Exploring an ANN Modeling Approach that Combines Accessibility and Mobility into a Single Trip Potential Index for Strategic Planning

Antônio Nélson Rodrigues da Silva
Departament of Transportation, São Carlos School of Engineering, University of São Paulo
Av. Trabalhador São-carlense, 400, 13566-590, São Carlos - SP - BRAZIL
E-mail address: anelson@sc.usp.br

Archimedes Azevedo Raia Jr.
Departament of Civil Engineering, University Federal of São Carlos
Via Washington Luis, km 235, 13565-905, São Carlos - SP - BRAZIL
E-mail address: raiajr@power.ufscar.br

Charlie Williams Rengifo Bocanegra
Departament of Transportation, São Carlos School of Engineering, University of São Paulo
Av. Trabalhador São-carlense, 400, 13566-590, São Carlos - SP - BRAZIL
E-mail address: crengifo@sc.usp.br

ABSTRACT

The objective of this work was the development of a modeling approach that uses Artificial Neural Networks (ANN) in the estimation of a trip potential index for strategic planning. In the application described in this paper for introducing the proposed approach, a mean separation accessibility index has been at first estimated for all households and subsequently linked to mobility variables taken from an O-D survey. The output variables, i.e. trip characteristics (number and length), were also taken from the same O-D survey or calculated in a Geographic Information System-environment. The input variables identified here as relevant to the model were: size and income of the household, which may be associated to mobility, and the accessibility indicator itself. As the practical use of the ANN model first created was somehow limited due to the need of the software used for training the network, the trained network was thereafter replicated in an electronic spreadsheet. That makes it a tool able to conduct, with speed and flexibility, several different analyses. The application presented here makes clear that the way the results are then visualized helps even in understanding the logic behind the model. For strategic planning, the methodology presented in this work seems to be a step forward in relation to traditional accessibility models and it may be a useful tool for urban and transportation planners and decision-makers. In addition, the case studied here stressed the fact that urban citizens in developing countries need not only physical accessibility, but also better mobility conditions.

Keywords: strategic planning; trip potential; accessibility; mobility; Artificial Neural Networks.

5242 words + 4 pictures (250 words each) + 4 tables (250 words each) = 7242 words

INTRODUCTION

Several Brazilian cities have undergone a very fast growth process in the last decades. The intense population migration towards the cities has been produced basically by the perspective of economic development. As a consequence, the automobile use is rapidly and continuously growing. That brings some serious impacts on public transportation systems, which have to survive in a scenario of a continuously decreasing demand combined with growing costs. In addition, these impacts are not evenly distributed to all population groups. The low income groups are affected with much more intensity for two reasons: they are usually captive users of public transportation, and they expend a large share of their budget with transportation. This situation has undesirable costs from a social point of view and it is highly inefficient from an economic perspective. Even considering that the continuity of this conjuncture will not bring much improvement in the quality of life of the entire urban population, the planning strategies used nowadays somehow help to keep the situation unchanged.

In developing countries there is often a gap and, even worse than that, sometimes a conflict between the transportation planning objectives and the real needs of the population. In most cases, equity issues should be more effectively incorporated in the planning process, as a way to reduce the impacts of the transportation changes on the low income segments of the society. A first step to tackle this problem is to rethink the planning process itself.

Therefore, bearing in mind that new models or new approaches to the old models are required, the objective of this paper is to explore a model for strategic planning that joins in a single trip potential index, on one side, accessibility elements and, on the other side, factors that govern mobility. The proposed approach uses Artificial Neural Networks (ANNs) to generate the index in four steps: *i*) the estimation of spatial attributes of the trips described in a Origin-Destination survey, *ii*) the conception of preliminary models for the evaluation of the input-output variables; *iii*) the evaluation of the relevance of each input variable in the model predictions; and *iv*) the formulation of new models only with the variables with the best performance. These steps are described in that same order in the paper, following a brief review of the literature concerning the use of Artificial Neural Networks in transportation planning.

Considering that models built in a ANN environment can be somehow unclear for practitioners, especially if they have to use in their applications the same programs used for training the networks, we transferred the trained network to an electronic spreadsheet format. With the models in dynamic environments, either as electronic spreadsheets or as formulas embedded in GIS (Geographic Information System) databases, their application for strategic planing becomes straightforward. We demonstrated it in this study by applying the model to estimate trips for the entire city studied and by simulating the impact that different values of the input variables had on the output of the model. The conclusions are in the last section of the paper, prior to the list of references.

ARTIFICIAL NEURAL NETWORKS IN TRANSPORTATION PLANNING

Artificial Neural Networks are systems made of simple processing units (nodes) arranged in layers that have parallel connections to which weights are associated. Those characteristics make them able to estimate certain mathematical functions. Moreover, according to (1), a MLP (*Multilayer Perceptron*) neural network with two hidden layers is able to approximate any function. As a consequence, depending on the problem considered, the ANN might even have a better performance than traditional statistical techniques (2).

One of the first applications of ANNs in transportation planning was the work of (3), in 1989, which was meant to solve traffic engineering problems. According to (4), the technique has been largely used in transportation engineering in the 1990s for several subject areas such as: driver behavior, parameter estimation, pavement maintenance, vehicle detection classification, traffic pattern analysis, freight operations, traffic forecasting, transport policy and economics, air and maritime transport, submarine vehicles, metro operations and traffic control. The conclusion of (4) was that ANNs show a great promise as a useful tool for analyzing non-linear problems, which are common in the transportation field.

Applications of ANN in transportation planning and research are now widespread in developed countries as shown, for example, in (5). In one of the most recent applications of the technique in transportation planning, the authors compared multiple regression models and ANN for modeling mobility in a Spanish urban area. Both techniques presented similar results, but the ANN models worked with less variables than the multiple regression model.

The ability to work with incomplete data makes the ANNs specially attractive for planning in developing countries, in which some studies are already under way. That was the case, for example, of (6) and (7), in India. In Brazil, (8) used ANN as an alternative to a Logit model, (9) used it to evaluate and to classify transportation projects, and (10) studied the possibility of drawing from the networks the behavior of users when selecting a transportation mode. Another application was described in (11), in which ANN and multiple regression models have been used to identify the impact of transportation accessibility on urban land values. The good performance of the ANNs in all those works suggested that it could also be an alternative to improve some strategic planning methods, particularly the accessibility indicators. One of the problems with traditional accessibility measures is that they usually do not include mobility characteristics of the users, as observed in (12). The importance of investigating transportation accessibility and mobility is stressed in recent works, such as in (13), (14) and (15). While most authors, such as (12), still use a multiple regression approach to model mobility elements and accessibility measures, (13) and (15) suggested to improve the previous approach with ANNs, as further developed in this paper.

STUDY DATA

The main source of data for this study was an Origin-Destination survey carried out in the Brazilian city of Bauru, which had over 300,000 inhabitants. The city has been divided in 98 traffic analysis zones (TAZs) for the survey. Interviews were conducted in a sample of 4,000 households, that is about 4.5% of the total number of households in the city (16). The survey recorded data of 23,314 trips using four transportation modes: car/motorcycle (as driver), car/motorcycle (as passenger), bus, and walk/bicycle. Around 8,000 trips were considered in this study, because of problems in finding the addresses of either origins or destinations. The information about the trips was grouped in two different ways for the analyses conducted here: the total number of trips per household and the total distance traveled per household. These two groups of output data generated eight sets of data when combined with each of the four transportation modes. Total trips (regardless of transportation mode) were used to form two additional sets of data, summing up ten groups of data for the study.

The survey included questions about income and education level, age, occupation, driver's license possession, and gender of the residents. For the purpose of this study, in which the analyses were conducted at the household level, the education variables were the proportion of household members in three levels. The same number of classes was used to classify the respondents by age. A similar approach was used for the driver's license variable, which was taken as the proportion of household members legally allowed to drive. None of the survey questions concerned accessibility, but the knowledge of the exact locations of the households made possible to associate them to accessibility values estimated for the entire city in a GIS-environment.

Spatial Data and Accessibility Measurement

Databases containing the location of TAZs, street centerlines (along with their names), and blocks (along with their numbers) of the city have been built in a Geographic Information System for transportation. GIS tools have been used to find the origin and destination addresses and to calculate the travel distances for all origin-destination pairs. As the automatic address matching process was able to locate 80% of the 4,000 household locations, a manual revision was necessary to locate the other points summing up 99% of households located. The database with street centerlines was used to build a network, that was essential to calculate the accessibility measure. The measure used was the Mean Separation Index (18), which is a measure of the effort of overcoming the spatial separation between zones (Equation 1). When travel distance is used, MSI is nothing but the average travel distance to all possible destinations.

$$A_i = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n} T_{ij} \tag{1}$$

where A_i is the normalized accessibility of points i; T_{ij} is the perceived cost (e.g., travel time) to the traveler between points i and j; and n is the total number of points used in the application.

Smaller values indicate smaller average distances and therefore better accessibility. As a consequence, the households located in the central region of the city have the best accessibility and those located at the periphery, which is usually the region where low-income groups live in Brazilian cities, have the worst accessibility. Following, the accessibility values of the households were joined with the other input and output variables in a same database in order to build exploratory models for an evaluation of the variables.

FORMULATION OF THE ANN MODELS

After the selection of the mobility-related variables from the O-D data and the calculation of travel distances and accessibility values for each household, preliminary ANN models were built. Two output variables have been tested: *the total number of trips per household* and *the total trip length per household*. The input data used in the models were all at the household level: number of residents, income level, proportion of residents in each of the three education levels considered, proportion of residents in each of the three age groups considered, proportion of residents with driver's license, and accessibility values at the trip origins.

The *EasyNN* package, that is a commercial software produced by Stephen Wolstenholme, in England, was used in this study to perform neural network calculations. The neural network model in this case is a multi-layered perceptron (MLP) with up to three hidden layers that uses a backpropagation learning algorithm for establishing the appropriate network weights. For the construction of the ANN models, three sets of data were randomly selected for training (50% of the cases), validation (25% of the cases) and testing (25% of the cases). This division was carried out three times for each of the ten original sets of data, always randomly, generating thirty different sets of data. Several networks were built for different combinations of learning rates, *momentum* and number of neurons in the hidden layer. After a network is trained, its performance can be measured by the comparison of the predicted values and the actual observations. There are several different measures used for that purpose, including relative estimation errors, that may be calculated as follows:

$$\mathrm{RE} = \left| \frac{\mathrm{OBS} - \mathrm{EST}}{\mathrm{OBS}} \right| \times 100 \tag{2}$$

where RE are the relative estimation errors, OBS are the actual observations, and EST are the values estimated by the model. This definition expresses the error as a percentage of the true or observed value.

In general, the performance of the networks using *the total trip length per household* as the output variable was inferior than that of the networks using *the total number of trips per household* as the output variable. Consequently, only the results of the 15 trained networks having *the total number of trips per household* as the output variable were presented here. Table 1 shows these results, as well as the number of nodes in the single hidden layer, the target errors, the mean relative errors, and the R^2 values for each case.

Afterwards, using the approach suggested by (18) the relevance of the input variables was evaluated for each set of data. The method, that has been successfully used in the works of (19) and (20), assumes that W_{ij} (i = 1,...,k; j = 1,...,p) are the weights in the connections of the neurons i of the input layer and j of the hidden layer. Similarly, it assumes that W_{rs} (r = 1,...,p) are the weights in the connections of the neurons r of the hidden layer and s of the output layer. Therefore, the weights W_{rs}, no matter if they are positive or negative, may be added to the weights W_{ij}^{*} using the following expression:

$$W_{ij}^{*} = \left\{ \left| W_{ij} \right| / S_j \right\} \left(\left| W_{rs} \right| \right) \tag{3}$$

where

$$S_j = \sum_{i=1}^{k} \left| W_{ij} \right| . \tag{4}$$

Next, the adjusted weights W_{ij}^{*} for the input nodes are added to each of the hidden layer nodes, and $\sum_{j=1}^{p} w_{ij}^{*}$ is obtained. Finally, the relevance Rv_i of each input variable can be estimated by equation (5). The results, given as percentages for each of the eleven variables considered in the preliminary models are shown in Table 2 for all 15 sets of data presented in Table 1.

$$Rv_i = \frac{\sum_{\mathrm{j=1}}^{\mathrm{p}} w_{ij}^{*}}{\sum_{i=1}^{k} \sum_{j=1}^{p} w_{ij}^{*}} \tag{5}$$

The R^2 values and the relative errors presented in Table 1 reflect the relatively poor performance of the preliminary models, except for the group with the models considering trips of the transportation modes altogether. In that case, a visual analysis of a chart in which predicted values were plotted against actual values showed that there were no points extremely away from the 45 degrees reference line (Figure 1). Consequently, the model with the best performance was tested again after the exclusion of the input variables with low relevance (Table 2). Tests with different combinations of input variables suggested that a model with less input variables would not compromise the performance of the model. Therefore, although some of the input variables related to mobility initially appeared to be more relevant than the accessibility indicator, the final model configuration had only the following inputs: income, household size (number of people per household), and accessibility. After conducting again all steps already described in this section, the performance of the new models with the three sets of data were evaluated, as shown in Table 3. The relevance of the input variables in the new models are shown in Table 4. The values of Table 4 show that, on average, the two mobility-governing variables were both more relevant than the accessibility variable.

It is interesting to notice that although the performance of the models was not improved by the reduction in the number of input variables from 11 to 3, the new conditions have not worsened the results as well. The R^2 values and the relative errors were quite similar in all cases. In other words, the models with 3 input variables are preferred to the models with 11 input variables, especially because the 3 variables considered are easy to obtain. Income and number of people per household, that clearly influence the mobility of people, are collected in the official Census surveys regularly conducted in all Brazilian cities. In addition, the accessibility measure does not require anything but the graphical representation of the street centerlines. Models with the three input variables shown here could be calibrated (or trained, in the ANN terminology) with the occasional O-D data and used to generate a trip potential indicator for strategic planning, as shown in the next sections.

EXPLORING THE BEST MODEL

Accordingly to what has been hitherto discussed, the model selected for further development and application had three input variables (income, household size, and accessibility) and one output variable (total number of trips per household). Having decided that, the next step was to reproduce the trained network in an electronic spreadsheet. The process started with the identification of the activation function adopted in the original NN simulator. In the case of the simulator *EasyNN*, the activation function is the logistic function that can be described by equations (6) and (7), which are applied in the connections between the input layer (in that case, normalized input data) and the hidden layer. Similar equations are applied to the connections between the hidden layer and the output layer in order to find the output values produced by the network.

$$F(u_j) = \frac{1}{1 + e^{-u_j}} \tag{6}$$

$$u_j = a_{0j} + \sum_{i=1}^{I} a_{ij} x_i \tag{7}$$

Where:

a_{ij} = weight associated to the connection from input node *i* to the hidden node *j*

x_i = value of the input node *i*

a_{0j} = Bias of the node *j* in the hidden layer

I = number of input nodes

After the network is trained, the values of weight and bias, respectively associated to each connection and node and required by equations (6) and (7), are provided by the NN simulator, making it simple to replicate the model in an electronic spreadsheet. In the present study, after the trained network was replicated in an electronic spreadsheet, two activities were carried out: the estimation of the trip potential index for the entire city studied, and a general analysis of the impact of each one of the input variables on the output values.

The estimation of the trip potential index for the entire city

The simplified ANN model calibrated in the previous section was then used with data aggregated at the zonal level in order to get the trip potential estimates for the entire urban area. As the available Census data from 1991 were older than the O-D data used in the calibration of the models we decided to use the latter also in the example. As we assumed that the performance of the model has already been tested and accepted based on the relative errors and R^2 values, the difference then was that the input variables were considered as average values per zone, exactly in the same format that they would be found in a Census report. The distribution of the input values throughout the urban area of Bauru can be seen in the thematic maps presented in Figure 2.

From Figure 2(a), for example, it is easy to see that the smaller households are mainly clustered around the city center. In addition, Figure 2(b) shows that the higher average income values are concentrated in the southeastern region of the city. The accessibility distribution in the zones is not presented here, because the pattern is quite simple to grasp. Due to the nature of the indicator applied, households located in the central region of the city have the best accessibility and the accessibility level decreases in concentric rings moving outwards.

The output obtained when the model was fed with input variables aggregated at the zonal level is an overview of the potential number of trips that would likely be generated all over the city. That makes it a powerful and promising tool for strategic planning. Although the results of the model have already been tested in the previous section, it was now interesting to see how the predictions at the zonal level were. From Figure 3, where the average numbers of actual household trips per zone are depicted, no clear spatial pattern can be easily drawn. Central TAZs generated in many cases the same number of trips as did peripheral TAZs. The comparison of parts of Figures 2(b) and 3(a), however, may suggest that the income level really had some influence on the number of trips, as one could anticipate. The southeastern zones, which had the highest average income values had also some of the larger figures of actual average trips, even considering that the accessibility level was not so high in that part of the urban area. This was likely due to the high automobile availability in those households. This may be a clear example of how the accessibility values alone are not good enough to explain actual trip patterns. This example stressed the importance of having models that combine accessibility and mobility factors for the prediction of travel patterns.

The results of the proposed model are shown in Figure 3(b). At first, a visual comparison of the actual trip patterns and the trip potential values shows that they were reasonably similar. To examine it in more detail, the results of a numerical comparison were analyzed and they were also acceptable. Several estimates had a variation of below 20% when compared with the actual trip values. In addition, the variation of the predicted values was below 40% in most zones. Only one zone had an extremely large variation, which may be explained by the fact that the zone was almost uninhabited.

Analyzing the input variables of the model

With the trained network replicated in an electronic spreadsheet, it was easy to perform several simulations with the model in order to see how the output values would be affected by changes in the input values.

Results of this kind of analysis are summarized in Figure 4. When analyzing one input variable, the values of the other two were assumed to be constant and close to the average: R$7,000 as the income per household, 5 persons as the household size, and 7.89 km as the average accessibility.

A detailed analysis of the graphs suggest some aspects that have to be stressed here. In Figure 4(a), for example, a first look suggests that the number of trips increased with the income values. This only happens above a certain value, however, what might indicate that the model should be used with a certain reserve for higher income values. Or it might be an indication that the data used for training the network was not reliable. Particularly in the case of high income values, some respondents may have forgotten certain trips or intentionally distorted income-related information. In the case of Figure 4(b), the observed trend is perfectly aligned with what should be expected: the bigger the household size, the larger the number of trips per household. It is important to stress, however, the fact that the trendline is irregular, what was a likely benefit of the Neural Network modeling approach. In Figure 4(c), the tendency is coherent and regular: higher accessibility values are associated with lower number of trips. It is important to remember that the accessibility values here are nothing but the average distance from any node to all other node in the street network. Therefore, higher values mean longer distances and they are supposed to have a negative effect on the trips.

CONCLUSIONS

The study conducted here showed that Artificial Neural Networks were a good alternative to build the intended models. The best preliminary models using *the total number of trips per household* as the output variable were tested again after the exclusion of the input variables with low relevance. The remaining input variables were: income, household size, and accessibility. The reduction in the number of input variables did not worsened the performance of the models, what makes them more interesting than the more complex, preliminary models.

Although the results obtained in the case study were good enough for strategic planning purposes, changes in the model formulation could further improve its estimations. The accessibility measure used as an input variable, for example, was based only on physical attributes of the network. A gravity-type index would probably give a better picture of the actual accessibility in the urban area. In that case, differences in the attractiveness of the zones would be taken into account. Unfortunately, there were no data available to consider this kind of measure in the present study. Other conditions could be tested in order to improve the performance of the models, for example, filtering the input data to analyze subsets of trips by purpose. Even though, the simplified ANN model seems to be quite useful for strategic planning considering that the input variables left are easy to get, for instance, from Census data.

It is important to emphasize the role that GIS has played in this study. Without it, the acquisition of the spatial attributes of the O-D survey and the estimation of the accessibility values would be very difficult and time-consuming. In addition, its mapping capabilities made easy to understand and to compare the results of the model application with the actual trip values.

The ANN *software* used is user-friendly and it is an affordable option even for developing countries standards. Other positive outcome of the ANN approach is the possibility of getting to know the relevance of the input variables while building the models. This made very easy to redesign them by the exclusion of less relevant input variables. In the case studied here, the selection of the most relevant variables stressed the assumption that accessibility and mobility should be examined together in transportation planning analyses. In general, the approach developed here seems to be a step forward in relation to traditional accessibility models and it may be a useful tool for urban and transportation planners and decision-makers. The approach makes clear that urban citizens (particularly those in developing countries) need not only physical accessibility, but also better mobility conditions.

Based on the results presented and analyzed in this study, the use of electronic spreadsheets to replicate the trained network seems to be a simple and efficient way to bring the ANN modelling capabilities closer to practitioners and decision makers. The sort of sensitivity analises conducted even might even help to understand the logic behind the model or to detect inconsistencies, as that observed here in the case of the input variable income, for example.

Acknowledgments

The authors would like to express their gratitude to the Brazilian agencies FAPESP (Foundation for the Promotion of Science of the State of São Paulo), CAPES (Post-Graduate Federal Agency), and CNPq

(Brazilian National Council for Scientific and Technological Development), which have supported this research in different phases.

References

1. Cybenko, G. Continuous valued neural networks with two hidden layers are sufficient. *Technical Report*, Department of Computer Science, Tufts University, Medford, MA, USA, 1988.
2. Falas, T. Neural networks in empirical accounting research: an alternative to statistical models. *Neural Network World,* Vol. 5, No. 4, 1995, pp. 419-432.
3. Nakatsuji, T., and T. Kaku. Application of Neural Network Models to Traffic Engineering Problems. *Proceedings of the Infrastructure Planning*, Vol.12. pp. 297-304, 1989.
4. Dougherty, M. A Review of Neural Networks Applied to Transport. *Transportation Research*, Vol. 3B, No. 4, 1995, pp. 247-260.
5. Sánchez, T.R., and J.R.M. Folgado. Modelos de Generación de Viajes. Análisis Comparativo de la Utilización de Redes Neuronales y Regresión Lineal Múltiple. In: Colomer, J.V., and A. García. (Ed.) *Proceedings of the IV Congreso de Ingeniería del Transporte*, Valencia, Espanha. Vol. I, pp. 189-196, June 2000.
6. Rao, C.V.S., K.V.K. Rao, and P.K. Sikdar. Micro-simulation Model for Service Employment Location Using ANN. In: Rizzi, P. (Ed.) *Proceedings of the 6th International Conference Computers in Urban Planning and Urban Management* (in CD-ROM), FrancoAngeli, Venice, Italy, September 1999.
7. Rao, C.V.S., and K.V.K. Rao. A Micro-analytic Land Use Model using Artificial Neural Networks. *Proceedings of the 9th World Conference on Transport Research* (in CD-ROM), World Conference on Transport Research Society, Seoul, Korea, July 2001.
8. Carvalho, M.C.M. Previsão de Demanda de Transportes: uma Abordagem Não-linear Usando Redes Neuronais. *Proceedings of the XIII ANPET Conference*, ANPET, São Carlos, Brazil. Vol. I, pp. 187-200, November 1999.
9. Furtado, A.N.D. Uma Nova Abordagem na Avaliação de Projetos de Transportes: o Uso de Redes Neurais Artificiais como Técnica para Avaliar e Ordenar Alternativas. Unpublished Ph.D.

Dissertation, Department of Transportation, São Carlos School of Engineering, University of São Paulo, 1998.

10. Wermersch, F.G., and E. Kawamoto. Uso de Redes Neurais Artificiais para Caracterização do Comportamento de Escolha do Modo de Viagem. *Proceedings of the XIII ANPET Conference*, ANPET, São Carlos, Brazil. Vol. I, pp. 31-34, November 1999.
11. Brondino, N.C.M., and A.N.R. Silva. Combining Artificial Neural Networks and GIS for Land Valuation Purposes. In: Rizzi, P. (Ed.) *Proceedings of the 6th International Conference Computers in Urban Planning and Urban Management* (in CD-ROM), FrancoAngeli, Venice, Italy, September 1999.
12. Tagore, M.R., and P.K. Sikdar. A New Accessibility Measure Accounting Mobility Parameters. Paper presented at the *7th World Conference on Transport Research.* Book of Abstracts, University of New South Wales, Sydney, Australia, 1995, pp. 10-7.
13. Raia Jr., A.A. Acessibilidade e Mobilidade na Estimativa de um Índice de Potencial de Viagens Utilizando Redes Neurais Artificiais, Unpublished Ph.D. Dissertation, Department of Transportation, São Carlos School of Engineering, University of São Paulo, 2000.
14. Zuidgeest M., E. Akinyemi, and M. van Maarseveen. Challenging Sustainable Accessibility: Accessibility and Mobility Concepts in Dynamic Urban Environments. *Proceedings of the 9th World Conference on Transport Research* (in CD-ROM), World Conference on Transport Research Society, Seoul, Korea, July 2001.
15. Olaru D., and V. Dragu. Accessibility in Urban Areas. Case Study: Perth City. *Proceedings of the 9th World Conference on Transport Research* (in CD-ROM), World Conference on Transport Research Society, Seoul, Korea, July 2001.
16. Reyes, D.E.L. Um procedimento para determinação da matriz origem-destino para diferentes modos: método indireto baseado em modelo de escolha discreta. Unpublished Ph.D. Dissertation, Department of Transportation, São Carlos School of Engineering, University of São Paulo, 1999.
17. Allen, W.B., D.Liu, and S. Singer. Accessibility Measures of U.S. Metropolitan Areas. *Transportation Research,* Vol. 27B, No. 6, 1993, pp. 439-449.
18. Garson, D.G. Interpreting Neural-Network Connection Weights. *AI Expert*, April 1991. pp. 47-51.

19. Nath, R., B. Rajagolapalan, and R. Ryker. Determining the Saliency of Input Variables in Neural Network Classifiers. *Computers & Operational Research*, Vol. 24, No.8, 1997, pp. 767-773.

20. Brondino, N.C.M. Estudo da Influência da Acessibilidade na Valorização de Lotes Urbanos Através do Uso de Redes Neurais. Unpublished Ph.D. Dissertation, Department of Transportation, São Carlos School of Engineering, University of São Paulo, 1999.

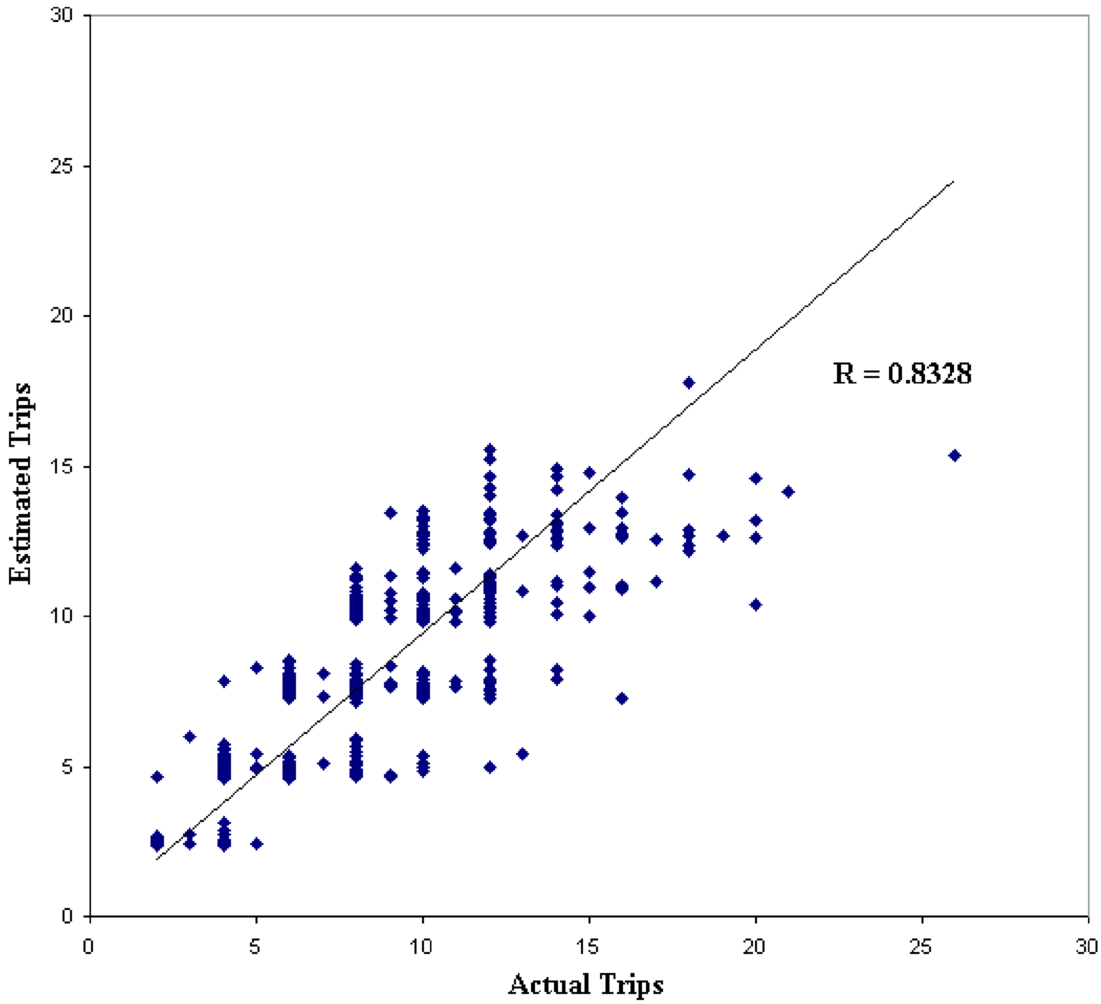
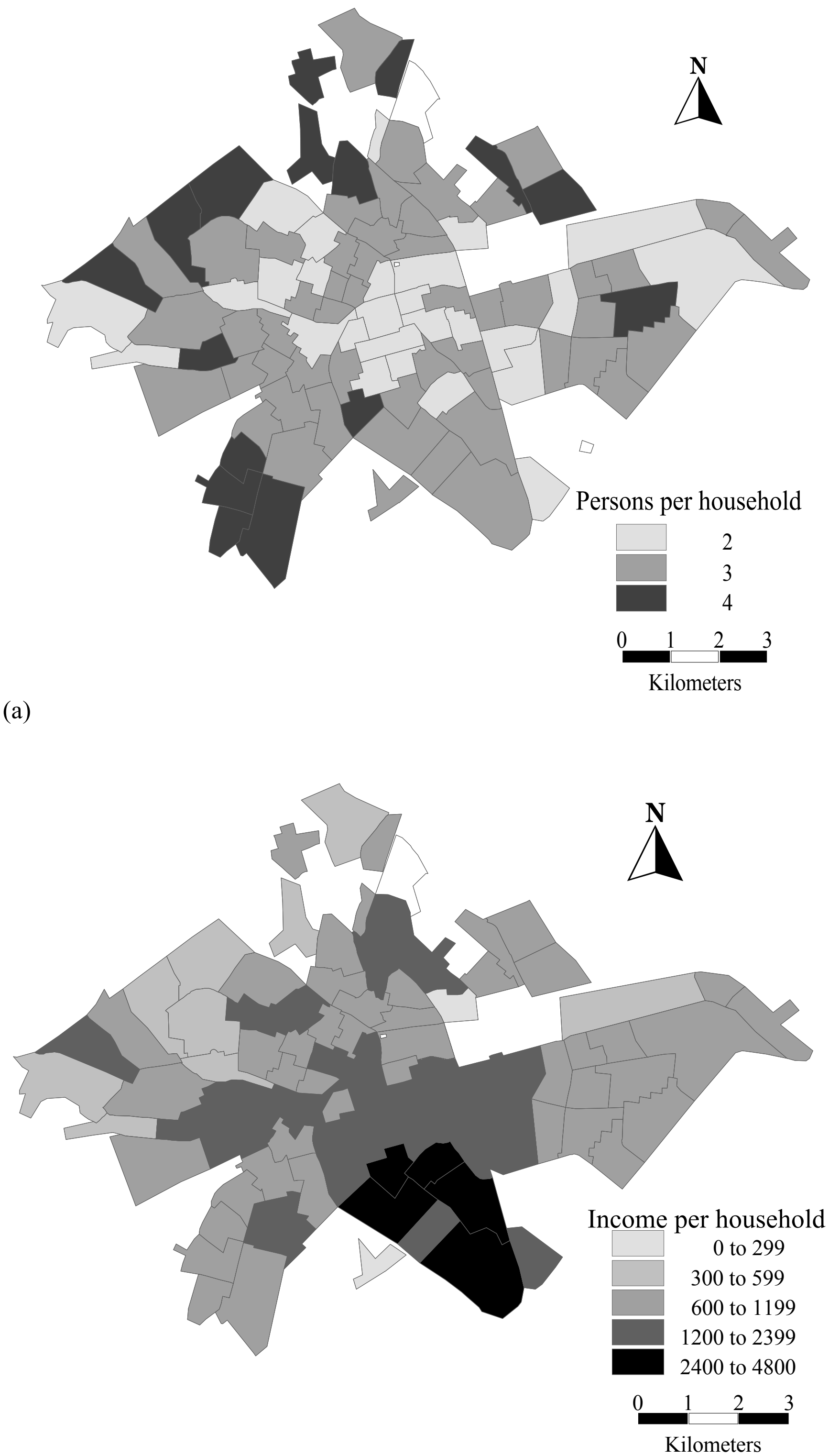
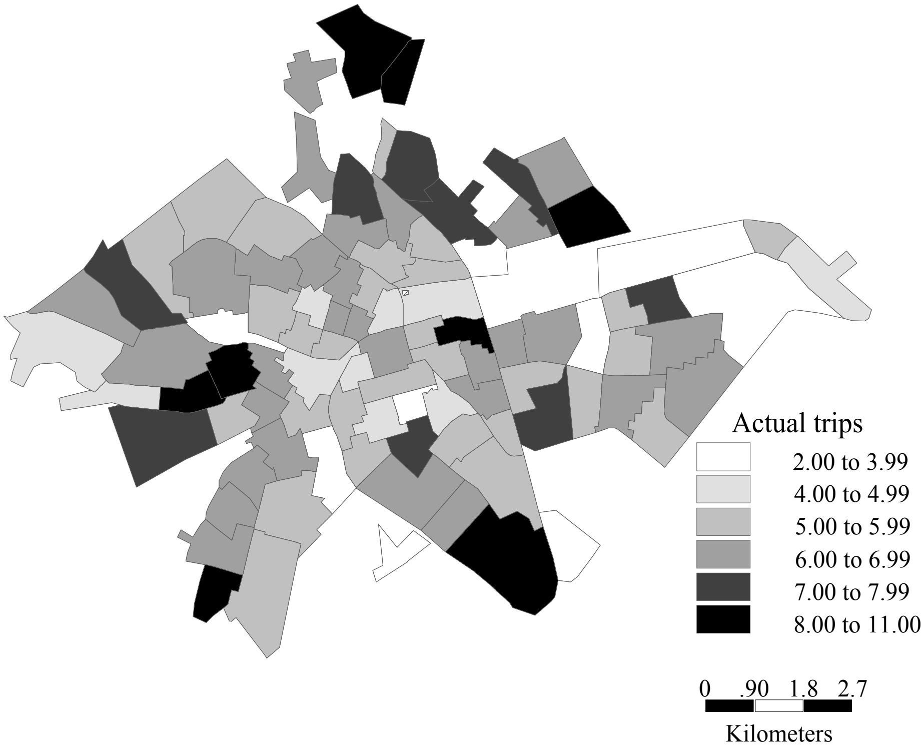


FIGURE 1 A comparison between actual and estimated trips

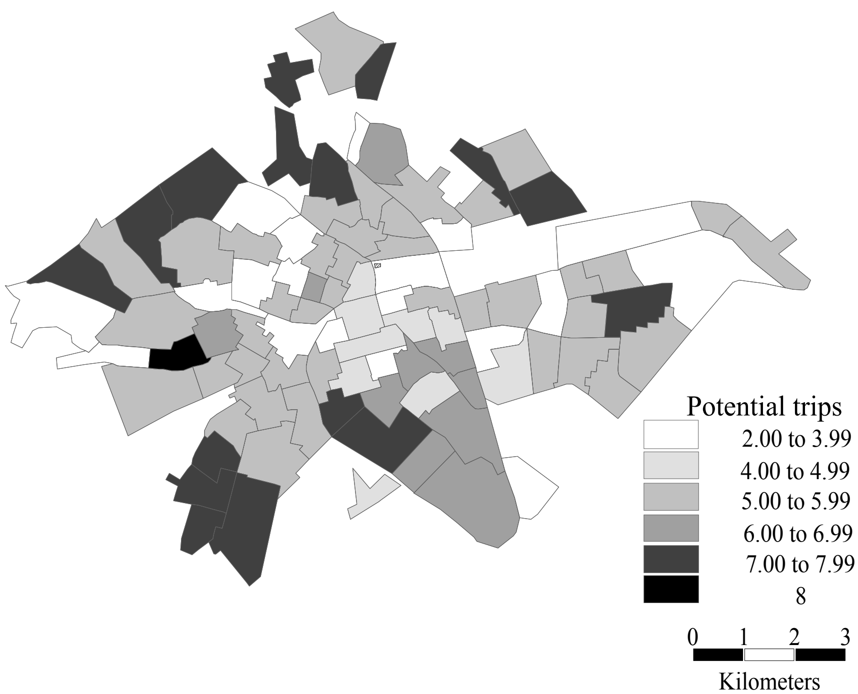


(b)

FIGURE 2 Average household size (number of persons per household) and income (R$) per TAZ

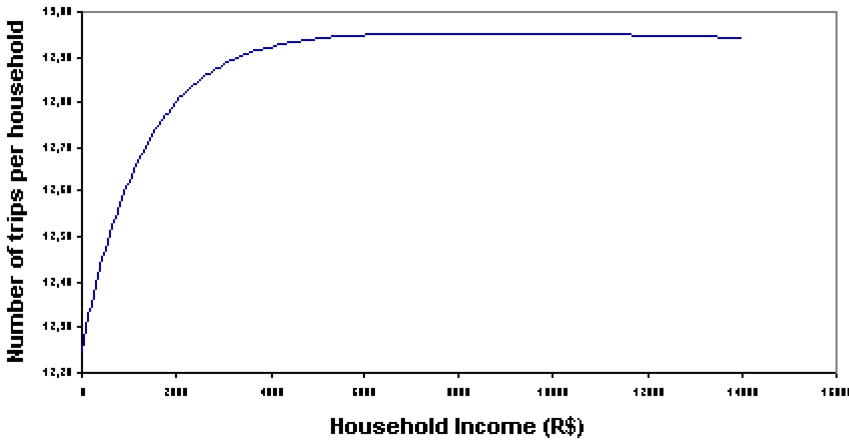


(a)

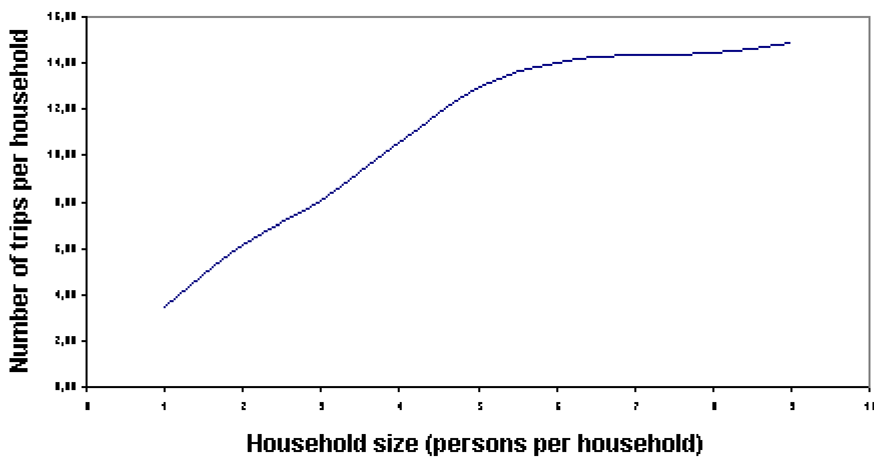


(b)

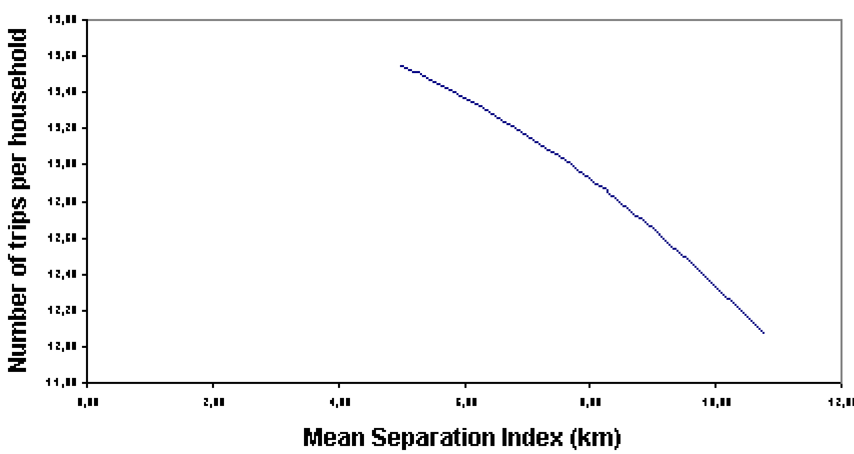
FIGURE 3 Average number of actual (a) and potential (b) trips per household in each TAZ



(a)



(b)



(c)

FIGURE 4 Influence of the input variables on the number of trips per household: (a) income, (b) household size, and (c) accessibility level

TABLE 1 Results of the Preliminary Models Having *the Total Number of Trips per Household* as the Output Variable

Transportation modes	Case	Nodes in the hidden layer	Target error	Mean relative errors	Mean relative error per mode	R^2 (%)
All	1	6	0.03	23.77		68.30
	2	6	0.03	20.36	22.01	65.64
	3	6	0.03	21.90		64.99
Car – Driver	1	9	0.03	34.35		39.28
	2	9	0.03	37.05	35.70	29.94
	3	9	0.03	35.71		28.46
Car – Passenger	1	7	0.03	28.59		25.77
	2	7	0.03	32.42	32.39	16.57
	3	7	0.03	36.17		12.51
Bus	1	6	0.03	20.18		56.03
	2	6	0.03	18.57	20.89	52.85
	3	6	0.03	23.93		60.21
Non-motorized	1	4	0.03	30.68		30.08
	2	4	0.03	29.10	27.95	35.90
	3	4	0.03	24.07		44.15

TABLE 2 Relevance of the Input Variables in the Preliminary Models Having *the Total Number of Trips per Household* as the Output Variable

Transportation Modes	Cases	Income	Household Size	Education level			Age			Driver's License	Accessibility
				A*	B**	C***	13 or younger	14-60	60 or older		
All	1	3.39	38.97	2.33	5.82	6.02	10.85	10.30	11.58	1.54	9.17
	2	8.79	39.97	3.89	16.81	5.12	13.51	2.08	2.39	4.23	3.20
	3	4.82	35.57	6.24	8.62	7.32	17.79	4.83	4.58	3.62	6.61
Car – Driver	1	11.89	12.04	11.28	7.94	6.53	13.86	5.89	9.23	11.84	9.49
	2	21.06	19.16	3.43	10.69	6.39	7.01	5.05	9.90	7.18	10.13
	3	14.27	16.05	8.39	11.02	13.27	12.24	3.17	3.17	9.51	8.91
Car – Passenger	1	12.75	45.12	6.96	2.70	4.81	5.29	3.12	1.67	9.00	8.58
	2	7.59	31.45	4.81	5.00	10.81	7.37	6.57	7.16	7.40	11.83
	3	9.33	13.84	6.39	4.28	10.85	7.18	11.18	14.15	10.04	12.76
Bus	1	7.69	24.68	14.73	5.37	13.00	8.92	10.26	4.32	2.97	8.07
	2	7.91	18.61	16.60	6.94	12.55	11.93	9.95	5.58	4.73	5.20
	3	12.92	32.29	5.11	6.31	11.49	8.30	6.56	5.78	7.65	3.59
Non-motorized	1	9.52	38.09	4.07	5.79	5.74	13.33	5.27	4.40	7.99	5.79
	2	6.89	36.92	8.45	7.46	15.42	5.49	5.05	2.74	7.98	3.60
	3	11.59	63.49	1.88	0.75	1.97	3.00	3.38	5.01	3.64	5.29
Average		10.03	31.08	6.97	7.03	8.75	9.74	6.18	6.11	6.62	7.48

* Up to 8 years of school
** Finished High school or attending University
*** Finished University

TABLE 3 Results of the Simplified Models Having *the Total Number of Trips per Household* as the Output Variable

Transportation modes	Cases	Nodes in the hidden layer	Target error	Mean relative errors	Mean relative error per mode	R^2 (%)
All	1	4	0.03	22.70		69.32
	2	4	0.03	20.39	21.65	65.42
	3	4	0.03	21.87		65.25

TABLE 4 Relevance of the Simplified Models Having *the Total Number of Trips per Household* as the Output Variable

Transportation modes	Case	Income level	Household size	Accessibility
All	1	29.17	64.38	6.46
	2	14.40	68.66	16.94
	3	17.63	70.52	11.85
Average		20.40	67.85	11.75