was able to accurately predict age-0 perch density in Lake Pareloup on the basis of the environmental variables in the model, whereas the MLR model and the GAM proved to have serious shortcomings, as indicated by both their *r* values and their $SD_{PE}-SD_T$ ratios. In addition, the predictive accuracy of the models was checked by means of the validation test on the 183 null samples. In that test, most of the perch densities predicted by the MLR model and the GAM were aberrant, whereas the ANN model reliably predicted null samples, with most predicted densities close to null. However, the ANN model predicted that some null samples would be densely populated. Again, these samples could correspond to areas in which the environment is suitable for perch but which the species happens not to have colonized, so that these predictions constitute estimates of the potential fish density in those samples.

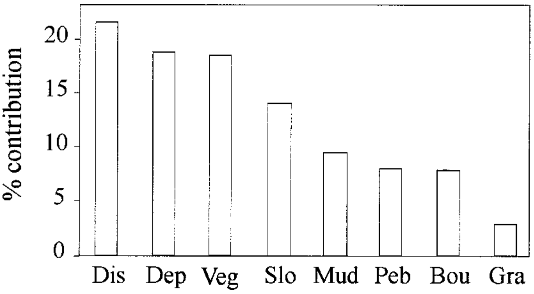

FIGURE 5.—Percentage contributions of the eight independent variables obtained by applying the Garson−Goh algorithm to the ANN model results. Abbreviations are given in the caption to Figure 1.

In terms of the influence of the individual environmental variables on the predictions of age-0 Eurasian perch density as assessed through the Garson−Goh algorithm, the results were consistent with those obtained using the Lowess fitting of the observed data. These results indicate the importance of flooded vegetation, distance from the bank, depth, and bottom slope on the habitat use of age-0 perch and demonstrate the suitability of the ANN approach for describing nonlinear interactions between variables in complex ecological systems.

Both analyses (Lowess and the comparison of the MLR, GAM and ANN model results) showed that the relationships between perch abundance and environmental characteristics are highly nonlinear, which explains why traditional linear statistical methods often fail to accurately define fish habitat use. Furthermore, ANNs constitute an efficient predictive tool that also provides relevant ecological information. Consequently, ANNs and Lowess can be used together as predictive and explanatory tools when common statistical methods are limited by the nonlinearity of the data. More particularly, this study shows that overall trends in fish distribution can be accurately assessed by Lowess and ANNs using a few relevant environmental variables, thus providing insights into the ecological meaning of such a distribution on a small spatial scale.

Acknowledgments

We are grateful to H. Collier for correcting the English version.

References

Bain, M. B., J. T. Finn, and H. E. Booke. 1985. A quantitative method for sampling riverine microhabitats by electrofishing. North American Journal of Fisheries Management 5:489–493.

Baran, P., S. Lek, M. Delacoste, and A. Belaud. 1996. Stochastic models that predict trout population densities or biomass on a mesohabitat scale. Hydrobiologia 337:1–9.

Beecher, H. A., T. H. Johnson, and J. P. Carleton. 1993. Predicting microdistributions of steelhead (*Oncorhynchus mykiss*) parr from depth and velocity preference criteria: test of an assumption of the instream flow incremental methodology. Canadian Journal of Fisheries and Aquatic Sciences 50:1380–1387.

Brabrand, A., and B. Faafeng. 1994. Habitat shift in roach (*Rutilus rutilus*) induced by the introduction of pike-perch (*Stizostedion lucioperca*). Internationale Vereinigung für theoretische und angewandte Limnologie Verhandlungen 25:2123.

Brosse, S. 1999. Habitat, dynamique spatiale, et structure des communautés pisciaires en milieu lacustre. Etude de la retenue de Pareloup (Aveyron, France). [Habitat, spatial dynamics, and structure of fish communities in a lacustrine environment: a study of Pareloup Reservoir (Aveyron, France)]. Doctoral dissertation. University of Toulouse 3, Toulouse, France.

Brosse, S., F. Dauba, T. Oberdorff, and S. Lek. 1999a. Influence of some topographical variables on the spatial distribution of lake fish during summer stratification. Archiv für Hydrobiologie 145:359–371.

Brosse, S., J. F. Guégan, J. N. Tourenq, and S. Lek. 1999b. The use of artificial neural networks to assess fish abundance and spatial occupancy in the littoral zone of a mesotrophic lake. Ecological Modelling 120:299–311.

Brosse, S., S. Lek, and F. Dauba. 1999c. Predicting fish distribution in a mesotrophic lake by hydroacoustic survey and artificial neural networks. Limnology and Oceanography 44:1293–1303.

Cailleux, A. 1954. Limites dimentionnelles des noms des fractions granulométriques. [Dimensional limits of the names of granulometric fractions]. Bulletin de la Société Géologique de France 4:643–646.

Christensen, B., and L. Persson. 1993. Species-specific antipredator behaviours: effect of prey choice in different habitats. Behavioral Ecology and Sociobiology 32:1–9.

Cleveland, W. S. 1979. Robust locally-weighted regression and scatterplot smoothing. Journal of the American Statistical Association 74:829–836.

Copp, G. H. 1989. Electrofishing for fish larvae and juveniles: equipment modifications for increased efficiency with short fishes. Aquaculture and Fisheries Management 20:453–462.

Copp, G. H. 1992. An empirical model for predicting microhabitat of 0+ juvenile fishes in a lowland river catchment. Oecologia 91:338–345.

Dewey, M. R. 1992. Effectiveness of a drop net, a pop net, and an electrofishing frame for collecting quantitative samples of juvenile fishes in vegetation. North American Journal of Fisheries Management 12:808–813.

Diehl, S. 1993. Effects of habitat structure on resource availability, diet, and growth of benthivorous perch, *Perca fluviatilis*. Oikos 67:403–414.

Dubois, C., C. Richeux, J. N. Tourenq, and C. Lejolivet. 1994. Régime et comportement alimentaire du gardon (*Rutilus rutilus*, L.) et de la perche (*Perca fluviatilis*, L.) de la retenue de Pareloup. I. Les alevins. [Diet and feeding behavior of the roach *Rutilus rutilus* and Eurasian perch *Perca fluviatilis* in Pareloup Reservoir. I. Alevins.] Hydroécologie Appliquée 6: 227–242.

Efron, B. 1983. Estimating the error rate of a prediction rule: some improvements on cross-validation. Journal of the American Statistical Association 78:316–331.

Eklöv, P. 1997. Effects of habitat complexity and prey abundance on the spatial and temporal distributions of perch (*Perca fluviatilis*) and pike (*Esox lucius*). Canadian Journal of Fisheries and Aquatic Sciences 54:1520–1531.

Eklöv, P., and S. F. Hamrin. 1989. Predator efficiency and prey selection: interactions between pike (*Esox lucius*), perch (*Perca fluviatilis*) and rudd (*Scardinius erythrophthalmus*). Oikos 56:149–156.

Eklöv, P., and L. Persson. 1996. The response of prey to the risk of predation: proximate cues for refuging juvenile fish. Animal Behaviour 51:105–115.

Englund, G. 1997. Importance of spatial scale and prey movements in predator caging experiments. Ecology 78:2316–2325.

Fischer, P., and R. Eckmann. 1997. Spatial distribution of littoral fish species in a large European lake, Lake

Constance, Germany. Archiv für Hydrobiologie 140:91–116.

Garson, G. D. 1991. Interpreting neural network connection weights. Artificial Intelligence Expert 6:47–51.

Geman, S., E. Bienenstock, and R. Doursat. 1992. Neural networks and the bias/variance dilema. Neural Computation 4:1–58.

Gliwicz, Z. M., and A. J. Warsaw. 1992. Diel migrations of juvenile fish: a ghost of predation past or present? Archiv für Hydrobiologie 124:385–410.

Goh, A. T. C. 1995. Back-propagation neural networks for modeling complex systems. Artificial Intelligence Engineer 9:143–151.

Grossman, G. D., A. De Sostoa, M. C. Freeman, and J. Lobon-Cervia. 1987. Microhabitat use in a Mediterranean riverine fish assemblage: fishes of the lower Matarrana. Oecologia 73:490–500.

Guégan, J. F., S. Lek, and T. Oberdorff. 1998. Energy availability and habitat heterogeneity predict global riverine fish diversity. Nature (London) 391:382–384.

Hammer, C. 1985. Feeding behaviour of roach (*Rutilus rutilus*) and the fry of perch (*Perca fluviatilis*) in Lake Lankau. Archiv für Hydrobiologie 103:61–74.

Hasler, A. D., and J. R. Villemonte. 1953. Observations on the daily movements in fishes. Science 118:321–322.

Hastie, T. J., and R. J. Tibshirani. 1990. Generalized additive models. Chapman and Hall, London.

Imbrock, F., A. Appenzeller, and R. Eckmann. 1996. Diel and seasonal distribution of perch in Lake Constance: a hydroacoustic study and in situ observations. Journal of Fish Biology 49:1–13.

Jacobsen, L., and S. Berg. 1998. Diel variation in habitat use by planktivorous in field enclosure experiments: the effect of submerged macrophytes and predation. Journal of Fish Biology 53:1207–1219.

Jacobsen, L., and M. R. Perrow. 1998. Predation risk from piscivorous fish influencing the diel use of macrophytes by planktivorous fish in experimental ponds. Ecology of Freshwater Fish 7:78–86.

James, F. C., C. E. McCulloch, and D. A. Wiedenfeld. 1996. New approaches to the analysis of population trends in land birds. Ecology 77:13–27.

Kohavi, R. 1995. A study of the cross-validation and bootstrap for accuracy estimation and model selection. *In* Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI). Montreal, Canada.

Laffaille, P., S. Brosse, S. Gabas, and S. Lek. 2001. Fish spatial distribution in the littoral zone of Lake Pareloup (France) during summer. Archiv für Hydrobiologie 153:127–144.

Lek, S., M. Delacoste, P. Baran, I. Dimopoulos, J. Lauga, and S. Aulagner. 1996. Application of neural networks to modelling non-linear relationships in ecology. Ecological Modelling 90:39–52.

Lek, S., and J. F. Guégan. 1999. Artificial neural networks as a tool in ecological modelling: an introduction. Ecological Modelling 120:65–73.

Lek, S., and J. F. Guégan. 2000. Artificial neuronal networks: applications to ecology and evolution. Springer-Verlag, Berlin.

Machacek, J., and J. Matena. 1997. Diurnal feeding patterns of age-0 perch (*Perca fluviatilis*) and roach (*Rutilus rutilus*) in a steep-sided reservoir. Archiv für Hydrobiologie 49:59–70.

Mastrorillo, S., F. Dauba, and A. Belaud. 1996. Utilisation des microhabitats par le vairon, le goujon et la loche franche dans trois rivières du sud-ouest de la France. [Utilization of microhabitats by the. . . in three rivers in southwestern France.] Annales de Limnologie 32:185–195.

Mastrorillo, S., S. Lek, F. Dauba, and A. Belaud. 1997. The use of artificial neural networks to predict the presence of small-bodied fish in a river. Freshwater Biology 38:237–246.

Nelva, A., H. Persat, and D. Chessel. 1979. Une nouvelle méthode d'étude des peuplements ichtyologiques dans les grands cours d'eau par échantillonnage ponctuel d'abondance. [A new method of studying colonizations by fish in large rivers by point-sampling abundance.] Comptes Rendus de l'Académie des Sciences Serie III 289:1295–1298.

Pennington, M. 1996. Estimating the mean and variance from highly skewed marine data. U.S. National Marine Fisheries Service Fishery Bulletin 94:498–505.

Persson, L., and P. Eklöv. 1995. Prey refuges affecting interactions between piscivorous perch and juvenile perch and roach. Ecology 76:70–81.

Rossier, O. 1995. Spatial and temporal separation of littoral zone fishes of Lake Geneva (Switzerland–France). Hydrobiologia 300/301:321–327.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1986. Learning representations by back-propagating error. Nature (London) 323:533–536.

Savino, J. F., and R. A. Stein. 1989. Behavioural interactions between fish predators and their prey; effects of plant density. Animal Behaviour 37:311–321.

Statistica. 2000. Statistica for Windows, computer program manual. StatSoft, Inc., Tulsa, Oklahoma.

ter Braak, C. J. F., and C. W. N. Looman. 1995. Regression. Pages 29–77 *in* R. G. H. Jongman, C. J. F. ter Braak, and O. F. R. Van Tongeren, editors. Data analysis in community and landscape ecology. Cambridge University Press, Wageningen, The Netherlands.

Townsend, C. R., J. L. Harper, and M. Begon. 2000. Essentials of ecology. Blackwell Scientific Publications, Malden, Massachusetts.

Trexler, J. C., and J. Travis. 1993. Non-traditional regression analyses. Ecology 74:1629–1637.

Winfield, I. J. 1990. Predation pressure from above: observations on the activities of piscivorous birds at a shallow eutrophic lake. Hydrobiologia 191: 223–231.