

Review and comparison of methods to study the contribution of variables in artificial neural network models

Muriel Gevrey^{a,*}, Ioannis Dimopoulos^b, Sovan Lek^a

^a CESAC UMR 5576, CNRS-University Paul Sabatier, 118, route de Narbonne, 31062 Toulouse cedex, France

^b Department of Health and Welfare Unit Administration, Technological Educational Institute of Kalamata, Antikalamos, 24100 Kalamata, Greece

Abstract

Convinced by the predictive quality of artificial neural network (ANN) models in ecology, we have turned our interests to their explanatory capacities. Seven methods which can give the relative contribution and/or the contribution profile of the input factors were compared: (i) the ‘PaD’ (for Partial Derivatives) method consists in a calculation of the partial derivatives of the output according to the input variables; (ii) the ‘Weights’ method is a computation using the connection weights; (iii) the ‘Perturb’ method corresponds to a perturbation of the input variables; (iv) the ‘Profile’ method is a successive variation of one input variable while the others are kept constant at a fixed value; (v) the ‘classical stepwise’ method is an observation of the change in the error value when an adding (forward) or an elimination (backward) step of the input variables is operated; (vi) ‘Improved stepwise a’ uses the same principle as the classical stepwise, but the elimination of the input occurs when the network is trained, the connection weights corresponding to the input variable studied is also eliminated; (vii) ‘Improved stepwise b’ involves the network being trained and fixed step by step, one input variable at its mean value to note the consequences on the error. The data tested in this study concerns the prediction of the density of brown trout spawning redds using habitat characteristics. The PaD method was found to be the most useful as it gave the most complete results, followed by the Profile method that gave the contribution profile of the input variables. The Perturb method allowed a good classification of the input parameters as well as the Weights method that has been simplified but these two methods lack stability. Next came the two improved stepwise methods (a and b) that both gave exactly the same result but the contributions were not sufficiently expressed. Finally, the classical stepwise methods gave the poorest results.

© 2002 Published by Elsevier Science B.V.

Keywords: Artificial neural networks; Backpropagation; Non-linear relationships; Variables contribution; Sensitivity analysis; Perturbation; Partial derivatives; Trout; Habitat modelling

1. Introduction

The relationships between variables in ecology are almost always very complicated and highly non-linear. One of the most appropriate methods to illustrate this seems to be Artificial Neural

* Corresponding author. Tel.: +33-561-55-8687; fax: +33-561-55-6096.

E-mail address: gevrey@cict.fr (M. Gevrey).

Networks (ANNs). In fact, this method is very powerful in dealing with non-linear relationships (Lek et al., 1996b). A large number of authors have underlined the interest of using ANNs instead of linear statistical models (Paruelo and Tomasel, 1997; Ramos-Nino et al., 1997; Manel et al., 1999; Starrett and Adams, 1997; Özesmi and Özesmi, 1999). This method has become increasingly popular in the analysis of ecological phenomena (Skelton et al., 1995; Recknagel et al., 1997; Whitehead et al., 1997; Lu et al., 1998; Mastrorillo et al., 1998; Yang et al., 1998; Brosse et al., 1999; Lae et al., 1999; Lek et al., 1999; Liong et al., 2000b; Maier and Dandy, 2000). Convinced by the predictive power of ANNs and their ability to analyse non-linear relationships, we consider them interesting to study from their explanatory point of view. In fact, starting from input variables, ANNs have the capacity to predict the output variable but the mechanisms that occur within the network are often ignored. ANNs are often considered as black boxes. Various authors have explored this problem and proposed algorithms to illustrate the role of variables in ANN models.

Nevertheless, in most works, these methods are used to eliminate irrelevant input, and are, therefore, called pruning methods (Guo and Uhrig, 1992; Zurada et al., 1994; El-Keib and Ma, 1995; Engelbrecht et al., 1995; Hsu et al., 1995; Maier et al., 1998; Yao et al., 1998; van Wijk and Bouten, 1999; Kim et al., 2000; Liong et al., 2000a,b). First, the most significant explanatory variables are determined, then the variables which are below a fixed threshold are excluded from the network. This allows the size of the network to be reduced and thus minimises redundancy in the training data (Zurada et al., 1994). However, even if good prediction is required in ecology, knowing what contribution each variable makes is of a prime importance. It is this explanatory aspect of ANNs that we studied here. These methods were used to determine the influence of each input variable and its contribution to the output. They are not, therefore, pruning methods but procedures to estimate the relative contribution of each input variable.

Seven different methods that allow contribution analysis were used: (i) the ‘PaD’ (for Partial Derivatives) method consists in calculating the partial derivatives of the output according to the input variables (Dimopoulos et al., 1995, 1999); (ii) the ‘Weights’ method is a computation using the connection weights (Garson, 1991; Goh, 1995); (iii) the ‘Perturb’ method corresponds to a perturbation of the input variables (Scardi and Harding, 1999); (iv) the ‘Profile’ method is a successive variation of one input variable while the others are kept constant at a fixed value (Lek et al., 1996a,b); (v) the ‘classical stepwise’ method is an observation of the change in the error value when an adding (forward) or an elimination (backward) step of the input variables is operated (Balls et al., 1996; Maier and Dandy, 1996); (vi) ‘Improved stepwise a’ uses the same principle as the last one, but the elimination of the input occurs when the network is trained, the connection weights corresponding to the input variable studied are also eliminated; (vii) ‘Improved stepwise b’ also involves the network being trained and fixed step by step with one input variable at its mean value to note the consequences on the error.

Multiple linear regression (MLR) will be used as classical model to judge the prediction quality of ANNs. In addition the capacities of stepwise regression will be compared with those of the contribution procedures associated to ANNs.

In the present paper, Section 2 describes the ecological database used in our study. Section 3.1 presents the regression model used. Section 3.2 presents the neural model. Section 3.3 introduces the methods that allow the determination of the variable contributions. Section 4 presents the results obtained for all methods used. Finally, Section 5 discusses the contribution of the input through a comparison of the methods and some conclusions are drawn.

2. Database

The data used here were reported in Delacoste et al. (1993), Delacoste (1995) and Lek et al. (1996b). Sampling was done at 29 stations, distributed on six rivers, subdivided into 205 morphodynamic

units. Each unit corresponds to a zone where depth, current and gradient are homogeneous (Malavoi, 1989). The physical characteristics of the 205 morphodynamic units were measured in January, immediately after the brown trout reproduction period. They, therefore, most faithfully indicate the conditions met by the trout during its reproduction. With reference to the works of Ottaway et al. (1981), Shirvell and Dungey (1983) and Crisp and Carling (1989), ten physical habitat variables were measured (Table 1) (Delacoste, 1995; Lek et al., 1996a,b).

3. Methods

3.1. Multiple linear regression modelling

MLR being the method most frequently used in ecology, a comparison to ANNs was made in order to judge their predictive capacities. The stepwise multiple regression technique (Weisberg, 1980; Tomassone et al., 1983) was computed especially to define the significant variables and their contribution order. In fact the influence of each variable can be roughly assessed by checking the final values of the regression coefficients.

Table 1
Habitat variables measured to study brown trout reproduction (from Delacoste et al., 1993)

Variable	Type	Characteristics
Wi	i	Wetted width (m ²)
ASSG	i	Area with suitable spawning gravel for trout per linear meter of river (m ² /linear m)
SV	i	Surface velocity (m/s)
GRA	i	Water gradient (%)
Fwi	i	Flow/width (m ³ /s per m)
D	i	Mean depth (m)
SDD	i	Standard deviation of depth (m)
BV	i	Bottom velocity (m/s)
SDBV	i	Standard deviation of bottom velocity (m/s)
VD	i	Mean speed/mean depth (m/s per m)
R/M	d	Density of trout redds per linear meter of streambed (redds/m)

i, independent, d, dependent; the independent variables are non-correlated except SV and BV, R = 0.76.

Calculations were done using s PLUS[®] software release 4.5 on PC.

3.2. Neural network modelling

The multi-layer feed-forward network, that is the most popular of the much architectures currently available, was used. The network was trained using an error backpropagation training algorithm (Rumelhart et al., 1986). This algorithm adjusts the connection weights according to the backpropagated error computed between the observed and the estimated results. This is a supervised learning procedure that attempts to minimise the error between the desired and the predicted outputs.

The network used consisted of three layers: one input layer of ten neurons (one for each input variable), one hidden layer of five neurons (it is the number which gives the best prediction result) and one output layer of one neuron which is the output variable (Fig. 1).

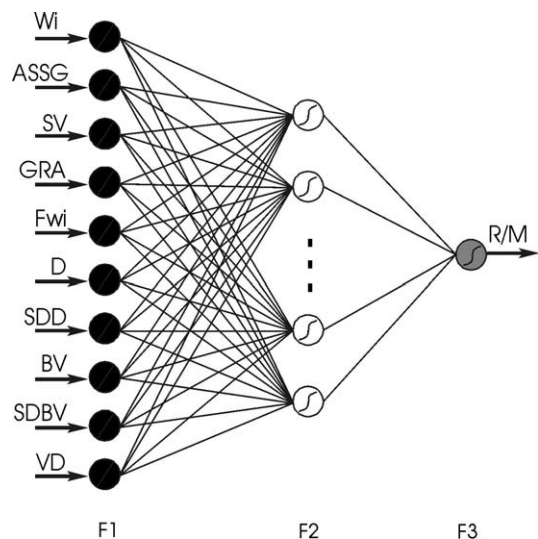


Fig. 1. Structure of the neural network used in this study. F1, input layer of neurons comprising as many neurons as variables at the entry of the system; F2, hidden layer of neurons whose number is determined empirically; F3, output layer of neurons with a single neuron corresponding to the single dependent variable.

Modelling was carried out in two steps:

- Firstly, testing the model to calibrate the model parameters: random selection was used to isolate a training set (3/4 of the records, i.e. 154) and an independent test set (1/4 of the records, i.e. 51). The model was first adjusted with the training set and then tested with the test set to determine the best ANN configuration (Geman et al., 1992).
- Secondly, applying the methods used to study the contribution of the different variables at the input on the already calibrated ANN model (during the first step) by using the whole data set.

3.3. Methods for testing the contributions of the different variables

3.3.1. 'PaD' method

Two results can be obtained by this method. The first is a profile of the output variations for small changes of each input variable and the second is a classification of the relative contributions of each variable to the network output.

To obtain the profile of the variations of the output for small changes of one input variable, we compute the partial derivatives of the ANN output with respect to the input (Dimopoulos et al., 1995, 1999). For a network with n_i inputs, one hidden layer with n_h neurones, and one output (i.e. $n_o = 1$), the partial derivatives of the output y_j with respect to input x_j (with $j = 1, \dots, N$ and N the total number of observations) are:

$$d_{ji} = S_j \sum_{h=1}^{n_h} w_{ho} I_{hj} (1 - I_{hj}) w_{ih}$$

(on the assumption that a logistic sigmoid function is used for the activation). When S_j is the derivative of the output neuron with respect to its input, I_{hj} is the response of the h th hidden neuron, w_{ho} and w_{ih} are the weights between the output neuron and h th hidden neuron, and between the i th input neuron and the h th hidden neuron.

A set of graphs of the partial derivatives versus each corresponding input variable can then be

plotted, and enable direct access to the influence of the input variable on the output. One example of an interpretation of these graphs is that, if the partial derivative is negative then, for this value of the studied variable, the output variable will tend to decrease while the input variable increases. Inversely, if the partial derivatives are positive, the output variable will tend to increase while the input variable also increases.

The second result of PaD concerns the relative contribution of the ANN output to the data set with respect to an input. It is calculated by a sum of the square partial derivatives obtained per input variable:

$$SSD_i = \sum_{j=1}^N (d_{ji})^2$$

One SSD (Sum of Square Derivatives) value is obtained per input variable. The SSD values allow classification of the variables according to their increasing contribution to the output variable in the model. The input variable that has the highest SSD value is the variable, which influences the output variable most.

3.3.2. 'Perturb' method

This method aims to assess the effect of small changes in each input on the neural network output. The algorithm adjusts the input values of one variable while keeping all the others untouched. The responses of the output variable against each change in the input variable are noted. The input variable whose changes affect the output most is the one that has the most relative influence. In fact, the mean square error (MSE) of the neural network output is expected to increase as a larger amount of noise is added to the selected input variable (Yao et al., 1998; Scardi and Harding, 1999).

These changes can take the form of $x_i = x_i + \delta$ where x_i is the selected input variable and δ is the change. δ can be increased in steps of 10% of the input value up to 50% (commonly used values).

The aim is to assess the effect of small changes in each input on the neural network output. We can then obtain a classification of the input variables by order of importance.

3.3.3. 'Weights' method

The procedure for partitioning the connection weights to determine the relative importance of the various inputs was proposed first by Garson (1991) and repeated by Goh (1995), see Appendix A. The method essentially involves partitioning the hidden-output connection weights of each hidden neuron into components associated with each input neuron. We suggested to simplify this algorithm that gives results identical to the algorithm initially proposed:

(1) For each hidden neuron h , divide the absolute value of the input-hidden layer connection weight by the sum of the absolute value of the input-hidden layer connection weight of all input neurons, i.e.

For $h = 1$ to nh ,

For $i = 1$ to ni ,

$$Q_{ih} = \frac{|W_{ih}|}{\sum_{i=1}^{ni} |W_{ih}|}$$

end,

end.

(2) For each input neuron i , divide the sum of the Q_{ih} for each hidden neuron by the sum for each hidden neuron of the sum for each input neuron of Q_{ih} , multiply by 100. The relative importance of all output weights attributable to the given input variable is then obtained.

For $i: 1$ to ni

$$RI(\%)_i = \frac{\sum_{h=1}^{nh} Q_{ih}}{\sum_{h=1}^{nh} \sum_{i=1}^{ni} Q_{ih}} \times 100$$

end.

3.3.4. 'Profile' method

This method was proposed by Lek (Lek et al., 1995, 1996a,b). The general idea is to study each input variable successively when the others are then blocked at fixed values. The principle of this algorithm is to construct a fictitious matrix pertaining to the range of all input variables.

In greater detail, each variable is divided into a certain number of equal intervals between its minimum and maximum values. The chosen number of intervals is called the scale. All variables except one are set initially, (as many times as required for each scale), at their minimum values, then successively at their first quartile, median, third quartile and maximum. For each variable studied, five values for each of the scale's points are obtained. These five values are reduced to the median value. Then the profile of the output variable can be plotted for the scale's values of the variable considered. The same calculations can then be repeated for each of the other variables. For each variable, a curve is then obtained. This gives a set of profiles of the variation of the dependent variable according to the increase of the input variables (see Fig. 2 with a scale of variation of 12). In this work, a range of the different scales possible was used, so the profiles were plotted for scales of 12, 24, 48, 96, 144 and 192.

3.3.5. Stepwise method

This method is the classical stepwise method that consists of adding or rejecting step by step one input variable and noting the effect on the output result. Based on the changes in MSE, the input variables can be ranked according to their importance in several different ways depending on different arguments. For instance the largest changes in MSE due to input deletions can allow these inputs to be classified by order of significance. In another approach the largest decrease in MSE can identify the most important variables, i.e. the most relevant to the construction of a network with a small MSE (Sung, 1998).

The two stepwise modelling approaches were adopted to assess the effect of the ten input variables used: first, the one by one addition of the input variables (forward stepwise), and second,

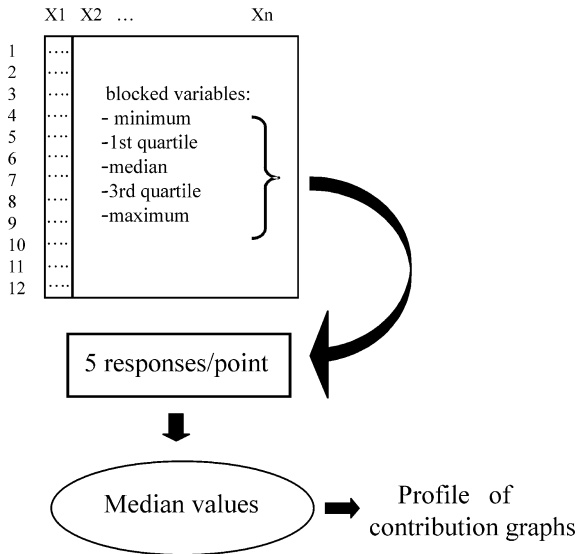


Fig. 2. Explanatory schema of the Profile method. When the X_1 variable is distributed on 12 variation levels between its minimum and its maximum value of initial data, the other variables are maintained fixed, successively at the minimum, first quartile, median, third quartile and maximum. At each X_1 value five responses are obtained, and it is the median value which is taken into account.

the elimination of the input variables (backward stepwise).

- 1) Forward stepwise: ten models were generated, each using only one of the available variables input. Then, nine models were generated, combining the variable that resulted in the smallest error (for a single input variable) with each of the remaining variables, this procedure was repeated using models with three input variables, four... until all the variables were added (Maier et al., 1998). The order of integration of the input variables in the network is the order of the importance of their contributions.
- 2) Backward stepwise: ten models were generated, each using only nine of the available variables each as inputs. The tenth missed out variable for which the resulting models gave the largest error, is the most important. Then, nine models were generated, combining eight

variables, i.e. the variables minus that eliminated just before and one of the other available inputs was eliminated in each model. This procedure was repeated using models with seven input variables, six... until the nine variables were all eliminated. The order of elimination of the input variables in the network is the order of the importance of their contribution.

3.3.6. 'Improved stepwise' methods

The major drawback of the classical stepwise method is that at each step a new model is generated and requires training. An improvement of this method consisted of building two others called improved stepwise method a and b, where only one model is used.

In methods that use a single trained model, each variable in turn is processed and the MSE examined. The variable that gives the largest MSE when eliminated is the most important. A classification of the variables can thus be made. (i) The Improved Stepwise a method consists of eliminating one variable and its corresponding weights. (ii) For the Improved Stepwise b method, all the values of one input are transformed to the same value, i.e. its mean.

3.3.7. Model stability measurement

In order to check the stability of each method, we repeated the training of the network ten times and noted the relative contributions of the input variables on the output obtained for each method and each trained network. We then calculated the mean contribution of each variable for the different methods. The ten training sessions allowed us to draw the standard error (S.E.) which gives an indication of the stability of the method.

4. Results

4.1. Multiple linear regression models

4.1.1. Predictive capacity

Using all ten available variables and the complete dataset, the equation of the MLR model and the determination coefficient were:

<i>R/M</i> =	1.3374	-0.02	Wi	+0.48	ASSG	-0.57	SV	-0.05	GRA	+1.31	Fwi	-0.01	D	-0.08	SDD	+0.01	BV	-0.01	SDBV	-0.02	VD
<i>t</i>		-1.1739		9.1384		-2.1596		-2.1472		1.6213		-1.7189		-1.0074		1.3274		-1.8332		-0.1991	
Sig.			0.2419	0.0000		0.0320		0.0330		0.1066		0.0872		0.3150		0.1859		0.0683		0.8424	
R² =	0.4692																				

Using forward–backward stepwise MLR, five variables were retained by the model: ASSG, SV, D, GRA, Fwi. The equation and the determination coefficient became:

structure can then be used for the second step, using the complete database for sensitivity analysis.

The result of the second step is $R^2 = 0.77$ ($P < 0.01$) testifying the predictive quality of the model.

<i>R/M</i> =	1.22	+0.46	ASSG		-0.65	SV		-0.05	GRA		+1.43	Fwi		-0.02	D
<i>t</i>		9.8774			-3.8614			-2.6433			1.8945			-3.8420	
Sig.		0.0000			0.0002			0.0089			0.0596			0.0002	
R² =	0.4542														

4.1.2. Explanatory capacity

MLR partial coefficients generally give an indication of environmental reality. Each coefficient is the partial derivative of the response of the model with respect to the variable of that coefficient; therefore, the influence of each variable can be assessed by checking the final values of the regression coefficients.

Concerning the complete model, only three variables contribute significantly. They are, in order of importance: ASSG, SV and GRA. The relationship between R/M and ASSG is positive while for SV and GRA the relationship with R/M is negative. The stepwise model retains four significant variables, which are ASSG, SV, D and GRA. Only ASSG has a positive relationship with R/M. So, the stepwise procedure does not lead to a very different conclusion from the complete one.

4.2. Artificial neural network models

4.2.1. Predictive capacity

The results of the first step (see Section 3.2) are $R^2 = 0.75$ ($P < 0.01$) for the learning set, and $R^2 = 0.76$ ($P < 0.01$) for the testing set. The results (determination coefficient) are as good in the learning set as in the testing set. The ANN

4.2.2. Contributions of input variables

Fig. 3 presents the derivative plots of the PaD method:

- a) The partial derivative values of R/M with respect to Wi are all negative: an increase of Wi leads to a decrease of R/M. For high values of Wi, the partial derivative values approach zero, thus R/M tend to become constant.
- b) The partial derivative values of R/M with respect to ASSG are all positive and very high for the low values of ASSG: R/M increases with the increase of ASSG and progressively this increase drops to become null for the highest value of ASSG.
- c) The partial derivative values of R/M with respect to SV are negative for low values of SV and near zero for the higher values. R/M decreases with the increase of SV till it becomes constant at high values of SV.
- d) The partial derivative values of R/M with respect to GRA are negative for low values of GRA and near zero for higher values. R/M decreases with the increase of GRA and progressively becomes constant.
- e) For low values of Fwi, the partial derivatives of R/M with respect to Fwi are positive,

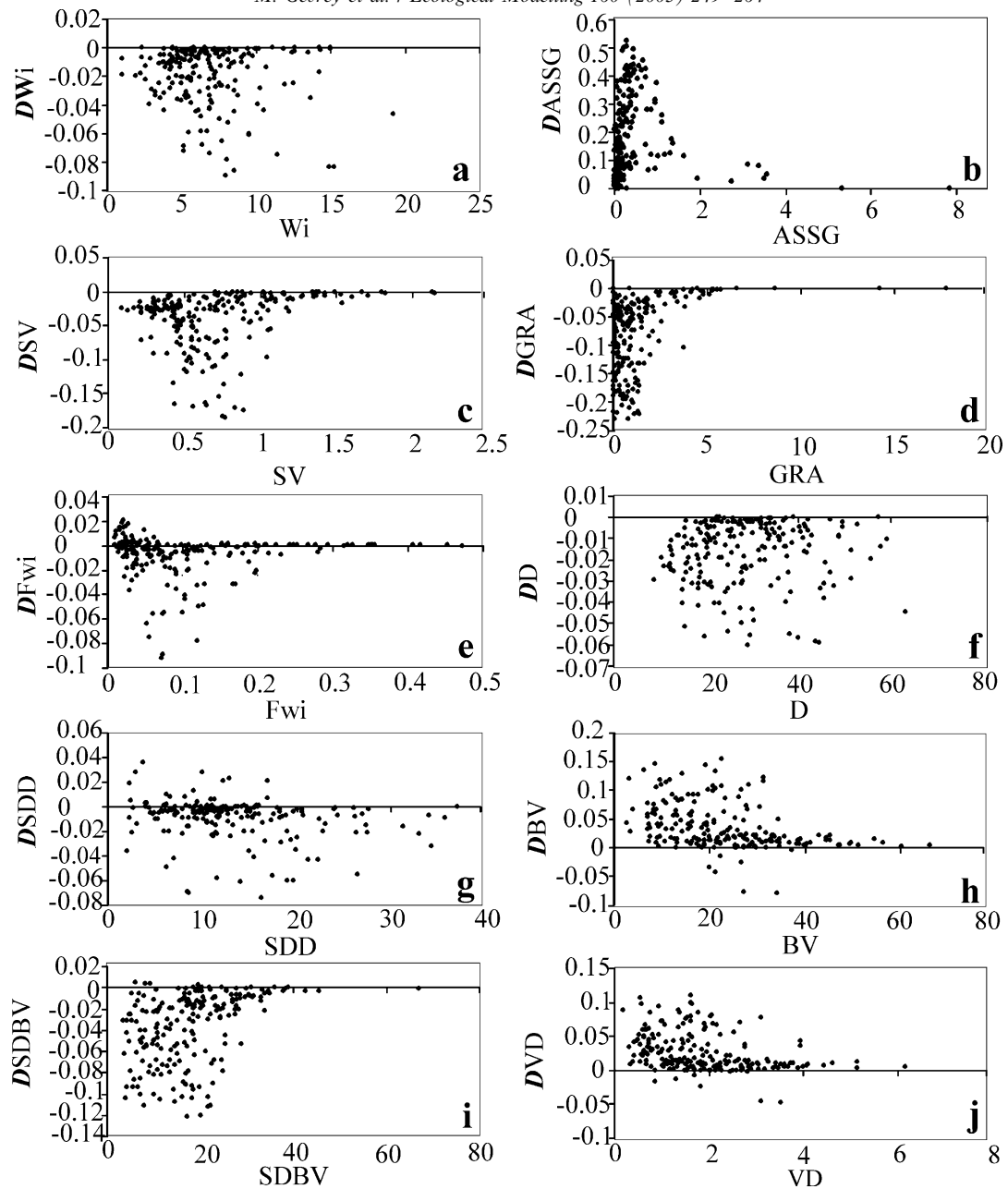


Fig. 3. Partial derivatives of the ANN model response (R/M) with respect to each independent variable (PaD algorithm, Derivatives Profile); (a) Wi; (b) ASSG; (c) SV; (d) GRA; (e) Fwi; (f) D; (g) SDD; (h) BV; (i) SDBV; (j) VD.

become rapidly negative then rise to reach null values for high Fwi: an increase of Fwi leads to a short increase of R/M and then a decrease which becomes attenuated and finally constant for high values of Fwi.

f) All the partial derivative values of R/M with respect to D are negative: an increase of D leads to a decrease of R/M.

g) The partial derivative of R/M with respect to SDD are positive and negative without a

precise direction, it is then not possible to come to a real conclusion about the action of SDD on R/M. It could, for instance, be due to an interaction between SSD and another variable.

- h) The partial derivative values of R/M with respect to BV are all positive: an increase of BV leads to an increase of R/M but to a lesser extent for the high values of BV.
- i) The partial derivative values of R/M with respect to SDBV are all negative: an increase of this variable leads to a decrease of R/M.
- j) The partial derivative values of R/M with respect to VD are almost all positive and near zero for high values of VD, an increase in this variable leads to an increase in R/M and R/M becomes constant for high values of VD.

Fig. 5a presents the relative contributions resulting from the application of the PaD method. The method is then very stable, whatever the model, and has a low confidence interval.

ASSG is the highest contributed variable (> 65%), followed by GRA. However, the contribution of the other variables is very low (Table 2) and the difference between SV, BV and SDBV is not significant, then come VD, Wi and at last D, Fwi and SDD between which the difference is again non-significant (Table 3).

The results of the Profile method are presented in Fig. 4. Graphs a to f represent the Profile method, respectively, for 12, 24, 48, 96, 144 and 192 scale intervals of input variables between their minimum and maximum. Each graph represents a different scale. It is interesting to notice the stability of the method whatever the scale. In fact, the profiles of the different variables always have the same shape, the difference is that the larger the scale is, the more the profile of the variables is marked.

In Fig. 4a, we can see that variables ASSG, GRA and SV are well expressed. ASSG is the variable, which has the greatest effect on the output as seen through the large range, i.e. which is the most important. An increase of ASSG leads to an increase of R/M and progressively, for the highest value of ASSG, the R/M values become

constant. An increase of GRA leads to a decrease of R/M. For low values of SV the values of R/M stay constant and progressively decrease. The same results are better expressed on a larger scale as presented in Fig. 4c. The values of R/M decrease with the increase of D; same result is observed for SDBV. For ASSG, GRA and SV the results are the same as in Fig. 4a with more details.

The relative contributions of each input variable can be expressed by the range values (maximum–minimum) of their contributions (see Table 2).

The contributions of the input variables given by the Weights method are presented in Fig. 5b. Compared with the PaD method, the confidence intervals are larger, testifying the greater instability of Weights method.

ASSG is the variable that makes the largest contribution followed by SV and GRA which both are not significantly different and then BV. The last five variables are not significantly different (Tables 2 and 3).

Fig. 5c shows the results obtained with the Pertub method with 50% perturbation. The error bars are large for the three most important variables, i.e. ASSG, SV and GRA, this method is not very stable.

ASSG is the most important variable following by GRA, SV, SDBV, BV, VD, Wi, D, SDD and finally Fwi (Table 2). There are, however, some differences between variable contributions that are not significant: between SV and SDBV, between BV, VD and Wi and between D and SDD (Table 3).

The results of the stepwise methods, forward and backward are presented in Table 2; the two methods do not give similar results except for the most important variable, which is ASSG in both cases. For instance, the forward stepwise method gives Wi, D, and SV after ASSG while the backward stepwise method gives GRA, SDBV, and SV after ASSG.

The results for improved stepwise methods are exactly the same and are presented in Table 2 and Fig. 5d where the small error bar indicates a good stability of the methods. ASSG is the variable which has the largest contribution followed by BV, and GRA, then come four not significantly different variables: SV, Wi, VD and SDBV, and

Table 2
Results of the sensitivity analysis: classification of the input variables according to the methods employed

Methods/Variables	Forward stepwise	Backward stepwise	PaD	Weights	Perturb	Improved stepwise a	Improved stepwise b	Profile 12	Stepwise MLR
Wi	2	5	7	7	7	5	5	8	–
ASSG	1	1	1	1	1	1	1	1	1
SV	4	4	3	2	3	4	4	3	2
GRA	5	2	2	3	2	3	3	2	4
Fwi	9	8	9	9	10	10	10	7	5
D	3	7	8	10	8	9	9	6	3
SOD	10	9	10	5	9	8	8	5	–
BV	7	6	4	4	5	2	2	9	–
SDBV	6	3	5	8	4	7	7	4	–
VD	8	10	6	6	6	6	6	9	–

–: not used.

Table 3

Semi-confusing matrix of Mann–Whitney pairwin comparison of variable contribution of four methods used: 1, PaD method; 2, Weight method; 3, Perturb method and 4, Improved stepwise (a and b) methods

	Wi	ASSG	SV	GRA	Fwi	D	SDD	BV	SDBV	VD
Wi	–	#	#(1, 2, 3)	#	#(1, 3, 4)	#	#(1, 3, 4)	#	#(1, 3)	#(1)
ASSG	–	–	#	#	#	#	#	#	#	#
SV	–	–	–	#(1, 3, 4)	#	#	#	#(2, 3, 4)	#(1, 2)	#(1, 2)
GRA	–	–	–	–	#	#	#	#	#	#
Fwi	–	–	–	–	–	#(3, 4)	#(3, 4)	#	#(1, 3, 4)	#(1, 3, 4)
D	–	–	–	–	–	–	#(2)	#	#	#
SDD	–	–	–	–	–	–	–	#	#(1, 3, 4)	#(1, 3, 4)
BV	–	–	–	–	–	–	–	–	#(2, 3, 4)	#(1, 2, 4)
SDBV	–	–	–	–	–	–	–	–	–	#(1, 3)
VD	–	–	–	–	–	–	–	–	–	–

#, with no number following means that the contribution of the two variables is significantly different with all the methods, otherwise the number of the methods for which it is different is noted.

finally SDD, Fwi and D which are not significantly different either (Table 3).

5. Discussion

MLR is the most commonly used method to analyse ecological data. It has been thoroughly statistically tested and is universally known. Its success comes from its easy use and its capacity to give predictive and explanatory results which make it very interesting. However, its incapacity to take into account non-linear relationships between the dependent variable and each independent variable is its principal drawback. That is why the use of ANNs is wholly justified in ecology where the relationships between variables are principally non-linear. Nevertheless, it is another drawback of MLR that is raised here: their inability to explain. In fact, determining the influence of the independent variable consists in checking the final values of the partial regression coefficients (Tomassone et al., 1993). However, this can be done only on the variables for which the coefficients are significant. Moreover, MLR gives only a coefficient with a sign and values for each independent variable, which can be translated by a direction of the relationship with the dependent variable, but no more information can be extracted from the results. This problem can then be avoided, by adding explanatory methods to

ANNs, which determine the contributions of the independent variables and the way they act on the dependent variable.

Explanatory methods have been developed with the idea to clarify the ‘black-box’ approach of ANNs. ANN models are able to make perfect predictions and are recognised as powerful in this field (Skelton et al., 1995; Recknagel et al., 1997; Whitehead et al., 1997; Lu et al., 1998; Mastroiello et al., 1998; Yang et al., 1998; Brosse et al., 1999; Lae et al., 1999; Lek et al., 1999; Liong et al., 2000b; Maier and Dandy, 2000). However, the relationships that occur between the variables are often complex in ecology, but also, very interesting to understand. It was then necessary to work on methods like contribution or sensitivity analysis to add power to ANNs in their explanatory capacity (Dimopoulos et al., 1995, 1999; Garson, 1991; Goh, 1995; Scardi and Harding, 1999; Lek et al., 1996a,b; Balls et al., 1996; Maier and Dandy, 1996).

In this work, the prediction results are satisfactory, testifying then a good prediction of trout density which is better with ANNs than with MLR, confirming the non-linearity of the relationship between the variables (Lek et al., 1996b).

From an ecological point of view, ASSG, the area with suitable spawning gravel for a trout redd, is the most significant variable whatever the method (linear or nonlinear). The presence of gravel is in fact a very relevant factor for trout

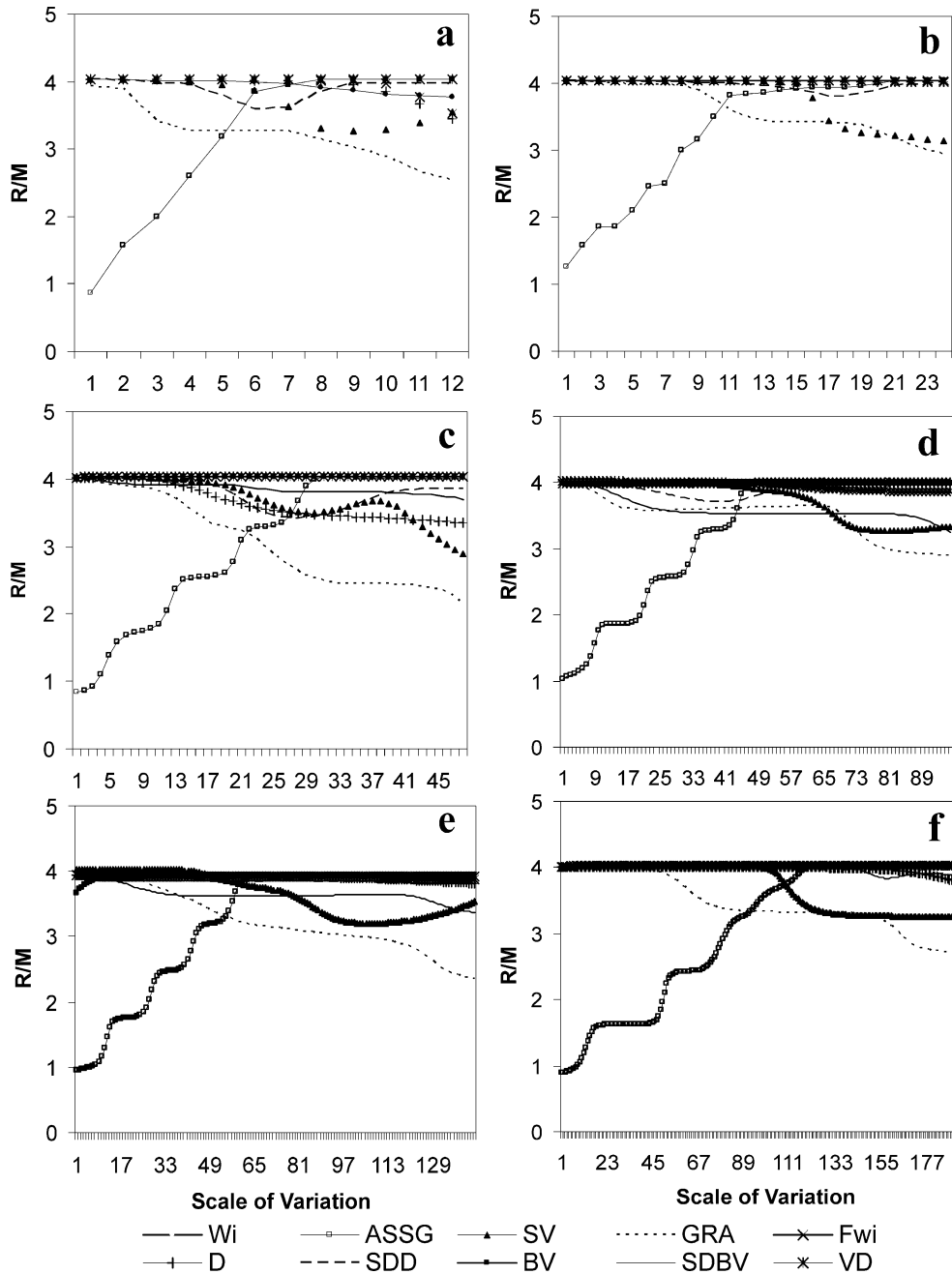


Fig. 4. Contribution of the ten independent variables (Wi, ASSG, SV, GRA, Fwi, D, SDD, BV, BV, SDBV, VD) used in the 10-5-1 ANN model for R/M, by the Profile algorithm. Six different scales were used, (a) 12, (b) 24, (c) 48, (d) 96, (e) 144, and (f) 192.

breeding (Rubin, 1995). Substrates with extreme gravel size disadvantage laying trout. Survival at the hatching stage has been found higher for

moderate values. A good particle size allows good water flow-through, keeping the egg oxygenation optimal. The gradient and water velocity is then

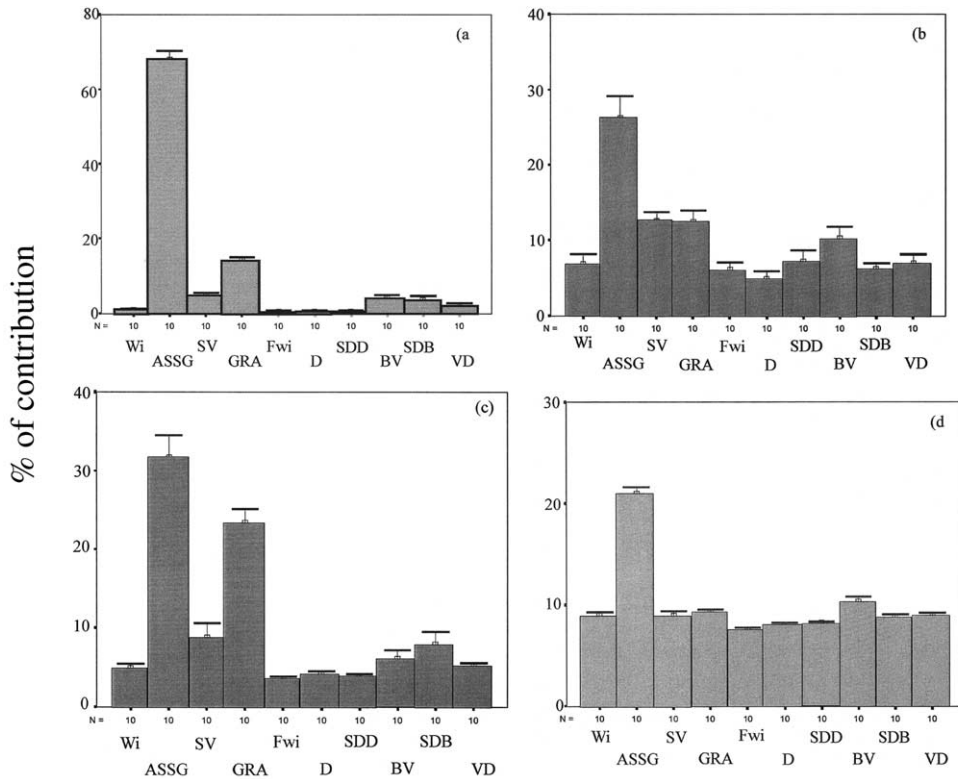


Fig. 5. Contribution of the ten independent variables (Wi, ASSG, SV, GRA, Fwi, D, SDD, BV, BV, SDBV, VD) used in the 10-5-1 ANN model for R/M, (a) in the PaD algorithm, relative contributions; (b) in the Weights algorithm; (c) in the Perturb algorithm for 50% perturbation; (d) in the improved stepwise (a and b) algorithm.

the most important variables. This is also in relation to the trout behaviour during egg laying. The ground has to be flat and the water velocity not too high. The model confirms that the habitat preferences of the trout population influence their spatial distribution.

The explanatory methods thus help to identify environmental factors affecting trout abundance and how these factors contribute to trout abundance. Several methods have been used in this study. Each of them can be criticised. While the PaD and the Profile methods provide two elements of information on the contribution of the environmental variables (order of contribution and mode of action), the other methods studied (Perturb, Weights, Classical stepwise, Improved stepwise)

are just able to classify the variables by order of importance of their contribution to the output. Comparing the PaD and Profile methods, the first is more coherent from a computation point of view. In fact, this method uses partial derivatives and works on the real values of the database variables while the Profile method takes the variables one by one and reduces their values into different scales of variation, that is, uses a fictitious matrix. Concerning the ability of all the methods to classify the input variables in order of importance, the results observed for each method are not always the same. Their different computation leads to different results. For instance, the necessity to use a new model for each variable selection skews the results of the classical stepwise

method. The confidence intervals plotted for each method indicate their stability. The PaD method gives the most stable results.

These methods of contribution analysis of input variables seem to be very interesting to use but it is important to underline the need of an ecologist's opinion regarding the ranking of importance of inputs and their mode of action on the output. If an ecologist's opinion is unavailable, two of these methods (for instance PaD and Perturb) should be used to analyse contribution of the inputs and compared. Analysis of the inputs by the two methods must not be different. If it is the case, the network may be poorly calibrated or the data may be very difficult to analyse.

In ecology, it is important to predict phenomena that occur in the studied environment. Moreover, the phenomena have to be understood and this is often very difficult due to their complexity. ANNs are tools that can resolve prediction problems, and this ANNs property is now well understood (Edwards and Morse, 1995; Colosanti, 1991). Adding new methods to ANNs allowing the analysis of the contributions of the different variables will help in understanding the ecological phenomenon and finally in finding solutions to act on it, restore it and improve the environmental conditions for life.

Acknowledgements

Funding for this research was provided by the EU project PAEQANN (N°EVK1-CT1999-00026).

Appendix A: Example illustrating the partitioning of weights

This appendix details the procedure for partitioning the connection weights to determine the relative importance of the various inputs, using the method proposed by Garson, 1991. The method essentially involves partitioning the hidden-output connection weights of each hidden neuron into components associated with each input neuron.

Consider the neural network with three input neurons, four hidden neurons and one output neuron with the connection weights as shown below, as an example.

Hidden neurons	Weights			
	Input 1	Input 2	Input 3	Output
Hidden 1	-1.67624	3.29022	1.32466	4.57857
Hidden 2	-0.51874	-0.22921	-0.25526	-0.48815
Hidden 3	-4.01764	2.12486	-0.08168	-5.73901
Hidden 4	-1.75691	-1.44702	0.58286	-2.65221

The computation process is as follows:

(1) For each hidden neuron i , multiply the absolute value of the hidden-output layer connection weight by the absolute value of the hidden-input layer connection weight. Do this for each input variable j . The following products P_{ij} are obtained:

	Input 1	Input 2	Input 3
Hidden 1	$P_{11} = 1.67624 \times 4.57857$	$P_{12} = 3.29022 \times 4.57857$	$P_{13} = 1.32466 \times 4.57857$
Hidden 2	$P_{21} = 0.51874 \times 0.48815$	$P_{22} = 0.22921 \times 0.48815$	$P_{23} = 0.25526 \times 0.48815$
Hidden 3	$P_{31} = 4.01764 \times 5.73901$	$P_{32} = 2.12486 \times 5.73901$	$P_{33} = 0.08168 \times 5.73901$
Hidden 4	$P_{41} = 1.75691 \times 2.65221$	$P_{42} = 1.44702 \times 2.65221$	$P_{43} = 0.58286 \times 2.65221$

(2) For each hidden neuron, divide P_{ij} by the sum for all the input variables to obtain Q_{ij} . For example for Hidden 1, $Q_{11} = P_{11}/(P_{11} + P_{12} + P_{13}) = 0.266445$.

(3) For each input neuron, sum the product S_j formed from the previous computations of Q_{ij} . For example, $S_1 = Q_{11} + Q_{21} + Q_{31} + Q_{41}$.

	Input 1	Input 2	Input 3
Hidden 1	$Q_{11} = 0.266445$	$Q_{12} = 0.522994$	$Q_{13} = 0.210560$
Hidden 2	$Q_{21} = 0.517081$	$Q_{22} = 0.228478$	$Q_{23} = 0.254441$
Hidden 3	$Q_{31} = 0.645489$	$Q_{32} = 0.341388$	$Q_{33} = 0.013123$
Hidden 4	$Q_{41} = 0.463958$	$Q_{42} = 0.382123$	$Q_{43} = 0.153919$
Sum	$S_1 = 1.892973$	$S_2 = 1.474983$	$S_3 = 0.632044$

(4) Divide S_j by the sum for all the input variables. Expressed as a percentage, this gives the relative importance or distribution of all output weights attributable to the given input variable. For example, for the input neuron 1, the

relative importance is equal to $(S_1 \times 100)/(S_1 + S_2 + S_3) = 47.3\%$.

	Input 1	Input 2	Input 3
Relative importance (%)	47.3	36.9	15.8

References

- Balls, G.-R., Palmer Brown, D., Sanders, G.-E., 1996. Investigating microclimatic influences on ozone injury in clover (*Trifolium subterraneum*) using artificial neural networks. *New Phytologist* 132, 271–280.
- Brosse, S., Guegan, J.-F., Tourenq, J.-N., Lek, S., 1999. 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.
- Colosanti, R.L., 1991. Discussions of the possible use of neural network algorithms in ecological modelling. *Binary* 3, 13–15.
- Crisp, D.T., Carling, P.A., 1989. Observations on sitting, dimension and structure of salmonid redds. *Journal of Fish Biology* 34, 119–134.
- Delacoste, M., 1995. Analyse de la variabilité spatiale de la reproduction de la truite commune (*Salmo trutta L.*). Etude à l'échelle du micro et du macrohabitat dans 6 rivières des Pyrénées centrales. Ph.D. of Institut National Polytechnique de Toulouse, p. 133
- Delacoste, M., Baran, P., Dauba, F., Belaud, A., 1993. Etude du macrohabitat de reproduction de la truite commune (*Salmo trutta L.*) dans une rivière pyrénéenne, La Neste du Louron. Evaluation d'un potentiel de l'habitat physique de reproduction. *Bulletin Français de la Pêche et de la Pisciculture* 331, 341–356.
- Dimopoulos, Y., Bourret, P., Lek, S., 1995. Use of some sensitivity criteria for choosing networks with good generalization ability. *Neural Processing Letters* 2, 1–4.
- Dimopoulos, I., Chronopoulos, J., Chronopoulou Sereli, A., Lek, S., 1999. Neural network models to study relationships between lead concentration in grasses and permanent urban descriptors in Athens city (Greece). *Ecological Modelling* 120, 157–165.
- Edwards, M., Morse, D.R., 1995. The potential for computer-aided identification in biodiversity research. *Tree* 10 (4), 153–158.
- El-Keib, A.A., Ma, X., 1995. Application of artificial neural networks in voltage stability assessment. *IEEE Transactions on Power Systems* 10, 1890–1896.
- Engelbrecht, A.P., Cloete, I., Zurada, J.M., 1995. Determining the Significance of Input Parameters Using Sensitivity Analysis. From Natural to Artificial Neural Computation. Springer, Malaga-Torremolinos, Spain.
- Garson, G.D., 1991. Interpreting neural network connection weights. *Artificial Intelligence Expert* 6, 47–51.
- Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. *Neural Computation* 4, 1–58.
- Goh, A.T.C., 1995. Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering* 9, 143–151.
- Guo, Z., Uhrig, R.E., 1992. Using modular neural networks to monitor accident conditions in nuclear power plants, SPIE. *Applications of Artificial Neural Networks III* 1709, 505–516.
- Hsu, C.T., Tzeng, Y.M., Chen, C.S., Cho, M.Y., 1995. Distribution feeder loss analysis by using an artificial neural network. *Electric Power Systems Research* 34, 85–90.
- Kim, S.H., Yoon, C., Kim, B.J., 2000. Structural monitoring system based on sensitivity analysis and a neural network. *Computer-aided Civil and Infrastructure Engineering* 155, 309–318.
- Lae, R., Lek, S., Moreau, J., 1999. Predicting fish yield of African lakes using neural networks. *Ecological Modelling* 120, 325–335.
- Lek, S., Belaud, A., Dimopoulos, I., Lauga, J., Moreau, J., 1995. Improved estimation, using neural networks, of the food consumption of fish populations. *Marine Freshwater Research* 46, 1229–1236.
- Lek, S., Belaud, A., Baran, P., Dimopoulos, I., Delacoste, M., 1996a. Role of some environmental variables in trout abundance models using neural networks. *Aquatic Living Resources* 9, 23–29.
- Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S., 1996b. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling* 90, 39–52.
- Lek, S., Guiresse, M., Giraudel, J.-L., 1999. Predicting stream nitrogen concentration from watershed features using neural networks. *Water Research* 33, 3469–3478.
- Liong, S.Y., Lim, W.H., Kojiri, T., Hori, T., 2000a. Advance flood forecasting for flood stricken Bangladesh with a fuzzy reasoning method. *Hydrological Processes* 14, 431–448.
- Liong, S.Y., Lim, W.H., Paudyal, G.N., 2000b. River stage forecasting in Bangladesh: neural network approach. *Journal of Computing in Civil Engineering* 14, 1–8.
- Lu, R.S., Lai, J.L., Lo, S.L., 1998. Predicting solute transfer to surface runoff using neural networks. *Water Science and Technology* 38, 173–180.
- Maier, H.R., Dandy, G.C., 1996. The use of artificial neural networks for the prediction of water quality parameters. *Water Resources Research* 32, 1013–1022.
- Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resource variables: a review of modelling issues and applications. *Environmental Modelling and Software* 15, 101–124.
- Maier, H.R., Dandy, G.C., Burch, M.D., 1998. Use of artificial neural networks for modelling *Cyanobacteria anabaena* spp. in the river Murray, South Australia. *Ecological Modelling* 105, 257–272.

- Malavoi, J.R., 1989. Typologie des faciès d'écoulement ou unités morpho-dynamiques d'un cours d'eau à haute énergie. *Bulletin Français de la Pêche et de la Pisciculture* 315, 189–210.
- Manel, S., Dias, J.M., Ormerod, S.J., 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological Modelling* 120, 337–347.
- Mastrorillo, S., Dauba, F., Oberdorff, T., Guegan, J.-F., Lek, S., 1998. Predicting local fish species richness in the Garonne river basin. *Comptes Rendus de l'Académie des Sciences Serie III Sciences de la Vie* 321, 423–428.
- Ottaway, E.M., Carling, P.A., Clarke, A., Reader, N.A., 1981. Observation on the structure of brown trout, *Salmo trutta Linnaeus*, redds. *Journal of Fish Biology* 19, 593–607.
- Özesmi, S.L., Özesmi, U., 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecological Modelling* 116, 15–31.
- Paruelo, J.M., Tomasel, F., 1997. Prediction of functional characteristics of ecosystems: a comparison of artificial neural networks and regression models. *Ecological Modelling* 98, 173–186.
- Ramos-Nino, M.E., Ramirez-Rodriguez, C.A., Clifford, M.N., Adams, M.R., 1997. A comparison of quantitative structure-activity relationships for the effect of benzoic and cinnamic acids on *Listeria monocytogenes* using multiple linear regression, artificial neural network and fuzzy systems. *Journal of Applied Microbiology* 82, 168–176.
- Recknagel, F., French, M., Harkonen, P., Yabunaka, K.-I., 1997. Artificial neural network approach for modelling and prediction of algal blooms. *Ecological Modelling* 96, 11–28.
- Rubin, J.-F., 1995. Estimating the success of natural spawning of salmonids in streams. *Journal of Fish Biology* 46, 603–622.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by backpropagation error. *Nature* 323, 533–536.
- Scardi, M., Harding, L.W., 1999. Developing an empirical model of phytoplankton primary production: a neural network case study. *Ecological Modelling* 120 (2–3), 213–223.
- Shirvell, C.S., Dungey, R.G., 1983. Microhabitats chosen by brown trout for feeding and spawning in rivers. *Transactions of the American Fish. Society* 112, 355–367.
- Skelton, P.H., Cambay, J.A., Lombard, A., Benn, G.A., 1995. Patterns of distribution and conservation status of freshwater fishes in South Africa. *South Africa Journal of Zoology* 30, 71–81.
- Starrett, S.K., Adams, G.L., 1997. Using artificial neural networks and regression to predict percentage of applied nitrogen leached under turfgrass. *Communications in Soil Science Plant Analytical* 28, 497–507.
- Sung, A.H., 1998. Ranking importance of input parameters of neural networks. *Expert Systems with Applications* 15, 405–411.
- Tomassone, R., Lesquoy, E., Miller, C., 1983. *La régression, Nouveaux Regards sur une ancienne méthode Statistique*. INRA, Paris.
- Tomassone, R., Dervin, C., Masson, J.P., 1993. *Biométrie. Modélisation de phénomènes biologiques*. Masson.
- van Wijk, M.T., Bouten, W., 1999. Water and carbon fluxes above European coniferous forests modelled with artificial neural networks. *Ecological Modelling* 120, 181–197.
- Weisberg, S., 1980. *Applied Linear Regression*. Wiley, New York.
- Whitehead, P.G., Howard, A., Arulmani, C., 1997. Modelling algal growth and transport in rivers: a comparison of times series analysis, dynamic mass balance and neural network techniques. *Hydrobiologia* 349, 39–46.
- Yang, C.C., Prasher, S.O., Tan, C.S., 1998. An artificial neural network model for water table management systems, pp. 250–257 in *Drainage in the 21st century: food production and the environment: proceedings of the seventh International Drainage Symposium*. Larry C. Brown, symposium proceedings Editor. St. Joseph, Mich, USA, 738 p.
- Yao, J., Teng, N., Poh, H.L., Tan, C.L., 1998. Forecasting and analysis of marketing data using neural networks. *Journal of Information Science and Engineering* 14, 843–862.
- Zurada, J.M., Malinowski, A., Cloete, I., 1994. Sensitivity analysis for minimization of input data dimension for feedforward neural network. *ISCAS'94. IEEE International Symposium on circuits and Systems*, vol. 6, IEE Press, London, pp. 447–450.