Neuro-fuzzy Modelling of Workers Trip Production

Meysam Ahmadpour[1], Dr Wen Long Yue[2], Morteza Mohammadzaheri[3]
[1] Postgraduate student, ISST – Transport Systems, University of South Australia
[2] Program Director, Senior Lecturer, ISST – Transport Systems, University of South Australia
[3] PhD candidate, Dept. of Mechanical Engineering Adelaide University

ABSTRACT

This paper attempts to introduce the application of Neuro fuzzy techniques for fulltime worker trip production estimations in Adelaide metropolitan area using the household/person characteristics such as age, vehicle ownership and distance from CBD. In the last 30 years, several linear regression models have been developed for this purpose. These models' linear structure does not seem suitable to predict highly nonlinear behaviour of urban transport systems. Consequently, intelligent modelling methods, as powerful nonlinear tools, have attracted much attention in the prediction of trip productions. In 1993, fuzzy logic and artificial neural networks were combined and neuro-fuzzy technique was emerged to model engineering systems. Since then this technique has been improved drastically and utilized to model a wide variety of complicated engineering systems. In this research, the aforementioned method is employed for modelling person/worker trip productions. After subtractive clustering, a meaningful relation between distance of residence from CBD area and workers' trip productions was not observed in this research. The modelling was accomplished with and without this factor and this view was justified. At the end, a fuzzy inference system was achieved which explains persons behaviour with a reasonable error range.

Key words: Neuro fuzzy, Modelling, Trip generation

1. INTRODUCTION

Since the recognition of transport systems as one of the major aspects of the economic growth (Taaffe, Morrill & Gould 1963), many countries have devoted an increasing amount of budget and attention on the development of their transport systems. This improvement is mainly concerned with the evaluation of the existing infrastructure and services as well as the planning process including travel demand modelling.

The fundamentals of transport modelling were developed in the U.S.A in 1950's, and after the initial 20 years of important theoretical developments, then the development of the main stream techniques has been evolutionary rather than revolutionary (Bates 2000) Recently, economic theories have been employed in modelling techniques and at the same time computing power has increased significantly. Consequently, the scale and details of the problems to be analysed by modelling techniques have expanded.

In transport planning activities, many models have become available in forecasting travel demand whereas the history of travel demand modelling has been largely dominated by the introduction of the so called four stage model (McNally 2000). The Four Stage Model (FSM) is one of the strongest tools available for transport modellers in order to forecast the future demand and the performance of a transport system. However according to McNally (2000), the FSM basically functions best if it is used to evaluate large-scale infrastructure rather than more subtle and complex policies involving the management and control of existing infrastructure or the introduction of policies that directly influence travel behaviour.

In the FSM, likewise many other transport planning processes, the generation of a travel demand-forecasting model, also known as origin-destination (O-D) matrices estimation, should be of the main interest for the planner before conducting any transport studies. In this model the generation of O-D matrix is achieved through the two initial steps of Trip Generation (and production) and Trip Distribution while the two last steps , known as Modal Split and Trip Assignment, are generally applied to predict the road user's approach on what mode of transport to use and what route of travel to take (McNally 2000)

Given the importance of the O-D matrices, several models have been proposed to model each of the two first stages, in particular, trip generation stage. The importance of the trip generation stage mainly comes from its nature that attempts to define the magnitude of total daily travel in the model system at household or zonal level for various trip purposes (McNally 2000). Since travel demand and perhaps trip generation behaviours are mainly related to decision-making process, most research works in this area have tried to focus on the behaviour of each household/individual (at disaggregate level) rather than average individual/zonal (Yaldi 2008).Thus in many cases the decision to choose between aggregate and disaggregate approaches seems to be one of the most critical issues in modelling (Ortuzar, Dios & Willumsen 1994). However, there are some successful approaches in which disaggregated sub-models could be applied to generate overall aggregated models (Gu, Holyoak & Benham 2007).

For this purpose, several mathematical models (mostly regression models), based on the area of study, have been introduced to model the trip generation behaviour for individual households rather than individual persons. Supernak, Talvitie, and DeJohn (1983) tried to apply a person category trip-generation model in which a homogeneous group of persons was used as an analysis unit instead of household units. In this study the most important descriptors of a person's mobility were found to be age, employment status and automobile availability.

Since most of early trip generation models were based on mathematical models and given the complicated nature of human based activities and the inadequacy of common mathematical models (in particular regression models) to deal with most human based activities, fuzzy inference systems (FIS) based on fuzzy logic theory can be employed in the processes (Teodorvic 1998).

The concept of this approach (fuzzy modelling approach) is based on the fact that, travel demand modelling is related to the decisions made by the humans. These decisions are largely made based on linguistic information. Linguistic terms (which are the basis of fuzzy inference systems) are largely characterized by uncertainty, ambiguity and imprecision that common mathematical models including regression models in particular simply fail to address (Zimmermann 1996; Teodorovic and Vukadinovic 1998).

There are many examples of using fuzzy logic in different stages of travel modelling including four step model. They are listed as follows:

- The trip generation problem solved using fuzzy logic by KalicÂ and TeodorovicÂ (1997);
- Trip distribution problem solved using fuzzy logic by KalicÂ and TeodorovicÂ(1997);
- Mode choice problem solved using fuzzy logic by KalicÂ and TeodorovicÂ (1996) are only a handful number of many successful studies carried out using fuzzy logic.

In this paper the modelling of worker trip production is addressed using neuro-fuzzy technique.

2. FUZZY LOGIC AND NEURO-FUZZY NETWORKS

As the main difference with binary logic, in fuzzy logic, fuzzy or linguistic values are designated to variables (i.e. temperature, velocity), like cold, fast. These fuzzy values are defined using membership functions. A 'membership function' is a function which receives the crisp (numeric) value of a variable (i.e. 25ºC) and returns another number in the range of [0, 1] namely 'membership grade'. This membership grade determines that how much the numeric value (i.e. 25ºC) accords to membership function (i.e. cold). For instance, the membership grade of 25ºC for fuzzy value of cold is 0.1 (according a specific definition of cold). The definition of membership functions is usually done by the designer or through a training process.

A fuzzy inference system (FIS) is a set of fuzzy rules. If all the variables both in *antecedent* and *consequent* of the fuzzy rules are fuzzy, the FIS is called "Mamdani Type". If the *consequents* of the fuzzy rules contain crisp (mathematical) relations or values, the FIS is known as "Sugeno-type". Figure 1 shows a structure of a linear Sugeno type fuzzy inference system (FIS). In this structure, *antecedents* of rules contain fuzzy sets (membership functions) and *consequents* are first order polynomial (a linear crisp function). The neuro–fuzzy network structure is presented in Figure 2. In neuro-fuzzy (neural network) structure shown in figure 2, the membership functions have been replaced by activation functions of layer 1. Layers 3, 4 and 5 play the role of averaging function in figure 1 (FIS), and layer 2 produces the weights of the rules.

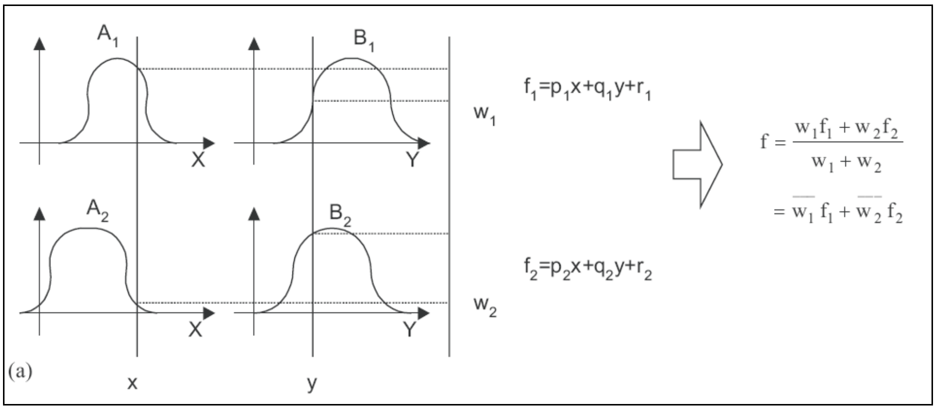

Figure 1: A Sugeno-type fuzzy inference system

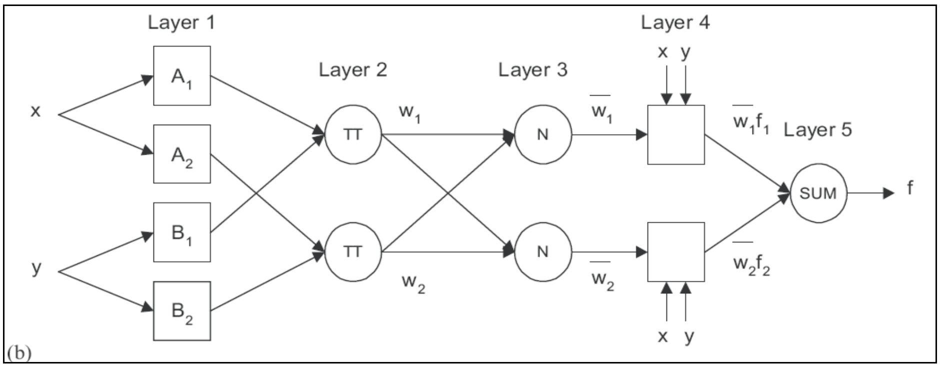

Figure 2: A Sugeno-type neuro-fuzzy network

The training procedure involves both gradient error back propagation (to adjust membership function coefficients) and LSE, to adjust linear output parameters (Jang, Sun & Mizutani 2006; Ghaffari, Mehrabian & Mohammadzaheri 2007; Jang 2007).

In fuzzy inference systems, if all the inputs have the identical number of fuzzy values, the number of fuzzy rules equals the number of fuzzy values of each input (e.g. 3) powered by number of inputs, if all the input space is covered in FIS. Therefore, sometimes, so many rules are needed to cover all the input space. Training such FIS's is too time consuming and practically impossible. In order to reduce the number of fuzzy rules with minimum accuracy loss, a method namely *subtractive clustering* is applied (Jang 2006; Ghaffari, Mehrabian & Mohammadzaheri 2007; Jang 2007). In this method, rules with most probable antecedents in the recorded data of actual system are selected.

In this method, for each set of input data ($x_i, i = 1,.., n$) in an *m*-dimensional space, a density value is calculated as below (the neuro-fuzzy model has *m* inputs):

$$D_j = \sum_{j=1}^{n} \exp\left(-\frac{\left\| x_i - x_j \right\|^2}{(r_a/2)^2} \right). \tag{1}$$

Where: $\|x_i - x_j\| = \sqrt{\sum_{k=1}^{m} (x_i^k - x_j^k)^2}$ (distance) (2)

r_a = range of influence (a positive constant).

Then the point with highest density is defined as the centre of the first cluster. The centre of the first cluster is named *C1*, and its density is named D_{C1}.A cluster is a hyper sphere with the centre of *Ci* and radius of r_a in an *m*-dimensional space. Later on, each cluster is used as the antecedent of a rule. After the definition of the centre of the first cluster, the density of other points is re-defined as:

$$D_i = D_i - D_{C1} \exp\left(- \frac{\|x_i - x_{C1}\|^2}{(r_b/2)^2} \right). \quad (3)$$

Where, r_b = Squash factor

If the redefined density of any point exceeds "accept ratio" it is defined as the centre of a cluster, and then, the density of other points redefined again:

$$D_i = D_i - \sum_{l=1}^{p} D_{Cj} \exp\left(- \frac{\|x_i - x_{Cj}\|^2}{(r_b/2)^2} \right) \quad (4)$$

Where *p*= the number of already defined clusters, D_i = the initially calculated density

After any stage of re-definition, the density of the centres of previously defined clusters are re-calculated. If their density yields lower than "reject ratio", those clusters are eliminated. This process continues till the clusters do not change in two sequential stages. The model derived from subtractive clustering is used as the initial neuro-fuzzy model for training.

3. PROBLEM STATEMENT

Since the importance of the trip generation as the fundamental stage for the FSM, several mathematical models have been introduced to model the trip generation at household levels. Additionally the trip produced by household units may be better modelled if the model is broken down to more fundamental units as persons. (Supernak, Talvitie & DeJohn 1983)

In this case many studies up to date have tried to apply travel demand models in order to predict trip generation behaviour which involves Work to Home (WTH) or Home to Work (HTW) trips. As it is obvious these type of trips are generally produced by working members of households. The main factor which causes the travel demand modeller to pay much attention to these type of trips is the fact that these type of trips compose the largest proportion of trips produced in the rush-hour and the number of trips produced by working members of the house within the working days (Benham 2001). However prediction of trips generated by working members of household involves a great deal of complexity as it includes many socioeconomic factors including household income, age, working status and so on. (Mannering, Kilareski & Washburn 2005).

In this study, the data gathered from 1999 Metropolitan Adelaide Household Travel Survey was used. The survey was in the form of face to face interview over a period of more than 5 months in 1999 in which a random sample of 9000 households (2 percent of all households in the Adelaide Statistical Division) were drawn by the State Electoral Commission in equal proportions from each electorate in the ASD (Benham 2001). At the end of the survey, the information of a total number of 5615 household was available. In this survey, the participants were interviewed for two consecutive days with regards to household characteristics, person characteristics, vehicle characteristics, trip characteristics and the details used in this paper. At the end of this survey details for 27076 people were achieved. Within this information, the data of 2112 full time-workers were available for at least one of the days of Tuesday and Wednesday. The data was adopted to model all the trips produced by working members of the surveyed household on the days Tuesday and Wednesday for the survey duration. In the modelling process and based on the sufficiency of the data sets to carry out training and checking steps, 312 data sets were used as the checking data and 1800 data sets were used for training purpose. In the data processing stage, the data associated with full time working members of the household were extracted in regards to the following details of each full time working member:

- *Age:* the age of the full time working member of the household
- *Distance to the CBD*: the actual distance of the household to the CBD of Adelaide in kilometres
- *Number of Vehicle available to the household*: the actual number of the vehicle available to the household regardless of the fact that whether the person is the driver of the vehicle or not
- *Number of vehicle available to the Person*: the number of vehicle that the person is stated as its main driver
- *Number of trips*: the total number of trips made by the person of all types and regardless of mode of transport and its nature.

The above data were chosen based on data availability and their relation with two important factors of the person status, household location status and the financial status (of both households and individuals). In this research, it is also tried to understand the influence of each of the aforementioned factors on the trips produced by the working member of the household interviewed on the days Tuesday and Wednesday of the survey. Additionally, a predictive fuzzy model was conducted to model the trip production behaviour of the fulltime working members of the household.

4. PROBLEM SOLUTION/MODELIING

In the modelling process the application of Sugeno-Type Fuzzy Inference system was used. Gaussian membership functions were selected to be used in the antecedent (premise) and linear relations were employed in the consequent part of the rules. Also "AND" function was product and weighted average was used to return the final result of FIS.

In this paper the subtractive clustering information was utilized to generate a Sugeno-type fuzzy inference system. First, the type of membership functions of the input variables was defined, then subtractive clustering find their parameters so that the training data was covered appropriately with a relatively low number of rules. The following clustering parameters were used in this research: Range of Influence = 0.5, Squash Factor =1, Accept Ratio =0.1 and Reject Ratio = 0.01.

As a result of the aforementioned process, four fuzzy rules were assigned to the fuzzy inference system. These fuzzy rules include four membership functions for Age, one membership function for Distance to CBD, four membership functions for Vehicles available to Household, and four membership functions for Vehicle available to Person.

At the final step it was concluded that the Distance to CBD factor was of the least possible influence on the trips produced by full time working persons. This fact was completely justified when this factor was excluded from the input sets and the model was trained without participation of this factor.

5. SIMULATION RESULTS AND DISSCUSSIONS

After training, a FIS was achieved with four rules and the following inputs to the final FIS:

- *Age:* u_1
- *Number of Vehicle available to the household*: u_2
- *Number of vehicle available to the Person*: u_3
- *Number of trips*: *y*

A typical fuzzy rule is expressed in this form:

- R_i*:* if u_1 is ${}_iA_1$ and u_2 is ${}_iA_2$ and u_3 is ${}_iA_3$ and u_4 is ${}_iA_4$ then $y = p_i u_1 + q_i u_2 + r_i u_3 + s_i$

Where R_i is the i th rule of the FIS and ${}_iA_j$ is the membership function of the j th input in the i th rule.

- p_i , q_i , r_i and s_i are the coefficients of consequent linear relation of the i th rule.

Equation 5 returns a membership grade of ${}_iA_j$:

$$\mu_{{}_iA_j}(x) = \exp[-\frac{(x - {}_ic_j)^2}{2\,{}_i\sigma_j^{\,2}}] \qquad (5)$$

Where c is the centre of the membership function and σ represents the support of the membership function (Jang, Sun & Mizutani 2006). Table 1 represents the values of the aforementioned parameters; also, figure 3 shows a schematic of the achieved Sugeno-type fuzzy inference system.

Table 1 : Achieved values for the Gaussian membership functions

	${}_ic_1$	${}_ic_2$	${}_ic_3$	${}_i\sigma_1$	${}_i\sigma_2$	${}_i\sigma_3$	p_i	q_i	r_i	s_i
R_1 (i=1)	40	2.124	.910	14.32	1.708	.284	0.09043	-0.6061	5.646	-2.893
R_2 (i=2)	1.997	2.06	.857	14.32	1.476	.6997	0.002725	0.02278	6.026	-0.08231
R_3 (i=3)	82.01	1.956	.991	14.31	1.277	.8538	0.01069	0.03946	2.65	-0.7097
R_4 (i=4)	22.01	4.792	2.231	14.33	1.792	.6175	0.04643	0.09753	-2.412	7.077

According to the results of this modelling process, the mean of absolute error is 2.0441 trips when the distance to CBD is considered (above FIS) and 2.0805 trips when the distance to CBD is not considered (with four fuzzy rules).

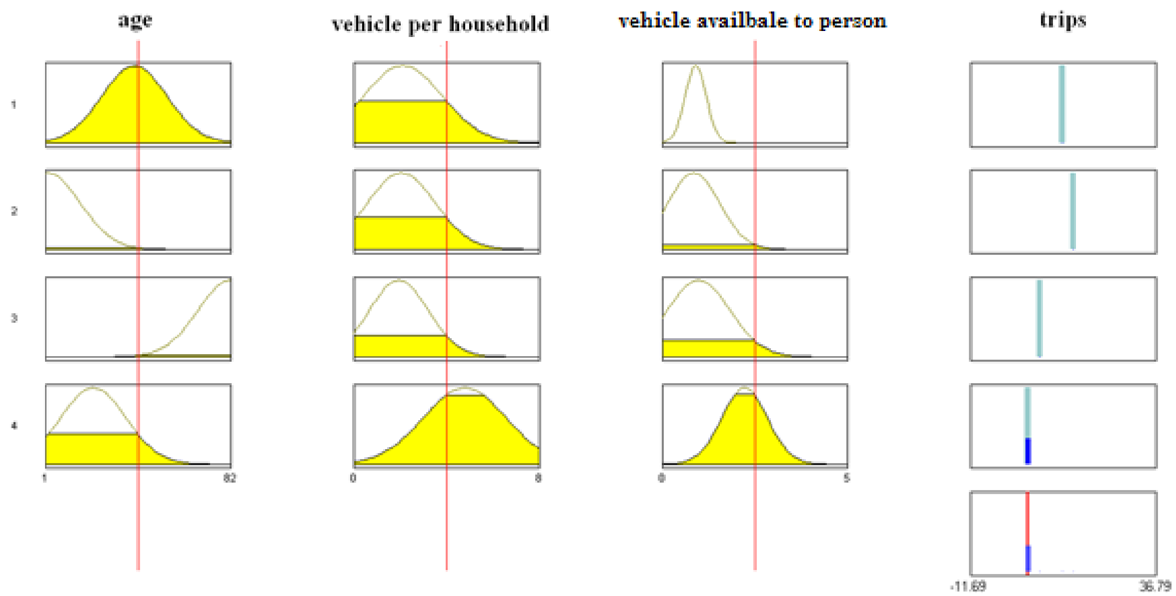

Figure 3: Sugeno-type fuzzy inference system achieved through modelling process

In order to discuss the result of this modelling practice it can be pointed out that the diversity of the membership functions designated to a variable in fuzzy rules indicates the significance of its effect on the output of the model (trips).

In this modelling practice, the household distance to CBD resulted in only one membership function for all four fuzzy rules which was interpreted as its very low influence on the models outcome. Therefore distance to CBD factor could be excluded from the modelling process.

As shown in Fig.3, age has the most important effect on the number of trips produced by full-time workers and the two other factors of Vehicle per household and Vehicle per person can be considered influential factors on trip making behaviour of full-time workers

6. CONCLUSION

In this paper, neuro-fuzzy method is employed for travel demand modelling. The achieved fuzzy inference system (including linguistic values) not only can model the system appropriately (similar to intelligent black-box models) but also can show the effect of inputs on the output qualitatively. For example, in this research, subtractive clustering was used to generate the antecedents of the rules and show the effect of inputs on the output (by assigning more fuzzy membership functions to more effective inputs). In this modelling practice, Subtractive clustering did not show a meaningful relation between the distance of residence to CBD and the trip production by full time workers, so this factor was eliminated. In the end, a four-rule Sugeno-type fuzzy inference system was generated, having 48 parameters , to predict full time workers trip production based on their age, and household/person vehicle ownership.

It is also concluded that the main advantage of neuro-fuzzy technique is that both human knowledge in the form of linguistic terms and input-output data can be utilized in modelling, and the model can be highly nonlinear from mathematical aspect. However, neuro-fuzzy modelling is not applicable in the absence of input-output data.

REFERENCES

Bates, J (2000), 'History of Demand Modelling', in *Handbook of transport modelling*, eds. D A, Hensher and KJ Button, Emerald Group Publishing, New York, pp. 11-32.

Benham, G (2001), *Processing of 1999 Metropolitan Adelaide Household Travel Survey*, PKK Environmnet and Infrastructure Pty Ltd, Adelaide.

Ghaffari, A, Mehrabian, AR and Mohammadzaheri, M (2007), 'Identification and Control of Power Plant De-Super Heater Using Soft Computing Techniques', *Engineering Applications of Artificial Intelligence,Special Issue in Applications of A.I. in Process Engineering,* vol. 20, no. 2, pp. 273-278.

Gu, A, Holyoak, N and Benham, G (2007), *MASTEM:Base Model Report*, Parsons Brinckerhoff Australia Pty Limited, Adelaide.

Jang, JR 2007, *Fuzzy Logic Toolbox User's Guide*, The MathWorks.

Jang, JR, Sun, C and Mizutani, E (2006), *Neuro-Fuzzy and Soft Computing*, Prentice-Hall of India, New Delhi.

KalicÂ , M and TeodorovicÂ , D (1996), *Solving the trip distribution problem by fuzzy rules generated by learning from examples*, Zlatibor,Yugoslavia.

KalicÂ , M and TeodorovicÂ , D (1997), *A soft computing approach to trip generation modeling*, Budva, Yugoslavia.

McNally, M (2000), 'THE FOUR-STEP MODEL', in *Handbook of transport modelling*, eds. D A, Hensher and KJ Button, Emerald Group Publishing, New York, pp. 35-51.

Mannering, F, P.Kilareski, W and Washburn, S (2005), *Principles of Highway Engineering and Traffic Analysis*, 3 edn, John Wiley and sons, Inc, United States of America.

Ortuzar, J, Dios, Jd and Willumsen, L (1994), *Modelling transport*, 2 edn, John Wiley and Sons, West Sussex, England.

Supernak, J, Talvitie , A and DeJohn, A (1983), 'Person-Caregory Trip-Generation Model', *Transportation Research Record 944, Transportation Forecasting: Analysis and Quantitative Methods*, no. 944, pp. 74-83.

Taaffe, EJ, Morrill, RL and Gould, PR (1963), 'Transport Expansion in Underdeveloped Countries: A Comparative Analysis', *Geographical Review,* vol. 54, no. 4, pp. 503-529.

Teodorovic, D and Vukadinovic, K (1998), *Traffic Control and Transport Planning: A Fuzzy Sets and Neural Networks Approach*, Kluwer Academic Publisher, Massachusetts, USA.

Yaldi, G (2008), 'Developing Fuzzy-Neuro Model for Travel Demand Modelling', Transport System Centre, PhD Research Proposal thesis, University of South Australia.

Zimmermann, H-J (1996), *Fuzzy Set Theory--and Its Applications*, 3 edn, Kluwer Academic Publishers, Massachusetts.