Predicting the abundance of minnow *Phoxinus phoxinus* (Cyprinidae) in the River Ariège (France) using artificial neural networks

Sylvain Mastrorillo [(1)], Sovan Lek [(2)] and Francis Dauba [(1)]

[(1)] *Laboratoire d'Ingénierie agronomique, Equipe Environnement aquatique et Aquaculture, Ecole Nationale Supérieure Agronomique de Toulouse, 145, avenue de Muret, 31076 Toulouse cedex, France, Email: mastrori@ensat.fr*

[(2)] *CNRS-UMR 5576 CESAC, Bât. 4R3, Université Paul Sabatier, 118, route de Narbonne, 31062 Toulouse cedex, France, Email : lek@cict.fr.*

Received November 15, 1996; accepted February 3, 1997.

Mastrorillo S., S. Lek, F. Dauba. *Aquat. Living Resour.*, 1997, **10**, 169-176.

Abstract

The study of abundance of small-bodied species of fish such as minnow is important because these species play an important role in the food-web dynamics of small streams. In this work, we propose the use of an Artificial Neural Network (ANN) to the modelling and prediction of abundance in minnow *Phoxinus phoxinus* using 10 environmental microhabitat variables: distance from the bank, percentage of boulders, pebbles, gravel, sand, mud, marl, cover respectively, depth and velocity. A total of 372 points were randomly chosen from a total of 465 electrofished point samples to establish a ANN model. A validation holdout of the training of the ANN was undertaken with testing on 93 other sampling points. On the test set, the prediction performance was 92 %. Our study showed the advantages of the back-propagation procedure of the neural network in the field of stochastic approaches to ecology of coarse fishes. The limitations of the neural network approaches as well as statistical and ecological perspectives are discussed.

Keywords: *Phoxinus phoxinus*, point abundance sampling, microhabitat, artificial neural networks, predictive model.

Prédiction de l'abondance du vairon Phoxinus phoxinus *(Cyprinidae) dans la rivière Ariège (France) au moyen des réseaux de neurones artificiels.*

Résumé

L'étude de l'abondance des petites espèces de poissons comme le vairon est nécessaire car ces espèces jouent un rôle important dans la dynamique de la chaîne alimentaire des rivières à salmonidés et cyprinidés rhéophiles. Dans cette étude, nous proposons l'utilisation des réseaux de neurones artificiels pour modéliser et prédire l'abondance du vairon *Phoxinus phoxinus* à l'échelle du microhabitat. Les poissons ont été échantillonnés par pêche électrique selon la méthode de l'échantillonnage ponctuel d'abondance. Dix variables du milieu ont été prises en compte pour décrire l'environnement ponctuel des poissons : la distance à la berge, le pourcentage de blocs, de galets, de gravier, de sable, de vase et de marne, la profondeur et la vitesse du courant. Sur un total de 465 échantillons ponctuels d'abondance réalisés, nous avons établi le modèle de réseaux de neurones utilisant la procédure de validation croisée : par le processus de tirage aléatoire, nous avons isolé 372 échantillons (soit 80 %) comme l'ensemble d'apprentissage et 93 échantillons restant (20 %) comme l'ensemble de test. Sur l'ensemble de test, la performance de prédiction a atteint 92 %. Notre étude a montré ainsi les avantages de l'algorithme de la rétropropagation de gradient du réseau de neurones pour une approche stochastique de l'écologie des poissons non-salmonidés.

Mots-clés : *Phoxinus phoxinus*, échantillonnage ponctuel d'abondance, microhabitat, réseau de neurones artificiels, modèle prédictif.

INTRODUCTION

Many ecologists have tried to link the abundance of animal populations to habitat characteristics (Verner *et al.* 1986), with a number of theoretical models of animal abundance based on habitat character proposed (McArthur *et al.*, 1966; Fretwell, 1972; Cody, 197 ; Tilman 1982; MacNally, 1983; Schoener, 1983). The estimation of animal abundance plays an important role in ecology, particularly in ichthyology. A variety of multivariate techniques have been used to investigate how habitat is related to abundance, including several methods of ordination and canonical analysis as well as univariate and multivariate linear, curvilinear, and logistic regressions. Complete and critical reviews of statistical methods by Fausch *et al.* (1988) and James and McCulloch (1990) have assumed that species-habitat relationships are smooth, continuous, and either linear or simple polynomial. These conventional techniques, based notably on multiple regression, are capable of solving many problems, but sometimes reveal serious shortcomings, in particular the fact that relationships between variables in the environmental sciences are often non-linear, whereas the methods generally used are based on linear principles.

Non-linear transformations of variables (logarithmic, power or exponential functions) can improve the results appreciably, but they are often only partially successful. In general, the relationships between species and environmental variables are non-linear or non-monotonous, and techniques based on correlation coefficients is often inappropriate (Ter Braak, 1986). The neural network, with the error backpropagation procedure, is the basis of an interesting methodology that could be used in the same field as regression analysis, particularly with the non-linear relations (Rumelhart *et al.* 1986). Nevertheless, few applications of this new technology have been published in the ecological sciences, in contrast to the physical or chemical sciences (Smits *et al.*, 1992; Albiol *et al.*, 1995; Lerner *et al.*, 1994; Faraggi and Simon, 1995). For example, artificial neural networks have been used to model the greenhouse climate (Seginer *et al.*, 1994), the identification of the major goals of underwater acoustics (Casselman, 1994), the prediction of density of brown trout (*Salmo trutta*) redds (Lek *et al.*, 1996a), and the prediction of density and biomass of trout (Baran *et al.*, 1996; Lek *et al.*, 1996b). Artificial neural networks may be applied to different kinds of problems such as pattern classification, interpretation, generalisation or calibration.

Until recently, riverine ichthyological investigations have concentrated on salmonids, with relatively little study on the species-habitat interactions of accompanying species to the salmonids, such as minnow (*e.g.* Copp, 1992; Mastrorillo *et al.*, 1996). Nonetheless, the problems of perturbations to an ecosystem are not always revealed by commercially exploited species, which generally have extended life spans. Small and young fishes are more sensitive to environmental change and therefore are better functional describers of ecosystems succession and perturbations (Copp *et al.*, 1991).

In the present paper, we examine the use of the neural networks principle for regression problem with aim of analysing the level of relationships between microhabitat variables and the abundance of the minnow *Phoxinus phoxinus* (Cyprinidae) in the River Ariège (South-Western France) and in doing so we propose a basis of the development of predicative model for estimating the abundance of the minnow using the neural network approach.

MATERIAL AND METHODS

The Ariège catchment drains an area of approximately 3 900 km^2 in South-Western France. The River Ariège has its source in Andorra and flows into the River Garonne upstream of Toulouse. Sampling was carried out on 4 sites (Fig. 1, Table 1). At each studied site, fish and environmental variables were sampled at numerous small points using Point Abundance Sampling by Electrofishing (PASE) (Nelva *et al.*, 1979; Copp and Peñáz, 1988; Persat and Copp, 1989; Copp and Garner, 1995). At low flow, the electrofishing was carried out using a generator-powered electroshocker (1 A, 300-400 V), with the operator moving on foot in an upstream direction along a trajectory perpendicular to the river bank (semi-transect towards the open river). The sampling points were chosen randomly, with the distance between each sampling point being about 5 m. When the depth or the water velocity became too high, the operator continued with a new sampling point 5 m upstream on a new semi-transect. At each point, the activated anode was submerged for a few seconds, with fishes collected using a dip net. The fishes were identified, measured and returned to

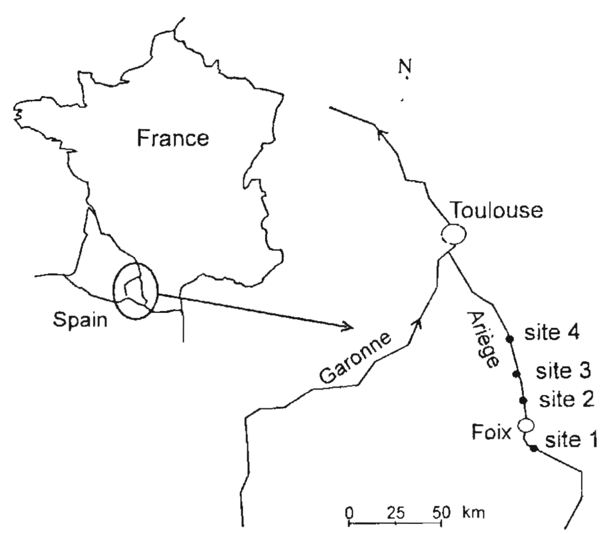

Figure 1. – Location of study sites (dots) on the River Ariège.

Table 1. – Main characters of study sites on the River Ariège, France.

Sites	Site code	Mean elevation m	Mean slope %	Mean width m	Catchment area km^2	Annual mean flow $m^3.s^{-1}$
Pont du Diable	site 1	450	1.60	22	1 145	4
Varilhes	site 2	320	0.30	60	1 435	44
Pamiers	site 3	270	0.50	52	1 500	46
Saverdun	site 4	220	0.16	43	1 800	53

the water. Immediately after processing of the fish at a given point, microhabitat character was evaluated using 10 environmental variables: distance from the bank, % cover, % boulders (> 20 cm), % pebbles (2-20 cm), % gravel (2 mm-2 cm), % sand (50 μm-2 mm), % mud (< 50 μm), % marl, depth and velocity. Bottom composition was evaluated as a percentage of the surface sampled. Depth and water velocity were measured at the centre of each sampling point. A total of 465 point samples was collected throughout the river catchment.

Classical statistic analysis

Univariate, bivariate and multivariate analysis of data were performed by the SPSS Software® "release 6.0 (Norusis, 1993) for Windows. The univariate analysis consisted of the determination of parametric (mean, standard deviation and coefficient of variation) and nonparametric (minimum, maximum, median and quartiles) statistical parameters. In the bivariate analysis, we studied the correlation between variables using Spearman's coefficients (values and probabilities of significant at 5 and 1 % of confidence intervals). In the multivariate analysis, the stepwise multiple linear regression (MLR) procedures were applied. The diagnosis of the studentized residuals (normality and independence) was used to test the validity of the determination coefficients obtained with each of the models (Weisberg, 1980; Tomassone *et al.*, 1983).

Artificial neural network (ANN)

A predictive model of minnow abundance from habitat characteristics was developed on one of the principles of ANN, the mathematical algorithm of backpropagation (Rumelhart *et al.*, 1986), transforming the activation into a non-linear type response. A network with backpropagation typically comprises three kinds of layers of neurons: an input layer, one or several hidden layers, and an output layer, which consists of one or several neurons (Fig. 2). All the neurons of a given layer, except those of the last, emit an axon to each neuron of the layer down stream. The network is said to be entirely interconnected by layers. In the majority of cases, ANN with one hidden layer and a sufficient number of nodes is capable achieving any mapping with an arbitrary degree of accuracy (Hornik *et al.*, 1989; Bhat and McAvoy, 1992).

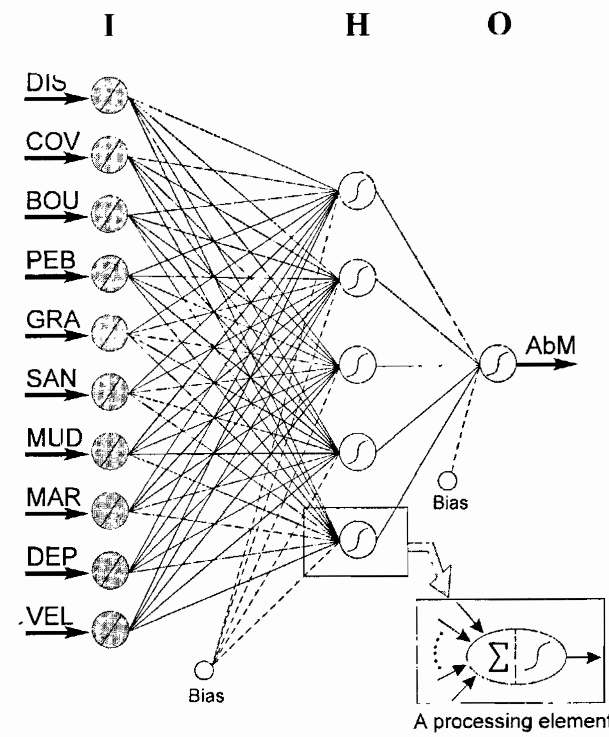

Figure 2. – Representation of the structure of the neural network used. Ten input nodes (I), five hidden layer nodes (H) and one output node (O) are shown. DIS: distance from the bank; COV: % cover; BOU: % boulders; PEB: % pebbles; GRA: % gravel; SAN: % sand; MUD: % mud; MAR: % marl; DEP: depth; VEL: velocity; Abm: abundance of minnow.

The input layer comprises n neurons coding the n elements of information ($X1 \ldots Xn$) at the input of the network. The number of neurons in the hidden layer is chosen by the user according to the reliability of the results desired. However, if one proceeds to model the network in a more complex form, i.e. with a large number of neurons, one can obtain a strong tendency to overfit the training dataset, resulting in a poor performance on the test set (Smith, 1994). Finally, the output layer comprises a single neuron responsible for the calculation of the value to be predicted. A modifiable weight is attributed to each connection between two successive neuron layers in the course of training, which is determined by the sets of data at the entry and exit.

The input layer having the number of nodes equal to the number of inputs in the problem at hand is merely used for distributing the input of the network. A neuron of the hidden layer and the output layer perform a weighted summation of the inputs received from all the upstream layer nodes. The summation is computed as follows:

$$a_j = \sum_{i=1}^{I} X_i W_{ji} \tag{1}$$

where a_j, activation of the jth downstream neuron; X_i, output value of the ith neuron of the previous layer; W_{ji}, synaptic weight of the connection between the ith neuron of the previous layer and the jth neuron of the current layer.

After computing the weighted sum, as shown in equation (1) the way that the output of hidden and output neurons is calculated from their net input depends upon the type of activation function being used in the ANN. Most backpropagation ANNs use the additive sigmoid type function, because of its non-linearity:

$$X_j = f(a_j) = \frac{1}{1 + \exp^{-a_j}} \tag{2}$$

The technique of backpropagation is related to supervised learning (to learn, the network has to know the reply that it should have given). It then modifies the intensity of the connection so as to minimise the error of the reply considered. The estimation of the error signal differs according to the layers considered. Many articles, notably by Rumelhart *et al.* (1986), Carpenter (1989) and Weigend *et al.* (1992) detail the algorithms of backpropagation of errors. Note finally that one can use parameters such as η (learning coefficient) and α (momentum), which serve to accelerate learning whilst avoiding the network from falling into local minima. The learning of the network continues until minimisation of the sum of the square of the errors (SSE) is given by the relationship:

$$SEE = \frac{1}{2} \sum_{j=1}^{N} (Y_j - \hat{Y}_j)^2 \tag{3}$$

where Y_j, value expected at the output of the network ("theoretical value"); $\hat{Y}$, value calculated by the network (neuron of the output layer); $j = 1 \ldots N$, number of recordings. The computing programs were performed under Matlab® environment for Windows®, a matricial calculation software.

Adaptation of the input data is necessary because they have orders of magnitude that differ greatly according to the variables. To standardise the measurement scales, inputs are converted into a reduced-centred variable. The dependent variable is also converted in the range [0 ... 1] to adapt it to the demands of the transfer function used (sigmoid function). To compare the results obtained with multiple linear regression and with neural network, a first application was made on the wole of the database (465 patterns, *i.e.* all point samples). Then, to justify the predictive quality of the models, the two procedures were applied to a set of 372 patterns (*i.e.* 80 %) randomly chosen in the total database to obtain multivariate models with regression analysis or for ANN training. After this first training phase, we presented the 93 new patterns (*i.e.* 20 %) to test the multivariate models or the neural network.

The performance criterion normally used is the coefficient of determination (r^2), to judge the quality of mathematical models. However, because of the rarity of strong values, we preferred to use an index of performance which is characterised by the number of individuals correctly predicted by the model. On the scatterplot of correlation graph between estimated and observed values by the model, the prediction will be good if points are placed inside the interval defined by the perfect fit line (coordinate 1:1) within more or less 10 %.

A disadvantage of ANN in comparison with MLR models is their lack of explanation power. MLR analysis can identify the contribution of each individual input in determining the output and also can give some measures of confidence about the estimated coefficients. On the other hand, currently there is no theoretical or practical way of accurately interpreting the weights in ANN. For example, weights cannot be interpreted as a regression coefficient, nor difficulty used to compute causal impacts or elasticities. Therefore, ANN are generally better suited for forecasting or prediction rather than for policy analysis. But in ecology, it is necessary to know the explanatory variable impacts. Some authors have proposed methods allowing of determination the impact of variables in entry (Garson, 1991; Goh, 1995; Lek *et al.* 1996*a*, *b*). In this work, an experimental approach can be used to determine the response of the model to each of the input variables separately by applying the technique describes by Lek *et al.* (1996*a*, *b*).

RESULTS

Of the 465 point samples collected in the River Ariège, 3 571 specimens were caught representing 12 fish species of which 1 740 were minnow (*i.e.* 48.7 % of the fish assemblage). The other species were: stone loach (*Barbatula barbatula*), gudgeon (*Gobio gobio*), brown trout (*Salmo trutta fario*), salmon (*Salmo salar*), dace (*Leuciscus leuciscus*), bullhead (*Cottus gobio*), brook lamprey (*Lampetra planeri*), chub (*Leuciscus cephalus*), sofie (sometimes referred to as south-western European nase) (*Chondrostoma toxostoma*), barbel (*Barbus barbus*) and pumpkinseed (*Lepomis gibbosus*).

The abundance of the minnow varied between sampling points. Minimum abundance, maximum abundance and mean abundance were respectively 0, 149 and 3.75 specimens, with a high variation coefficient of 303 %, a very current result in ecology.

Relationship between the abundance of minnow and environmental variables

Minnow abundance was significantly related to 7 variables: positively with % sand (r^2 = 0.346; p < 0.01), % cover (r^2 = 0.287; p < 0.01), distance from the bank (r^2 = 0.197; p < 0.01), % mud (r^2 = 0.116; p < 0.05), and negatively with velocity (r^2 = -0.186; p < 0.01), % pebbles (r^2 = – 0.143; p < 0.01) and depth (r^2 = – 0.100; p < 0.05). The correlation was not significant for the other three variables (% boulders, % gravel, % marl; p > 0.05).

Stepwise regression analysis

Without any transformation of variables, the usual method of multiple linear regression analysis was performed with the 465 point samples to see if a significant correlation could be obtained with this classical method. The stepwise procedure performed with SPSS selected 4 variables by 4 steps: step 1, % sand (r^2 = 0.12, $F_{1,463}$ = 62.96, p < 0.001), step 2, % cover (r^2 = 0.181, $F_{2,462}$ = 51.07, p < 0.001), step 3, depth (r^2 = 0.195, $F_{3,461}$ = 37.29, p < 0.001) and step 4, % gravel (r^2 = 0.203, $F_{4,460}$ = 29.23, p < 0.001). With all of the 10 environmental variables, we have obtained a determination coefficient of only 0.218 ($F_{9,455}$ = 14.07, p < 0.001). All MLR models are significant at 0.1 %, but low determination coefficient testify the low percentages of explained variance (< 20 %). The supplementary variable addition as compared to the stepwise regression contributes only very little to the improvement of results (22 % of explained variance).

Thus, we can affirm that these supplementary variables do not improve the model in a linear process. The classic modelling method suggests the non-linear transformation of the variables, independent and/or dependent, such as logarithmic, power or exponential transformations to improve the percentage of the explained variances of the model; but in this case, you have to pay attention to the biological meaning.

Artificial neural network (ANN)

Results of the ANN obtained with 500 iterations and 5 neurons on the hidden layer yielded a coefficient of determination (r^2) of 0.946 for the regression between observed and estimated values (Fig. 3), indicating that the ANN provides satisfactory results over the whole entire of values for the dependent variable. The points are well aligned on the diagonal of the perfect fit line (coordinate 1:1). Although weakly represented, the strong values of the output variable are aligned around this same perfect fit line, with a few outliers (Fig. 3). Some weak values were slightly underestimated. By applying an error interval of 10 %, a good classification rate of 97.2 % was obtained, *i.e.* only 13 observations were badly classified. At 5 %, the rate is 92.2 %, *i.e.* 36 observations badly classified due mainly to misjudgements.

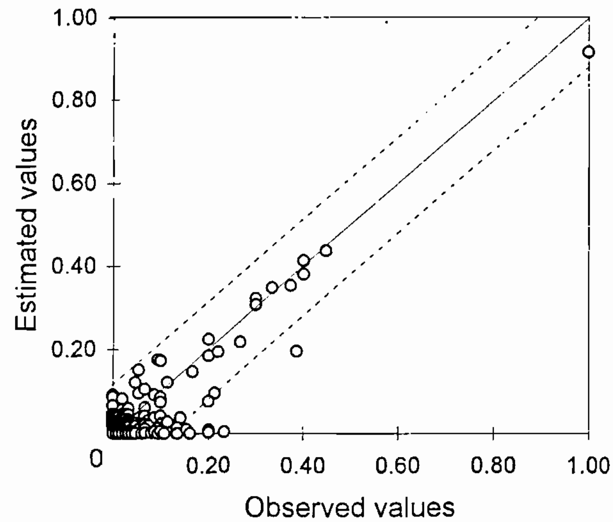

Figure 3. – Scatterplot between observed values and estimated values by the model. The solid line indicates the perfect fit line (coordinate 1:1), the two dashed lines indicate the limits of the interval of 10 % prediction error.

Residuals have an average of 0.005 and a standard deviation of 0.034 (Fig. 4 *a*), with more than half of the observations having an error of zero. The distribution of residuals does not seem to be normal, as there is an exaggerated clustering of residuals toward the centre (averaging of zero) and a straggling tail towards large positive values. Thus, the normality assumption may be violated. Lilliefors test of normality gave a maximum difference of 0.305, p < 0.001. The study of the relationship between residuals and values estimated by the model shows complete independence. The coefficient of determination was negligible (r^2 = 0.0008); the residuals are well distributed on either side the horizontal line (ordinate) representing the residual mean (Fig. 4 *b*).

The influence of the 10 independent environmental variables on the abundance of minnows in the ANN modelling is illustrated by 5 curves (Fig. 5), representing 4 sensitivity (or contribution) types:

Linear increasing: % gravel and % sand. The abundance of minnow increases gradually as the value of the independent variables increases.

Linear decreasing: distance, % pebbles and velocity. The abundance of minnow decreases gradually as the value of the independent variables increases.

Gaussian: % boulders, % mud, % cover and depth. The abundance of minnow increases as the independent variable approaches its median values (level 5 for % cover and % boulders; level 8 for depth; level 6 for % mud), and decreases as the variables values approach their extremes.

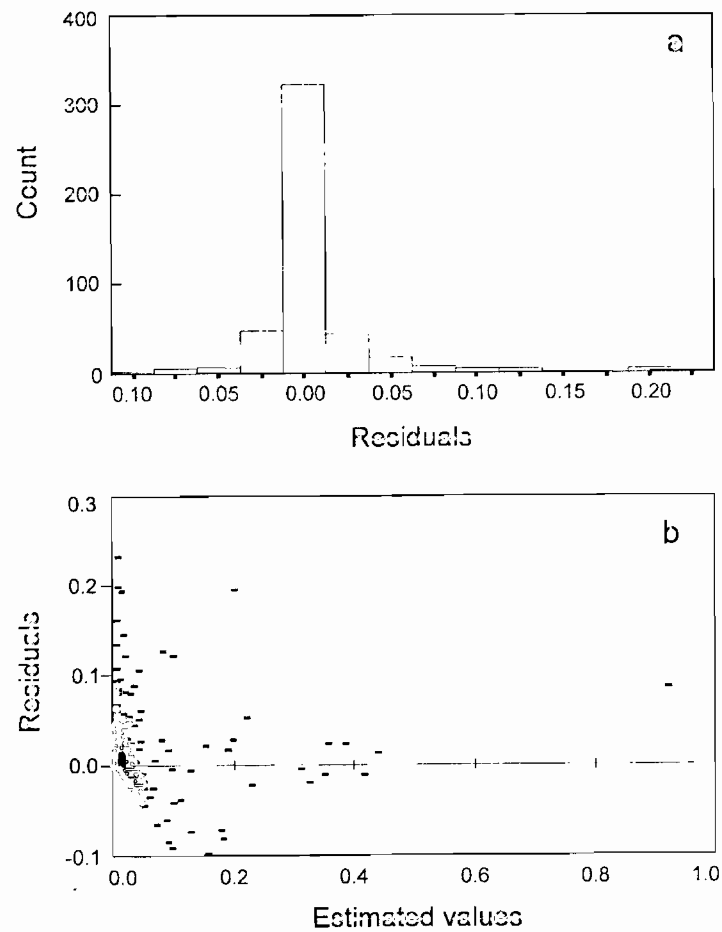

Figure 4. – a. Distribution of residuals (observed values – estimated values). b. Relationship between the residuals and the values estimated by the model.

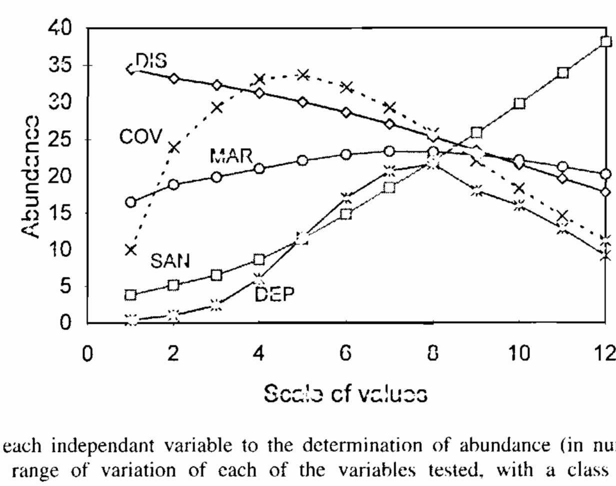

Figure 5. – Contribution profile for each independant variable to the determination of abundance (in numbers) of minnow by Artificial neural networks. The 12 values cover the range of variation of each of the variables tested, with a class interval that varied according to the variable. Boulders, pebbles, gravel, mud and velocity are not represented here.

Constant: % marl. The abundance of minnow remains relatively constant, at best mildly Gaussian, along the range of values for the independent variable.

The predictive power of the different models determined from five training fractions was tested on five test fractions from which they were independent. By applying an interval of error of 10 %, a good classification rate was obtained for the test data (92.5 %, 93.6 %, 89.3 %, 91.4 % and 93.0 %), with a mean good classification rate of 91.9 %, and a weak standard deviation (1.69). These results reflect great stability of the ANN model.

DISCUSSION

Until recently, riverine ichthyological investigations have concentrated on salmonids, with relatively little study on the species-habitat interactions of accompanying species to the salmonids, such as minnow (Copp, 1992; Mastrorillo *et al.*, 1996). The study of accompanying species is important, because they play an important role in the food-web dynamics of small streams. Our study on the prediction of minnow abundance in an important Pyrenean stream shows that variations in abundance are closely linked to a number of environmental variables such as depth, current velocity, bottom composition and instream cover, all variables retained for studies of fish habitat (Gorman and Karr, 1978; Angermeier, 1995; Copp, 1989, 1992; Grossman and de Sostoa, 1994; Pouilly, 1994*b*; Garner, 1996*a*, *b*).

In the Ariège, the abundance of minnow increased proportionally with the quantity of gravel and sand and decreased proportionally with the distance from the bank, the current velocity and the presence of pebbles. The strongest abundances were also observed at average values of depth, of instream cover, percentage of boulders and of mud. Our results corroborate the habitat preference curves for minnows with respect to water depth, velocity and bottom composition established from electrofishing studies by Pouilly (1994*a*) on three rivers of the Rhône catchment and the ecological profiles study by Mastrorillo *et al.* (1996).

The use of backpropagation in the ANN led to the development of stochastic tools to predict the abundance of minnows from habitat features on a microhabitat scale. The selection of input variables introduced in the modelling procedures, their ecological signification for a coarse fish such as minnow, and the constitution of a testing set of data to access models precision are important elements for this type of approach. A simple backpropagation of ANN is capable of modelling ecological problems involving nonlinear variables. After training a set of patterns, the ANN models were able to produce reasonably accurate prediction. The determination coefficients of test set were lower than in training set but still remained clearly significant.

The complexity of the relationships within an ecosystem, and more particularly between biotic and abiotic components, needs increasingly sophisticated analytical techniques. The ANN has demonstrated its learning and predicative capability for the modelling of cyprinid fish abundance such as minnow. However, to use this new procedure, one requires a very large database. Hence, artificial neural networks identify themselves as a potential tool for the solution of a wide variety of problems in the field of aquatic living resources.

Acknowledgments

We thank two referees, P. Boët and G.H. Copp, for their critical comments on this manuscript.

REFERENCES

Albiol J., C. Campmajo, C. Casas, M. Poch 1995. Biomass estimation in plant cell cultures: a neural network approach. *Biotechnol. Prog.* **11**, 88-92.

Angermeier P.L., R.A. Smogor 1995. Estimating number of species and relative abundances in stream-fish communities: effects of sampling effort and discontinuous spatial distributions. *Can. J. Fish. Aquat. Sci.* **52**, 936-949.

Baran P., S. Lek, M. Delacoste, A. Belaud 1996. Stochastic models that predict trouts population densities or biomass on macrohabitat scale. *Hydrobiologia* **337**, 1-9.

Bhat N.V., T.J. McAvoy 1992. Determining model structure for neural models by network stripping. *Comput. Chem. Eng.* **116**, 271-281.

Carpenter G.A. 1989. Neural network models for pattern recognition and associative memory. *Neural Networks* **2**, 243-257.

Casselman F.L., D.F. Freeman, D.A. Kerrigan, S.C. Lane, D.M. Magley, N.H. Millstrom, C.R. Roy 1994. A neural network-based underwater acoustic application. *In:* Proc. International Conference on Neural Networks, *IEEE* Orlando, 3409-3414.

Cody M.L. 1974. Competition and the structure of bird communities. Monogr. Popul. Biol. n° 7. Princeton University Press, Princeton NJ, 318 p.

Copp G.H. 1989. The habitat diversity and fish reproductive function of floodplain ecosystems. *Environ. Biol. Fish.* **26**, 1-26.

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

Copp G.H., M. Peňáz 1988. Ecology of fish spawning and nursery zones in the flood plain, using a new sampling approach. *Hydrobiologia* **169**, 209-224.

Copp G.H., J.M. Olivier, M. Peňáz, A.L. Roux 1991. Juvenile fishes as functional describers of fluvial ecosystem dynamics: application of the River Rhône, France. *Regul. Riv. Res. Manag.* **6**, 135-145.

Copp G.H., P. Garner 1995. Evaluating the microhabitat use of freshwater fish larvae and juveniles with point abundance sampling by electrofishing. *Folia Zool.* **44**, 145-158.

Faraggi D., R. Simon 1995. A neural network model for survival data. *Stat. Med.* **14**, 73-82.

Fausch K.D., C.L. Hawkes, M.G. Parsons 1988. Models that predict the standing crop of stream fish from habitat variables. U.S. Forest Service General Technical Report PNW-GTR, 213 p.

Fretwell S.D. 1972. Populations in a seasonal environment. Monogr. Popul. Biol. n° 5. Princeton University Press, Princeton N.J. 217 p.

Garner P. 1996*a*. Microhabitat use and diet of 0+ cyprinid fishes in a lentic, regulated reach of the River Great Ouse, England. *J. Fish Biol.* **48**, 367-382.

Garner P. 1996*b*. Suitability indices for juveniles 0+ roach (*Rutilus rutilus* L.) using point abundance sampling data. *Regul. Riv. Res. Manag.* **10**, 99-104.

Garson G.D. 1991. Interpreting neural-network connection weights. *Artif. Intel. Expert* **6**, 47-51.

Goh A.T.C. 1995. Back-propagation neural networks for modelling complex systems. *Artif. Intel. Eng.* **9**, 143-151.

Gorman O.T., J.R. Karr 1978. Habitat structure and streams fish communities. *Ecology* **59**, 507-515.

Grossman G.D., A. de Sostoa 1994. Microhabitat use by fish in the upper Rio Matarrana, Spain, 1984-1987. *Ecol. Freshw. Fish* **3**, 141-152.

Hornik K.K, M. Stinchcombe, H. White 1989. Multilayer feedforward neural networks are universal approximators. *Neural Networks* **2**, 359-366.

James F.C., C.E. McCulloch 1990. Multivariate analysis in ecology and systematics: panacea or Pandora's box? *Ann. Rev. Ecol. Syst.* **21**, 129-166.

Lek S., A. Belaud, P. Baran, I. Dimopoulos, M. Delacoste 1996*a*. Role of some environmental variables in trout abundance models using neural networks. *Aquat. Living Resour.* **9**, 23-29.

Lek S., M. Delacoste, P. Baran, I. Dimopoulos, J. Lauga, S. Aulanier 1996*b*. Application of neural networks to modelling nonlinear relationships in ecology. *Ecol. Model.* **90**, 39-52.

Lerner B., H. Guterman, I. Dinstein, Y. Romen 1994. Feature selection and chromosome classification using a multilayer perceptron neural network. *In:* Proc. International Conference on Neural Networks, *IEEE*, Orlando, 3540-3545.

Mastrorillo S., F. Dauba, A. Belaud 1996. Utilisation des microhabitats par le Vairon, le Goujon et la Loche franche dans trois rivières du sud-ouest de la France. *Ann. Limnol.* **32**, 75-85.

McArthur R.H., H. Reicher, M.L. Cody 1966. On the relation between habitat selection and bird species diversity. *Am. Nat.* **100**, 319-332.

McNally R.C. 1983. On assessing the significance of interspecific competition to guild structure. *Ecology* **64**, 1646-1652.

Nelva A., H. Persat, D. Chessel 1979. Une nouvelle méthode d'étude des peuplements ichtyologiques dans les grands cours d'eau par échantillonnage ponctuel d'abondance. *C. R. Acad. Sci. Paris*, sér. D **209**, 1295-1298.

Norusis, M.J. 1993. SPSS for Windows. Base system user's guide release 6.0. SPSS Inc. 828 p.

Persat H., G.H. Copp 1989. Electrofishing and point abundance sampling for the ichthyology of large rivers. *In:* Developments in electrofishing. I. Cowx eds. Fishing News Books. Blackwell, Oxford, 203-215.

Pouilly M. 1994*a*. Relations entre l'habitat physique et les poissons des zones à cyprinidés rhéophiles dans trois cours d'eau du bassin rhodanien: vers une simulation de la capacité d'accueil pour les peuplements. Thèse dr. Univ. Lyon 1, 256 p.

Pouilly M., Y. Souchon 1994*b*. Simulation de l'habitat physique du barbeau fluviatile (*Barbus barbus* L., 1758): choix des modèles biologiques et sensibilité de la réponse. *Bull. Fr. Pêche Piscic.* **334**, 213-225.

Rumelhart D.E., G.E. Hinton, R.J. Williams 1986. Learning representations by back-propagating error. *Nature* **323**, 533-536.

Schnoener T.W. 1983. Field experiments on interspecific competition. *Am. Nat.* **122**, 240-285.

Seginer I., T. Boulard, B.J. Bailey 1994. Neural network models of the greenhouse climate. *J. Agric. Eng. Res.* **59**, 203-216.

Smith M. 1994. Neural networks for statistical modelling. Van Nostrand Reinhold, New York 235 p.

Smits J.R.M., L.W. Breedveld, M.W. J. Derksen, G. Katermam, H.W. Balfoort, J. Snoek, J.W. Hofstraat 1992. Pattern classification with artificial neural networks: classification of algae, based upon flow cytometer data. *Anal. Chim. Acta* **258**, 11-25.

Teer Braak C.J.F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* **67**, 1167-1179.

Tilman D. 1982. Resource competition and community structure. Monogr. Popul. Biol. n° 17. Princeton University Press, Princeton NJ, 296 p.

Tomassone R., E. Lesquoy, C. Millier 1983. La régression, nouveaux regards sur une ancienne méthode statistique. INRA actualités scientifiques et agnomiques, 13, Masson, Paris 177 p.

Weigend A.S., B.A. Huberman, D.E. Rumelhart 1992. Predicting sunspots and exchange rates with connectionist networks. *In:* Nonlinear modelling and forecasting, SFI Studies in the sciences of complexity, Proc. vol. 12. M. Casdagli, S. Eubank eds. Addison-Wesley 395-432.

Weisberg S. 1980. Applied linear regression. John Wiley and Sons, New York, 324 p.