Role of some environmental variables in trout abundance models using neural networks

Sovan Lek [(1)], Alain Belaud [(2)], Philippe Baran [(2)], Ioannis Dimopoulos [(1)] and Marc Delacoste [(2)]

[(1)] *UMR 9964, Équipe de Biologie quantitative, Université Paul Sabatier, 118, route de Narbonne, 31062 Toulouse Cedex, France.*

[(2)] *Laboratoire d'Ingénierie Agronomique, Équipe Environnement Aquatique et Aquaculture, École Nationale Supérieure Agronomique de Toulouse, 145, avenue de Muret, 31076 Toulouse Cedex, France.*

Received June 26, 1995; accepted September 14, 1995.

Lek S., A. Belaud, P. Baran, I. Dimopoulos, M. Delacoste. *Aquat. Living Resour.*, 1996, **9**, 23-29.

Abstract

Neural networks provide a "black box" model for explaining and predicting trout abundance with 8 environmental variables. This work investigates the specific effect of each variable, by inputting fictitious configurations of explanatory variables and by checking the responses of the model. The comparison between this response of the model to environmental variables on one hand, and results from field observations on the other hand, shows similarities and indicates neural network modelling can be trusted. The elevation appears to be the major explanatory factor. The influence of shelters, bottom velocity and Froude number also play an important role. When considered separately, depth does not have a notable influence on the density of trout. Such confirmations of field observations suggest that these models can be used to obtain a clear identification and hierarchization of the factors influencing the abundance of trout and the mode of action of the factors. This approach can be extended to other applications in quantitative ecology in which non-linear relationships are usually observed.

Keywords: Rivers, habitat, environmental management, mathematical models, *Salmo trutta*, Pyrénées.

Rôle de quelques variables liées à l'habitat sur la densité des truites : modèle du réseau de neurones.

Résumé

Les réseaux neuronaux ont permis de disposer d'un modèle de type « boîte noire » explicatif et prédictif de l'abondance des truites selon 8 variables environnementales. Ce travail porte sur la recherche de l'effet propre de chacune des variables, en appliquant des configurations fictives des variables explicatives et en observant les réponses du modèle. La confrontation entre, d'une part, cette sensibilité du modèle aux variables environnementales et, d'autre part, les résultats directement issus de l'expérience de terrain, montre de nombreuses similitudes et accrédite la pertinence de la modélisation par les réseaux de neurones. L'altitude paraît le facteur explicatif prépondérant. Apparaissent aussi une influence de la surface occupée par les abris, de la vitesse au fond et du nombre de Froude. La profondeur considérée isolément ne manifeste pas d'influence notable sur la densité de truite. Ces confirmations des observations de terrain suggèrent d'utiliser ces modèles pour identifier et hiérarchiser les facteurs influant sur l'abondance des truites et leur mode d'action. Après validation par un retour au terrain, cette démarche pourrait être étendue à d'autres applications en écologie ou des relations de types non linéaires sont fréquentes.

Mots-clés : Rivières, habitat, aménagement environnemental, modèles mathématiques, *Salmo trutta*, Pyrénées.

INTRODUCTION

Biodiversity, density or biomass of a population are the result of a large number of variables. The interactions between these explanatory variables are difficult to formulate with conventional statistical tools.

To take a relatively simple example, the biomass and the density of the trout in a river, essentially depend on the quality of the habitat (Haury *et al.*, 1991). Habitat variables are commonly measured on different scales [microhabitat, mesohabitat, reach (Frissel *et al.*, 1986)] to estimate the carrying capacity of that environment. The carrying capacity can be assimilated to the density of fish observed in the absence of any particular disturbance, whether human or natural (Platts and Nelson, 1988).

Previous studies (Baran, 1995) showed that none of the environmental variables considered could, alone, explain the variability of biomass and/or the density of trout and that multivariate models must be used. Numerous models have been developed (Binns and Eiserman, 1979; Fausch *et al.*, 1988). Faced with this problem, and considering the urgent need for tools useful for living resource management, simulation techniques based on neural networks, recently applied in ecology, have been tested (Baran *et al.*, 1995).

Neural networks adjust the result of the density calculations to the values actually measured in the field by means of iterations and retropropagation of the error (Baran *et al.*, 1995). The results were considered satisfactory since, firstly, the model residuals were independent of the explained variables, and secondly, the model built up from a set of "training data" could predict trout abundance for another set of data obtained in the same geographic area. In principle, the neural network is the "black box" type model and does not clarify the participation of each of the explanatory variables.

In this paper, we probed the operation of this efficient "black box" by making it plot the response between the density of trout and each influencing variable. Since neural networks operate with non-linear relationships (Rumelhart *et al.*, 1986), they can be compared to those established previously from field experiments.

This article uses a simple method to evaluate the response of the explained variable (trout density) and the explanatory variables (variables of the physical environment), in order to investigate the validity of these response functions and to orientate future research.

MATERIAL AND METHODS

51 sites on 13 different rivers of the central French Pyrenean mountains (*fig.* 1) were subdivided into 232 morphodynamic units (Malavoi, 1989), which are considered as homogeneous physical habitats. The units ranged from 2 to 40 metres in length. *Table* 1 gives the general characteristics of the 13 streams studied. Angling activities were similar on the 13 streams. About 12 000 trouts were sampled.

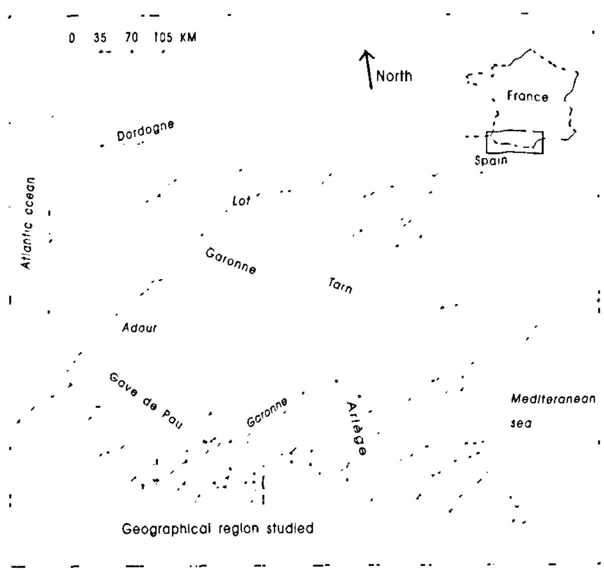

Figure 1. – Geographical location of the 13 streams studied.

Height variables of the physical habitat were measured in each morphodynamic unit at low flow (winters 1992 and 1993). Multiple pass electrofishing was used to estimate winter trout (*Salmo trutta* L.) populations (De Lury, 1951). *Table* 2 summarizes the explained (density of trout) and explanatory (environmental) variables.

Environmental variables

The 8 explanatory variables are listed in *table* 2. Elevation can be considered, over a relatively homogeneous area, to integrate climatic factors (temperature), morphodynamic characteristics (slope, substratum) and the quality of the water (mineralization). The mean values of depth and water velocity (at the surface and at depth where the trout lie) give an indication of the water flow in the mesohabitat. The Froude number of flow (F) is the mean column velocity (v) divided by the square root of the product of the depth (d) and gravitational acceleration (g), *i.e.*:

$$F = \frac{v}{(g \cdot d)^{1/2}} \qquad (1)$$

It is a useful descriptor of the state of the flow (Jowett, 1993). Direct observations of trout behaviour stress the influence of total cover (shelter, deep water area, and overhanging riparian vegetation) on the abundance of trout. Total shelter (undercut bank, logs, wetted brush and boulders) surface area was measured using the method of Binns (1982). Water depth was measured by sounding and water velocity

Table 1. – General characteristics of the 13 streams studied.

Streams	Tributary of	Elevation of source (m)	Elevation of confluence (m)	Mean gradient (%)	Drainage basin	Total length (km)
Gave d'Arrens	Gave de Pau	2 100	420	6.2	Gave de Pau	27
Gave du Marcadau	Gave de Pau	2 330	1 033	9.5	Gave de Pau	13.3
Gave d'Ossoue	Gave d'Héas	2 250	1 190	9.7	Gave de Pau	10.9
Gave d'Héas	Gave de Pau	2 050	1 071	11.2	Gave de Pau	8.7
Ru du Lac de Gaube	Gave du Marcadau	2 330	1 496	10.0	Gave de Pau	8.3
Adour de Lesponne	Adour	1 730	630	5.9	Adour	21.4
Neste d'Oueil	One	1 850	765	11.8	Garonne	9.2
Pique	Garonne	2 440	470	6.5	Garonne	30.5
One	Pique	1 550	600	8.9	Garonne	19.5
Lys	Pique	2 450	850	17.6	Garonne	9.1
Neste du Louron	Neste d'Aure	2 600	700	6.9	Garonne	30.1
Nistos	Neste d'Aure	1 510	430	6.2	Garonne	17.3
Ru d'Estibère	Neste d'Aure	2 360	1 880	12.6	Garonne	3.8

Table 2. – Studied variable table
(x: explanatory variables, y: explained variable).

Variable	Type	Definition	Method used
F	x	Mean Froude number	
D	x	Mean water depth (m)	Direct measurement
Vf	x	Mean bottom velocity ($m.s^{-1}$)	Flowmeter
Vs	x	Mean surface velocity ($m.s^{-1}$)	Flowmeter
Sc	x	Area of shelter (%)	Direct measurement
Cov	x	Area of total cover (%)	
Pool	x	Area of deep water (%)	
Alt	x	Altitude (elevation)	Direct measurement
Dha	y	Density of trout ($fish.ha^{-1}$)	Electrofishing

by an electromagnetic current meter (Marsh McBirney 201D) accurate to within ± 1 $cm.s^{-1}$.

Neural network

Based on the works of Rumelhart *et al.* (1986), Galant (1993) and Smith (1994), the design of the neural network was described in detail in Baran *et al.* (1995). The configuration chosen was that which gave the best adjustment ($R^2=0.879$): 8 neurons in the input layer (coding the 8 variables of the environment), 8 neurons in the hidden layer and one neuron in the output layer coding the density of the trout.

Response of the model to each variable

The complexity of functions implemented by neural networks make the analytical studies of each variable contribution difficult. An experimental approach can be used to determine the response of the model to each of the input variables separately by applying a typical range of variation of a single "free" variable to the model, while the other ("blocked" variables) are held constant. The contribution of each environmental variable to trout density was calculated at 12 values evenly spaced over the range between the minimum and the maximum that appeared in the set of data. The remaining "blocked" variables were provisionally set at an arbitrary level. As this level influenced the results, we set the 7 variables simultaneously together at their minimum value, first quartile, median, third quartile and maximum successively. Five responses were thus obtained for each of the 12 values of the "free" variable. They were further reduced to their median value. The median was chosen over the arithmetic mean because of it robustness. The operation was repeated for all 8 of the environmental variables [1].

RESULTS

Before investigating the responses, it was necessary to assess the utility of each of the 8 measured variables. *Table* 3 indicates the R^2 values for all combinations of 8 and 7 variables. These simulations were performed with a constant structure network with 8 neurons in the hidden layer and using 1 000 iterations. Removal of any variable causes the performance of the network to decrease. Thus, all the variables play a role in the construction of the model.

Table 3. – Performance of models (R^2) for all combinations of 8 and 7 variables at the input of the network.

Variable removed	R^2
None	0.879
F	0.841
D	0.847
Vf	0.855
Vs	0.848
Sc	0.825
Cov	0.864
Pool	0.861
Alt	0.777

[1] The simulating program in Matlab environment is available from S. Lek at let@citct.fr.

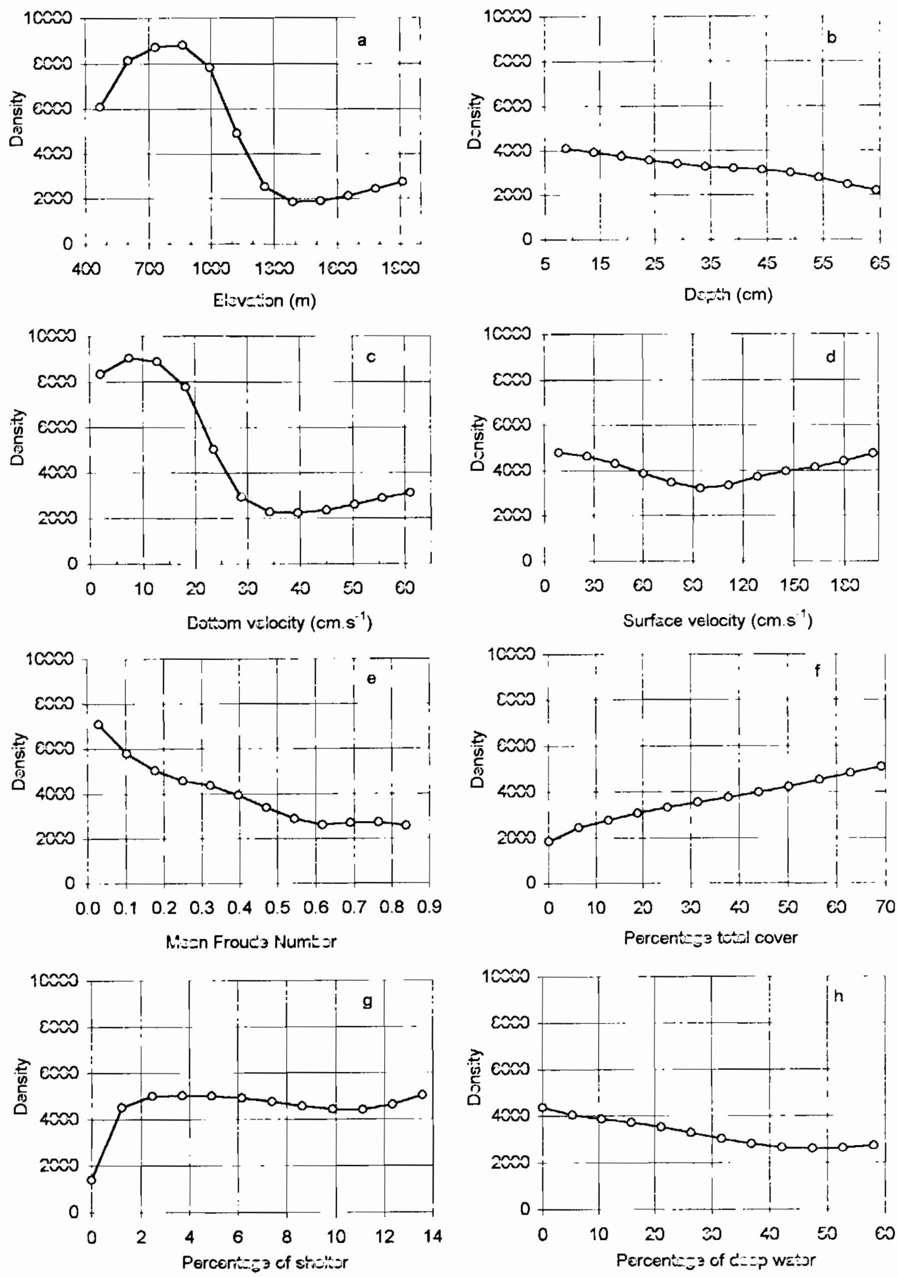

Figure 2. – Contribution profile (response) of each simulated variable to the prediction of trout density by neural networks.

Figure 2 shows the response of the trout density model to each of the 8 physical variables applied to the neural network. The 12 calculated points cover the observed range of variation of the corresponding variable: so the scales on the abscissa differ according to the variable. Several patterns of density response are evident. Predicted trout density increased with the percentage of total cover (*fig.* 2*f*) and decreased with mean Froude number (*fig.* 2*e*), depth and percentage of deep-water area (*figs.* 2*b* and *h*). The contribution of the percentage of shelter to fish density (*fig.* 2*g*) increased rapidly at first, but reached an asymptote for values greater than 2%. The contributions of elevation (*fig.* 2*a*) and bottom velocity (*fig.* 2*c*) were the dominant variables both in magnitude and complexity (see discussion). The response of trout density to surface velocity variation was biphasic and symmetric: inhibition of density at moderate velocities (about 90 cm.s^{-1}) and enhancement at higher and lower velocities.

DISCUSSION

Owing to the complexity of the algorithms implemented by neural networks, analytical study of the contribution of various explanatory variables on the explained variable is extremely difficult. Series of specific configurations of the input variables (within the limit of the observed values) bring out the role of each of them. The most obvious solution was to scan the range of variation of one of the variables while the others remained locked. This attempt to exhibit the influence of one variable using a model designed around maximum interdependence of the effects of all the variables must be considered as preliminary and requires criticism and validation each time a model is developed. The main question is whether all combinations of the fictitious situations applied were valid or whether situations that were not physically encountered must be rejected because they were not taken into account during the training phase of the model. For the moment we can only judge the applicability of this method by the relevance of the responses evaluated with respect to the experimental knowledge. Moreover, it is not possible to interpret the results and the derived responses too literally as they are a result of training and as such must reflect a certain amount of sampling error.

Numerous variables have been used to model the abundance of trout in streams (Fausch *et al.*, 1988; Baran, 1995). The 8 variables used here describe the quality of the physical habitat (Heggenes, 1988) and the elevation (Burton and Wesche, 1974; Scarnecchia and Bergersen, 1987; Haury *et al.*, 1991; Baran *et al.*, 1993) which are known to be important for the trout. This knowledge enables us to evaluate model response but is probably not sufficient to establish a definitive model for the abundance of trout using the characteristics of the environment.

Present understanding of the effects of the environmental variables on the abundance of trout results from direct observations and statistical analyses that were either obtained with univariate analysis (Lewis, 1969; Kennedy and Strange, 1982; Nielsen, 1986; Heggenes, 1988) or with multivariate analysis (Binns and Eiserman, 1969; Fausch *et al.*, 1988; Jowett, 1992).

Elevation is recognized as being a good integrator of the chemical and especially the thermal conditions (Baran *et al.*, 1993; Burton and Wesche, 1974; Scarnecchia and Bergersen, 1987). Most of the observations report a decrease in the abundance of trout with altitudes. For the Pyrenees (Baran, 1995), it appears to be the elevation range 800 to 1 000 m that offers summer temperatures of 12 to 14°C, considered as optimum for the brown trout according to the optimum temperature for the development of this species (Frost and Brown, 1967). Winter conditions, especially temperature, also play an important role in trout abundance (Calkins, 1989). In the Pyrenean region, winter and summer temperatures are strongly related. Consequently, the relationship between trout density and temperature may be explained by either winter or summer conditions.

According to the neural network model, the density of trout is very sensitive to elevation (*fig.* 1*a* and *table* 3), but in a way that is more complex than predicted. The calculated densities are maximum for the interval 600-1 000 m. The reduction in the density of trout downstream (below 500 m) is thought to indicate entry into the valley bottom zone where trout and cyprinids exist together. Between 1 000 and 1 300 m, the simulated abundance of trout decreased strongly, respecting the conventionally accepted climatic effect. The abruptness of this decrease could be caused by obstacles to the passage fish through gorges which often occupy, in the Central Pyrenees, the 900 to 1 300 m band. For the sites located at over 1 300 m and up to 1 900 m, the variations in elevation appear to have no particular effect on the densities calculated by the model (*fig.* 2*h*). Paradoxically, increasingly harsh climatic conditions no longer seem to play a key role. It appears that the model constructed by the neural networks develops a response to elevation that is more complex than the simple climatic gradient used until present both in the Pyrénées (Baran *et al.*, 1993) and in other studies (Burton and Wesche, 1974; Scarnecchia and Bergersen, 1987). This therefore is an example of curve fitting to data on a grand scale.

For physical habitat, percentage area of shelter and bottom water velocity play a substantial role in the model (*fig.* 2). A constant positive sensitivity of the trout density to the increase of the area of total shelter confirms the capital role of this variable in the spatial distribution of the trout (Lewis, 1969; Hunt, 1976; Fausch and White, 1981; Baran *et al.*, 1993). For the percentages of shelter under the banks, an increasing, asymptotic relationship is observed indicating that this variable has a greater influence on trout density for values lower than 2% than for higher values. Baran (1995) also stated, from classical statistical analysis, that trout density increases up to 2% of shelter and remains stable for higher percentages.

The water velocity on the stream bottom (the level where the trout are located) is the second key factor in the prediction of the density of trout populations. This variable is known in trout ecology as being determining in the choice of the microhabitat used for feeding (Shirvell and Dungey, 1983; Bachman, 1984; Heggenes, 1988). The model expresses a variation of the densities similar to a curve of habitat preference determined for individual trout (Bovee, 1978; Belaud *et al.*, 1989; Rincon and Lobon-Cervia, 1993). The model seems to take account of individual preferences of trout on the microhabitat scale.

The surface water velocity had little effect on the result of the density calculation (*fig.* 2*d*). This variable does not characterize the microhabitat chosen by the trout. Baran (1995) has shown that the surface velocity is a good criterion for mesohabitat classification and also that the trout density in a north Pyrenean valley is related to the type of mesohabitat. The different results from the neural network model suggest that the surface velocity is not an appropriate factor to depict the levels of occupation by trout of the different types of mesohabitat. The model seems to give a preference to bottom velocity, related to the microhabitat of the trout.

The average depth is, over the range studied, that which elicits the least response from the model. This is, however, one of the main factors for microhabitat preference (Bovee, 1978; Belaud *et al.*, 1989). The studies of Baran (1995), carried out on the morphodynamic scale of the mesohabitat, and which aimed to identify and to analyse, using conventional statistical methods (multivariate analysis), the factors influencing trout abundance, showed no relationship between the mean depth of the mesohabitat and the density or biomass of the trout.

The Froude number is a general indicator of the flow conditions at the mesohabitat scale (Jowett, 1993). Unlike surface velocity, the Froude number does influence trout density. Abundance decreased as F increased. This index, which combines mean velocity and depth, is in general agreement with the behaviour of the fish on the mesohabitat scale. That is, brown trout like the association of low velocities and deep water.

Automatic modelling by neural networks therefore leads to a system of prediction of the density of trout that seems to be in agreement with current knowledge of their biology and ecology and that seems to combine the micro and mesohabitat characteristics. The model confirms that the individual preferences for habitat greatly influence the spatial distribution of the trout population. This approach can therefore help to identify environmental factors affecting trout abundance.

Acknowledgements

The authors gratefully appreciate the anonymous reviewers for their valuable comments and suggestions resulting in the present form of the paper.

REFERENCES

Bachman R. A. 1984. Foraging behaviour of free-ranging wild and hatchery brown trout in a stream. *Trans. Am. Fish. Soc.* **113**, 1-32.

Baran P., M., Delacoste, J. M. Lascaux, A. Belaud 1993. Relations entre les caractéristiques de l'habitat et les populations de truites communes (*Salmo trutta* L.) de la vallée de la Neste d'Aure. *Bull. Fr. Pêche Piscic.* **331**, 321-340.

Baran P. 1995. Analyse de la variabilité des abondances de truite commune *Salmo trutta* L. dans les Pyrénées centrales françaises. Thèse dr. Institut National Polytechnique, École Nationale Supérieure Agronomique de Toulouse, 156 p.

Baran P., S. Lek, M. Delacoste, A. Belaud 1995. Stochastic models that predict trout population densities or biomass on macrohabitat scale. *Hydrobiologia*, in press.

Belaud A., P. Cheveroche, P. Lim, C. Sabaton 1989. Probability-of-use curves applied to brown trout (*Salmo trutta* L.) in rivers of Southern France. *Regul. Rivers Res. Manag.* **3**, 321-336.

Binns N. A. 1982. Habitat quality index: procedure manual. Wyoming Game and Fish Department, 209 p.

Binns N. A., F. M. Eiserman 1969. Quantification of fluvial trout habitat in Wyoming. *Trans. Am. Fish. Soc.* **108**, 215-228.

Bovee K. D. 1978. Probability-of-use criteria for the family salmonidae. Report U.S. Fish and Wildlife Service, FWS/OBS-78/07, Fort Collins, Colorado, USA, 80 p.

Burton R. A., T. A. Wesche 1974. Relationship of duration of flows and selected watershed parameters to the standing crop estimates of trout populations. University of Wyoming, Water Resources Research Institute, Water Resource Series 52. Laramie.

Calkins D. J. 1989. Winter habitats of Atlantic Salmon, hook trout, brown trout and rainbow trout. A literature review. Special report 89-34. US Forest Service, Dep. of Agriculture, 9 p.

De Lury D. B. 1951. On the planning of experiments for the estimation of fish populations. *J. Fish. Res. Board Can.* **18**, 281-307.

Fausch K. D., R. J. White 1981. Competition between brook trout (*Salvelinus fontinalis*) and brown trout (*Salmo trutta*) in a Michigan stream. *Can. J. Fish. Aquat. Sci.* **38**, 1220-1227.

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, 52 p.

Frost W. E., M. E. Brown 1967. The trout. Collins, St James Place, London, 286 p.

Gallant S. I. 1993. Neural network learning and expert systems. MIT Press, London, 365 p.

Haury J., D. Ombredane, J. L. Baglinière 1991. L'habitat de la truite commune (*Salmo trutta* L.) en eaux courantes.

La truite, biologie et écologie. In: J. L. Baglinière, G. Maisse, INRA Publ., Paris, 25-46.

Heggenes J. 1988. Physical habitat selection by brown trout (*Salmo trutta*) in riverine systems. *Nordic. J. Freshw.* 64, 74-90.

Hunt R. L. 1976. A long-term evaluation of trout habitat development and its relation to improving management-related research. *Trans. Am. Fish. Soc.* **105**, 361-364.

Jowett I .G. 1992. Models of abundance of large brown trout in New Zealand rivers. *N. Am. J. Fish. Manag.* **12**, 417-432.

Jowett I. G. 1993. A method for objectively identifying pool, run, and riffle habitats from physical measurements. *N. Z. J. Mar. Freshw. Res.* **27**, 241-248.

Kennedy G. J. A., C. D. Strange 1982. The dispersion of salmonids in upland streams in relation to depth and gradient. *J. Fish. Biol.* 20, 579-591.

Lewis S. L. 1969. Physical factors influencing fish populations in pools of a trout stream. *Trans. Am. Fish. Soc.* **115**, 183-195.

Malavoi J. R. 1989. Typologie des faciès d'écoulement ou unités morpho-dynamiques d'un cours d'eau à haute énergie. *Bull. Fr. Pêche Piscic.* **315**, 189-210.

Nielsen G. 1986. Dispersion of brown trout (*Salmo trutta* L.) in relation to stream cover and water depth. *Pol. Arch. Hydrobiol.* **33**, 475-488.

Platts W. S., R. L. Nelson 1988. Fluctuations in trout populations for land-use evaluation. *N. Am. J. Fish. Manag.* 8, 333-345.

Rincon P. A., J. Lobon-Cervia 1993. Microhabitat use by stream-resident brown trout: bioenergetic consequences. *Trans. Am. Fish. Soc.* **122**, 575-587.

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

Scarnecchia D. L., E. P. Bergersen 1987. Trout production and standing crop in Colorado's small streams, as related to environmental features. *N. Am. J. Fish. Manag.* **7**, 315-330.

Shirvell C. S., R. G. Dungey 1983. Microhabitat chosen by brown trout for feeding and spawning in rivers. *Trans. Am. Fish. Soc.* **112**, 355-367.

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