Prediction and explanation of sound quality indicators by multiple linear regressions and artificial neural networks

L. Brocolini[a], C. Lavandier[a], C. Marquis-Favre[b], M. Quoy[c] and M. Lavandier[b]

[a]Université Cergy Pontoise, 5 mail gay lussac, IUT Génie Civil, 95031 Cergy Pontoise, France
[b]Université de Lyon - ENTPE, rue Maurice Audin, 69120 Vaulx-En-Velin, France
[c]Equipes Traitement de l'Information et Systèmes, 6, avenue du Ponceau. F 95014 Cergy-Pontoise Cedex
laurent.brocolini@u-cergy.fr

The purpose of this study is to develop a predictive model of urban sound quality from field survey data using multiple linear regressions and artificial neural networks (ANNs). In order to determine a sound quality indicator, 320 passers-by were asked to assess their environment mainly from an acoustic point of view but also from a global perspective (visual and air quality). The investigation took place in two large cities in France, and involved 8 different kinds of typical sound environments: park, pedestrian street, boulevard, street and urban transitions such as transition between park and boulevard. In each place, passers-by had to evaluate 26 subjective variables on 11-point scales. The collected data were analyzed according to two methods: multiple linear regressions and artificial neural networks. The resulting models were compared. The regression model was more self-explanatory about the influence of each variable whereas the ANN model made it possible to differentiate the influence of each variable depending on the type of the environment.

1 Introduction

The study presented here aims at developing a global model predicting soundscape quality that would be relevant for different urban locations. Two distinct approaches have been compared. The first one was based on neural network methods, and the second one used multiple linear regressions.

Through a questionnaire, the soundscape quality was assessed by passers-by as the pleasantness of the sound environment. The soundscape was also characterized by other subjective variables. From this set of data it was possible to find a relationship between sound environment pleasantness and the other explanatory variables.

In addition to this relationship, and by relying on cross-validation method, a particular attention has been given to the ability of the established models to predict a sound pleasantness value when new data are provided.

2 Methodology

2.1 Locations and periods of the study

Field surveys were conducted in two large cities in France. The first site is located in Lyon, in the 6th district, near and in an urban park named "Parc de la Tête d'Or". This is a big urban park bordered by two large boulevards. The second site is located in the 5th district of Paris.

In each of these cities, surveys were conducted in four locations. In Lyon the four selected locations are situated as follows: one in the park, two on both sides of the main entrance and one near a boulevard ("Boulevard des Belges"). In Paris, the four locations are situated in a pedestrian street named "Rue Mouffetard", in a one-lane circulated street ("Rue de l'Epée de Bois") and finally near a boulevard ("Boulevard Monge"). The nomenclature used in this study for these eight locations is: "Park", "Transition (Park)", "Transition (Bld / Lyon)", "Boulevard (Lyon)", "Pedestrian street", "Transition (Pedestrian)", "Transition (Street)" and "Boulevard (Paris)". Figure 1 shows the exact position of these eight locations.

From a previous study [1] and based on acoustic measurements, these eight locations correspond to different kind of sound environments. In the "Park", the prominent sound sources are birds, nature, voices, walking sounds. In the "Pedestrian street", many people and a lot of commercial and working activities can be found. Of course, "Boulevards (Lyon and Paris)" gather a lot of road sound sources such as cars, motorbikes, heavy trucks, horns, etc. … And finally, in the "Transitions" there are not homogeneous sound environments. So, actually, many sound sources are present. Completed during different months of the year 2009 in Lyon and 2010 in Paris, every survey was carried out the afternoons, between 2 P.M and 8 P.M, Saturdays and Sundays excluded, in order to get some homogeneity in the different day studies [2,3].

Figure 1: Selected locations for investigation in Lyon and Paris.

2.2 Questionnaire

The aim of the survey was to gather people's perception of their environment at the place and during the time of the interview (about ten minutes). Based on previous studies [4,5,6,7], it mainly consisted in closed questions taking the form of semantic differential with a continuous graduated scale. Subjects had to answer using scale (see Figure 2).

The survey was made so that respondents first considered the environment from a global perspective, and then dwelled on various aspects of the environment (acoustic, visual, air quality) but still in their entirety, and finally ended with the identification of the sound sources. Thus the questionnaire could be divided into five parts.

For the first part of the questionnaire, subjects were asked to assess the overall environmental quality, telling in a few words why they thought it was pleasant or unpleasant, and to note this pleasantness on the presented scale (Figure 2).

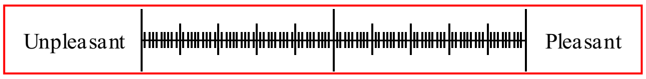

Figure 2: Global, sound, visual and air quality pleasantness scale.

Subsequently people were asked to focus on the sound environment as a whole, and then to answer questions about different characteristics. For a better understanding an explanation was given under each adjective (see Figure 3).

In the third part, subjects were asked to assess the visual pleasantness, the perception of air quality but also to evaluate the familiarity of the soundscape.

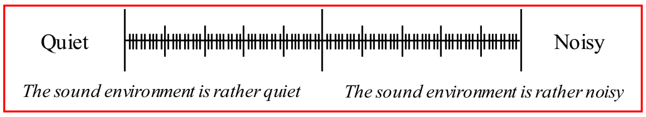

Figure 3: Example of question on the sound environment.

The fourth part of the questionnaire concerned the sound sources. Subjects were asked to focus on sound sources, to specify which ones they were able to identify, and for each one to estimate their loudness and their time ratio of presence (based on the duration of the survey). After that, they received a sound source list and had to clarify if they noticed or not these sound sources, and if so to rate their loudness and time ratio of presence (Figure 4).

Finally, subjects were asked if they thought the sound environment was suitable for their activity.

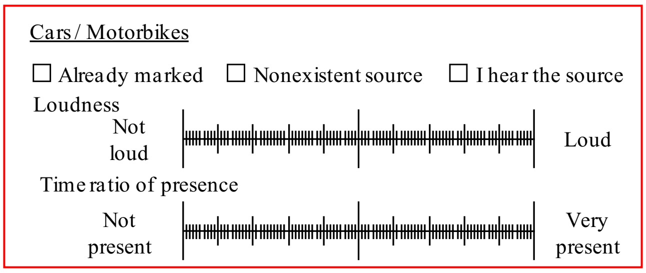

Figure 4: Example of sound sources marking.

Table 1 presents all the variables measured by each subject. Each variable was noted on the same scale and was linearly transformed into a value ranging from 0 to 10.

Table 1: Measured variables.

	Measured variables	
Pleasantness	(1) Sound pleasantness	
	(2) Global pleasantness	
	(3) Visual pleasantness	
	(4) Air quality pleasantness	
Soundscape global characteristics	(5) Quiet / Noisy	
	(6) Stable / Changing	
	(7) Lifeless / Lively	
	(8) Enveloping / Not Enveloping	
	(9) Surprising / Familiar	
	(10) Unsuitable / Suitable	
Sound Sources	(11) PL.LV (12) TP.LV	Cars / Motorbikes (Light Vehicles)
PL = Perceived Loudness	(13) PL.Mop (14) TP.Mop	Mopeds
	(15) PL.TB (16) TP.TB	Trucks/Buses
	(17) PL.H (18) TP.H	Horns
	(19) PL.Act (20) TP.Act	Activities
TP = Time ratio of presence	(21) PL.HP (22) TP.HP	Human Presence
	(23) PL.Bir (24) TP.Bir	Birds
	(25) PL.Nat (26) TP.Nat	Nature

2.3 Subjects

320 passers-by were interviewed (40 at each location). The only personal data collected on subjects were gender and age category, which was evaluated by the experimentater following three classes: adolescent, adult, senior. Although these variables have been identified, they have not been taken into account in the development of the models.

Indeed, studies have shown that the relationship between gender or age and evaluation of sound environment was not significant [8,9]. In this study, chi-square tests of independence showed that gender had no influence on the judgment of the soundscape quality ("Sound pleasantness") neither on the "Silence" variable. The subject responses for these variables ((1) and (5) in Table 1) were divided into three categories, namely for the "Sound pleasantness": unpleasant (value < 3.6), neutral (3.6 < value < = 6.7) and pleasant (value > 6.7), and for the "Silence" variable: noisy (value < 3.6), neutral (3.6 < value < = 6.7) and quiet (value > 6.7).

Distributions for men / women are presented in Figure 5. For the "Sound pleasantness", χ^2 is 0.57 with a p-value equal to 0.75. In the case of the "Silence", $\chi^2 = 3.61$ and $p = 0.16$. In both cases, "Silence" and "Sound pleasantness" were not dependent on gender.

The age category was divided into three classes: adolescent, adult and senior. From the 320 subjects, only 13 were seniors. However, in order to have a chi-square test of independence valid, each sample must be greater than 5, which was not the case in our study (only 2 seniors evaluated the location as noisy or quiet, only 4 seniors evaluated the location as pleasant). Thus, the dependence of pleasantness on subject age was not tested.

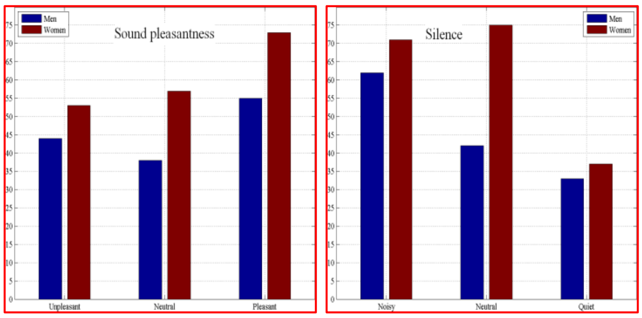

Figure 5: Distribution men / women for "Sound pleasantness" in three categories ($\chi^2 = 0.57$ / $p = 0.75$).and for "Silence" in three categories ($\chi^2 = 3{,}61$ / $p = 0{,}16$).

3 Analysis

3.1 Variable selection

Based on the Bravais-Pearson correlation coefficients between each pair of independent variables with a 95 % confidence interval, the following variables were not taken into account with the aim of proposing soundscape quality models: "Global pleasantness", "Air quality pleasantness", "Suitability" and "Stable/Changing".

Only one aspect of the sound source, either the perceived loudness or the time ratio of presence, was kept to build the models, because these variables were highly correlated. In a previous study on the contribution of the sources in the characterization of sound environments [10], the time ratio of presence seemed to be a better indicator than loudness to explain the perceived environmental sound quality. However, in the present study, the perceived loudness has sometimes more variability than the time ratio of presence. So we paid attention to the response distribution and the one with the higher variability was kept.

Table 2 presents the thirteen variables chosen to explain the sound pleasantness through the proposed models of multiple linear regression and artificial neural network.

Table 2: Variables chosen to explain sound pleasantness

	Nomenclature	Measured variables
Dependent variable	Snd.Pl	Sound pleasantness
Independent variables	Vis.Pl	Visual pleasantness
	Sil	Silence
	Liv	Liveliness
	Env	Envelopment
	Surp	Surprising
	TP.LV	Light vehicles
	PL.Mop	Mopeds
	PL.TB	Trucks & Buses
	PL.Hor	Horns
	TP.Act	Activities
	TP.Hum	Human presence
	TP.Bir	Birds
	TP.Nat	Nature

3.2 Models

The multiple linear regression (REG) analysis informs about the linear relationship between the dependent variable (in our case the sound pleasantness) and the 13 independent variables [11].

The artificial neural networks (ANN) are inspired by the structure of biological neurons. A database with inputs (independent variables) and targets (dependent variable) is presented. The network approximates with a non-linear function the relationship between inputs and targets and can modify this function to minimize the error between the calculated outputs and the target values [12]. In our study we used a backpropagation multilayer perceptron provided by the Neural Network Matlab Toolbox. This network had 13 input neurons (the measured variables), one output neuron (sound pleasantness) and 13 neurons in the hidden layer.

3.3 Procedure

In order to compute and test the predictive models we used a cross-validation technique. So, the database (320 subjects) was divided into two databases. One to set up the model (construction database) and one to test it (test database). For the neural network model, the first database was divided into a learning database and a validation database. The proportion of the three databases (learning, validation, testing) was respectively 50-20-30 %, i.e. 160 subjects for learning, 64 for validation and 96 for testing. The subjects were randomly assigned to each database.

To quantify a model quality (REG or ANN) in terms of adjustment and prediction, we used the determination coefficient (R^2) which will be the square Bravais-Pearson correlation coefficient when the measured sound pleasantness and the sound pleasantness calculated by a model will be compared. So, for one multiple linear regression, two determination coefficients were calculated. One for the construction database (R_c^2) and one for the test database (R_t^2).

However the choice of construction and test databases is really important and the regression and neural network results are strongly dependent on these databases. Two distinct construction databases can lead to different linear regression results. Also to optimize the choice of databases, 1000 sets of databases (construction and test) were randomly drawn. Among the 1000 database picking, we select one which gave the best couple (R_c^2, R_t^2). R^2 is always between 0 and 1, so best couple (R_c^2, R_t^2) is one which is closest to (1, 1) in terms of Euclidean distance.

Then, from the best construction and test databases, we were able to establish a model of neural network. Because there are random parameters in the ANN model, a same construction database will eventually lead to a different result to each new run. Also to compare ANN and REG models in terms of prediction, we decided to compute an ANN model from the best databases drawing and keep this one only if the couple (R_c^2, R_t^2) from this model was better than the one from the corresponding regression. In other words, we seek a neural network which satisfies R_c^2 (ANN) > R_c^2 (REG) and R_t^2 (ANN) > R_t^2 (REG). While this condition was not assumed, the neural network model was rejected and another one from the same construction and test databases was computed, until the condition was finally satisfied.

4 Results

4.1 Multiple linear regression

The equation (1) of the best multiple linear regression model obtained according to the procedure described in paragraph 3.3 is presented below.

$$\begin{aligned} Snd.Pl = 0.09 \; & + 0.55{*}Sil + 0.30{*}Vis.Pl \\ & + 0.15{*}Liv + 0.10{*}Env \\ & - 0.06{*}Surp - 0.12{*}TP.LV \\ & - 0.05{*}PL.Mop - 0.02{*}PL.TB \\ & - 0.08{*}PL.Hor + 0.00{*}TP.Act \\ & + 0.02{*}TP.Bir + 0.01{*}TP.Hum \\ & + 0.05{*}TP.Nat \end{aligned} \quad (1)$$

It appears in equation (1) that the most important variables explaining the sound pleasantness were the “Silence” (0,55), the “Visual pleasantness” (0,30), the “Liveliness” (0,15) and the “Time ratio of presence of light vehicles” (-0,12).

4.2 Comparison REG / ANN

One of the main goals of the project from which this study is part of was to check the advantage of artificial neural networks over multiple linear regressions in terms of prediction. To this end, we have tried to find an artificial neural network using the method described in paragraph

3.3, i.e. to retain an artificial neural network that satisfied the conditions R_c^2 (ANN) > R_c^2 (REG) and R_t^2 (ANN) > R_t^2 (REG). Unfortunately, after many attempts, it was not possible to satisfy this condition. Then, the criterion to keep an artificial neural network model was revised downward, and we tried to find a neural network model that satisfied the condition R_{all}^2 (ANN) > R_{all}^2 (REG). R_{all}^2 is the coefficient of determination which characterizes the correlation between the measured sound pleasantness and the calculated values obtained for the whole database denoted by "All" (320 subjects).

Table 3: R^2 values of both models.

	REG		ANN	
Database	R^2	p-value	R^2	p-value
Construction	0,5761	2,96E-43	0,6722	1,10E-55
Test	0,7044	1,30E-26	0,6957	5,05E-26
All	0,6076	1,45E-66	0,6744	1,78E-79

Table 3 presents the results of correlation coefficients obtained for each of the three databases (construction, test and all) for both selected models (REG and ANN). Comparing the R^2 values (construction and test) from REG and ANN models, it appears that the latter was better from a construction database point of view. On the other hand from a strictly predictive point of view, (using the test database), it is not possible here to say that a model is better than another because the values of R_t^2 are almost equal in the case of the two models. However, the values very close to these two models allow us to go further in the comparison.

4.3 Artificial neural network

Neural networks are generally considered as "black boxes". Nevertheless, it is possible for a local average subject to look at how the calculated sound pleasantness varies. A local average subject was here simply the mean of all subject answers in one location for each of the measured variables. Each local average subject was injected into the model by varying each of the variables one by one between 0 and 10 with a step of 0.1. The output plots, calculated from the two models, show the relationship between the dependent variable and each of the explanatory variables.

For instance, consider the regression equation presented in paragraph 4.1. If all variables except one ("Visual pleasantness" for example) are fixed to a mean value and that "Visual pleasantness" varies continuously from 0 to 10, result from this equation defines a straight line whose slope is equal to the "Visual pleasantness" coefficient in the equation, i.e. 0.30. The idea is therefore to apply the same method to the artificial neural network in order to visualize the relationship between sound pleasantness and the variables considered in the model.

Figure 6 shows the output of the ANN model with local average subjects of Lyon as inputs. Red curves represent the Park, green curves symbolize the Transition (Park), blue curves are the answer to the Transition (Bld / Lyon) and finally purple curves illustrate Boulevard (Lyon). The black line is the output of REG model with a global average subject as input. (i.e. considering all the locations studied). Outputs calculated with Paris local average subjects as inputs are presented in Figure 7.

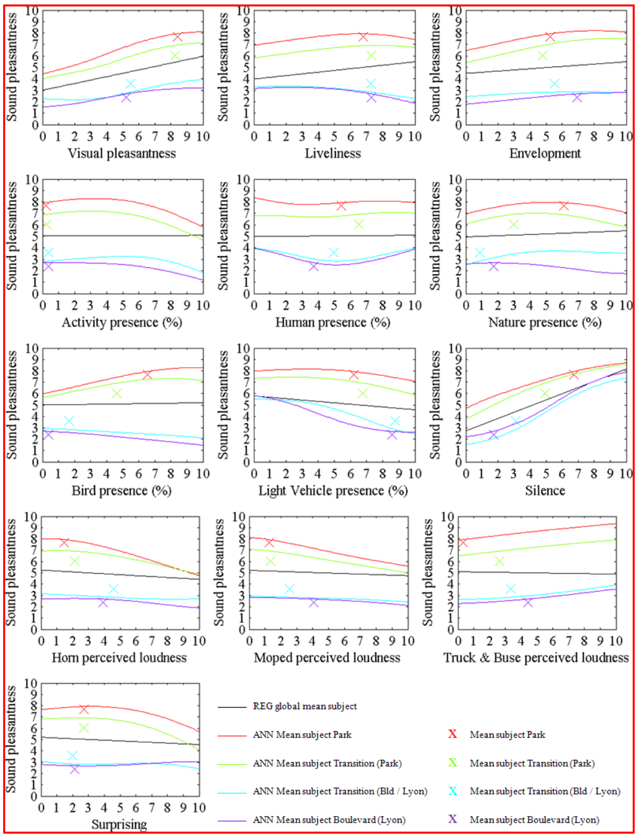

Figure 6: Output curves in Lyon.

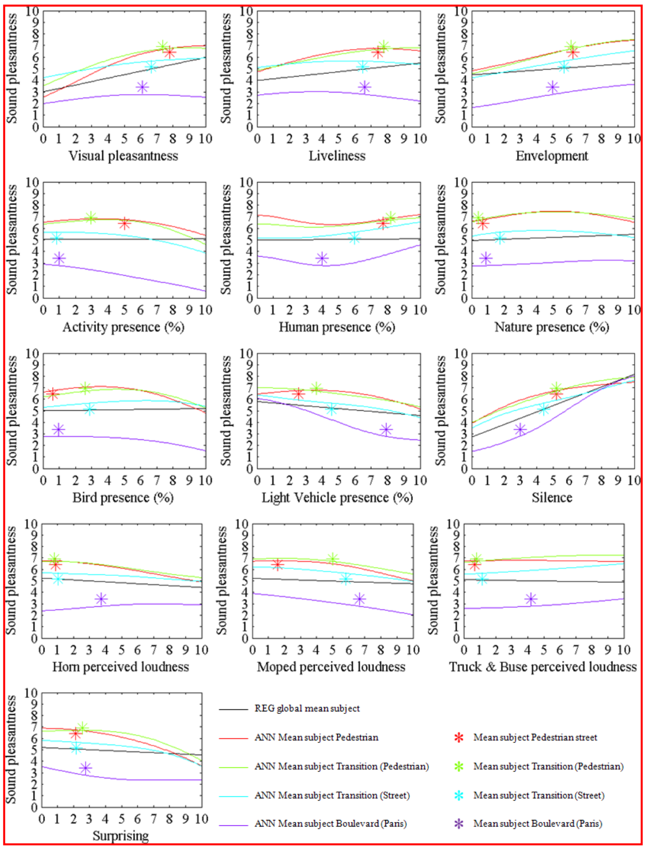

Figure 7: Output curves in Paris.

Because of the linearity of regression, whoever the subject chosen as input, the slope of the output line will be the same. Differences of slope around a local average subject may therefore be highlighted only with the artificial neural network. Although the relationship is not linear, each slope was calculated around average subjects in the range [-1 + 1]. The results of the slopes are presented in Table 4 for the variable "Visual pleasantness" as an example. The higher the slope, the more important the "Visual pleasantness".

Table 4: Slopes of both models for "Visual pleasantness".

	REG slope	ANN slope
Park	0,30	0,19
Transition (Park)	0,30	0,22
Transition (Bld / Lyon)	0,30	0,32
Boulevard (Lyon)	0,30	0,24
Pedestrian street	0,30	0,20
Transition (Pedestrian)	0,30	0,13
Transition (Street)	0,30	0,11
Boulevard (Paris)	0,30	0,00

The ANN model shows that this importance of "Visual pleasantness" is not the same according to the different locations. Indeed, influence of "Visual pleasantness" is the same with either REG or ANN models in location "Transition (Bld / Lyon)" (slope = 0.32). However, it is lower in "Transition (Pedestrian)" (0.13) and 0.00 on the "Boulevard" in Paris. Therefore, it seems that the influence of the vision is more important in the transition area in Lyon than in the Boulevard in Paris.

5 Conclusion and perspective

The purpose of this study was to develop two predictive sound quality models from data measured in 8 places with different sound environments. The first model is built with multiple linear regression and the second one with artificial neural network. The comparison of these two models shows that from a predictive point of view they are very similar. However, the advantage of artificial neural network is that it is possible to highlight the relative influence of each variable on the sound environment quality for a specific location. This influence is shown with the slopes calculated of each variable in the model. Unfortunately, these slopes depend on the selected "best" neural network. Other good artificial neural networks ($Rall^2$ (ANN) > $Rall^2$ (REG)) give different slopes and draw different interpretations. So in a future work, it is necessary to select a large number of good artificial neural networks, calculate the slopes of outputs for each of them and carry out an analysis of variance to compare means, in order to validate and to interpret the relative influence of each variable.

Acknowledgments

This project has been supported by the French Energy Agency ADEME in the frame of the PREDIT program.

References

[1] L. Brocolini, C. Lavandier, M. Quoy, C. Ribeiro, "Discrimination of urban soundscapes through Kohonen map", *Proceedings of Euronoise 09*, Edinburgh (2009)

[2] C. Lavandier C., B; Barbot, "Influence of the temporal scale on the relevance of acoustic parameters selected to characterize urban sound environments", *Proceedings of Euronoise 03* Naples (2003)

[3] L. Brocolini, C. Lavandier, C. Ribeiro, M. Quoy, "Optimization of the acoustic measurement duration in urban environment through Kohonen self-organizing map", *Proceedings of Inter Noise 2010*, Lisbon (2010)

[4] S. Viollon, "Influence des informations visuelles sur la caractérisation de la qualité acoustique de l'environnement urbain", *PhD thesis*, Université de Cergy-Pontoise (2000)

[5] M. Raimbault, "Simulation des ambiances sonores urbaines: intégration des aspects qualitatifs", *PhD thesis*, Université de Nantes (2002)

[6] D. Dubois, C. Guastavino, M. Raimbault, "A cognitive approach to urban soundscapes : using verbal data to access everyday life auditory categories", *Acta Acustica united with Acustica* 92, 865-874 (2006)

[7] C. Guastavino, "The ideal urban soundscape: investigating the sound quality of French cities", *Acta Acustica united with Acustica* 92, 945-941 (2006)

[8] L. Yu, "Soundscape Evaluation and ANN Modelling in Urban Open Spaces", *PhD thesis*, School of Architecture, University of Sheffield, United Kingdom (2009)

[9] S. Marry, "L'espace public sonore ordinaire. Les paramètres de la perception sonore dans les espaces publics. Contribution à une connaissance de l ambiance sonore", *PhD thesis*, Université de Grenoble, France (2011)

[10] C. Lavandier, B. Defréville, "The contribution of sound source in the assessment of urban soundscapes", *Acta Acustica united with Acustica* 92, 912-921 (2006)

[11] D. Howell, *Statistical Methods for Psychology*, Wadsworth Edition (2009)

[12] S. Thiria, Y. Lechevallier, O. Gascuel, S Canu, *Statistiques et méthodes neuronales*, Dunod (1997)