



Multi-route choice modelling in a metropolitan context: A comparative analysis using Multinomial Logit and Fuzzy Logic based approaches

Sowjanya Dhulipala^{1*}, Ashu Kedia², B. K. Katti³

¹PhD Student, Indian Institute of Technology Bombay, India

²PhD Student, University of Canterbury, New Zealand

³Emeritus Professor, SV National Institute of Technology, Surat, India

Abstract

Route choice plays a vital role in the traffic assignment and network building, as it involves decision making on part of riders. The vagueness in travellers' perceptions of attributes of the available routes between any two locations adds to the complexities in modelling the route choice behaviour. Conventional Logit models fail to address the uncertainty in travellers' perceptions of route characteristics (especially qualitative attributes, such as environmental effects), which can be better addressed through the theory of fuzzy sets and linguistic variables. This study thus attempts to model travellers' route choice behaviour, using a fuzzy logic approach that is based on simple and logical 'if-then' linguistic rules. This approach takes into consideration the uncertainty in travellers' perceptions of route characteristics, resembling humans' decision-making process. Three attributes – travel time, traffic congestion, and road-side environment are adopted as factors driving people's choice of routes, and three alternative routes between two typical locations in an Indian metropolitan city, Surat, are considered in the study. The approach to deal with multiple routes is shown by analyzing two-wheeler riders' (e.g. motorcyclists' and scooter drivers') route choice behaviour during the peak-traffic time. Further, a Multinomial Logit (MNL) model is estimated, to enable a comparison of the two modelling approaches. The estimated Fuzzy Rule-Based Route Choice Model outperformed the conventional MNL model, accounting for the uncertain behaviour of travellers.

Keywords: Route choice, Fuzzy logic, Fuzzy rule-based model, Multinomial logit model, India.

1. Introduction

Traffic assignment is the last and critical phase in the four-stage travel demand modelling process. The choice of route plays a key role as it helps to allocate trips on to the various routes between an origin and a destination. Also, the knowledge of travellers' route choice behaviour helps planners alleviate traffic-related problems, such as congestion, delays, accidents, and air pollution (De Palma and Picard, 2005). Route choice analysis, in general, is based on the econometric concept of Random Utility Maximization (RUM) where the choice of a route is influenced by the attributes

* Corresponding author: Sowjanya Dhulipala (sowjanyad14@gmail.com)

associated with that particular route, and travellers try to maximise their satisfaction by choosing an alternative that provides them with the highest amount of utility (Domencich and McFadden, 1975; Luce, 2005; McFadden, 1986). The most commonly used models based on RUM theory are Logit and Probit models. Selected initial research on these models can be found in Ben-Akiva et al. (1984), Bekhor et al. (2002), Prashker and Bekhor (2004), and Prato (2009). However, they explain very little about the psychological process of decision-makers, as they fail to effectively capture the unobserved portion of utilities associated with various alternatives (Gärling et al., 1994). This limitation leads to the consideration of other approaches, such as models based on various soft-computing techniques, namely, Fuzzy Logic, Artificial Neural Networks, and Neuro-Fuzzy hybrid.

The concept of Fuzzy logic was proposed by Zadeh in 1965 and is powerful in dealing with problems characterised by uncertainty. For instance, the route choice problem is characterised by uncertainties associated with decisionmakers' perceptions of route attributes. Models based on crisp choices are not capable of incorporating uncertainty and vagueness prevailing in the decisions. Fuzzy based models consider input variables in terms of linguistic membership functions rather than crisp inputs and are based on simple and logical *if-then* rules (Wang and Mendel, 1992). For example, consider 'travel time' as an attribute governing the choice between two routes A and B; and values of the variable 'travel time' are adopted as low, medium, and high; then the *if-then* rules can be framed as "*if travel time on route A is less than that on route B, then the preference of choosing route A is high*". As travellers' perceptions of route attributes are characterised by vagueness and ambiguity, a rule-based approach leads to a more realistic consideration of qualitative attributes, such as traffic congestion and comfort along the route.

Teodorovic and Kikuchi (1990) were the first to model the standard route choice problem with two alternative routes, using fuzzy logic, with reference to perceived travel times. Lotan and Koutsopoulos (1993) presented a fuzzy rule-based method for incorporating traffic information in the route choice process comprising three alternative routes. They considered travel time as a criterion for decision making. Furthermore, fuzzy logic was found to be promising in modelling the route choice in the presence of Advanced Traveller Information Systems (Teodorović et al., 1998). Henn (2000) proposed a route choice model with reference to travel cost, where route costs were represented as fuzzy sets. Mizutani and Akiyama (2000) applied a knowledge-based approach using fuzzy logic and created an advanced combined model of binary logit models with both linear and fuzzy logic based utility functions for modal choice between car and mass transit. Ridwan (2004) introduced a fuzzy preference-based model of route choice, where individual travellers' spatial knowledge about routes was taken into account, and fuzzy pair-wise comparisons were made between the available alternatives. Murat and Uludag (2008) presented a route choice model of the transportation network in Denizli, using a fuzzy logic model and logistic regression model.

Route choice of travellers depends on various factors that make the modelling process complex. Mainly, travel time, traffic safety, traffic congestion, travel cost, and environmental effects have been considered as attributes affecting the choice of route in the literature (See Arslan and Khisty, 2005, 2006; Hawas, 2004; Henn, 2000; Teodorovic and Kikuchi, 1990). Most of the studies considered either two routes with two or more route-attributes affecting the choice, or three routes with two attributes affecting the choice. In this study, three routes and three route-attributes are considered simultaneously to model travellers' choice of routes. Three different routes connecting a selected origin-

destination pair in Surat, a newly emerging metropolitan city in India, are selected to study the urban route choice behaviour under the mixed traffic condition. Two models, one based on fuzzy logic, and the other (multinomial logit model) based on RUM theory are estimated with reference to the two-wheeler riders, as two-wheeler is the predominant mode (i.e. a modal share of more than 50%) in the city. The performances of both models are compared and discussed.

The rest of the paper is organised as follows. Section 2 describes the physical and traffic characteristics of the study routes. Data collection and route-attributes are described in Section 3. Fuzzy logic-based modelling is discussed in Section 4. Section 5 discusses the Multinomial logit modelling exercise. Section 6 shows a comparison of both the modelling approaches. Conclusions are discussed in Section 7.

2. Study Routes

2.1 Physical characteristics.

Surat, a fast-growing metropolitan city in the state of Gujarat, India is chosen to study and model the route choice behaviour. Three alternative routes (named as *A*, *B*, *C*) between Ichchhanath (origin), an area in the South-west zone and the city railway station (destination), an area in the Central zone are considered. Route *A* passes through the Central Business District (CBD) covering a length of 8.0 km. Over 50% of this route is a two-lane two-way and 60% of the route is characterized by commercial activities. Route *B* passes through a non-CBD area covering a length of 9.3 km, with six-lane ring road portions characterized by consecutive flyover bridges. Route *C* covers a length of 9.0 km that also passes through the non-CBD area. Route *C* is a part of Route *B* but deviates from *B* for a length of around 3.0 km passing through residential and commercial areas. Figure 1 shows the study routes where OSPD forms Route *A*, OSPQD refers to Route *B* and OSQD forms Route *C*. The stretch PQD is the part of the inner ring road of the city. Figure 2 shows the photographs of a portion of Route *A* passing through the CBD area and the ring road portion of routes *B* and *C* passing through the non-CBD area. The physical characteristics of the study routes are summarized in Table 1.

Table 1: Physical characteristics of routes.

<i>Feature</i>	<i>Route A</i>	<i>Route B</i>	<i>Route C</i>
Length (km)	8	9.3	9
Total number of intersections	15	20	21
Road geometry	50% 2-Lane Two way, 50% 4-Lane Divided	55% 6-Lane divided, 45% 4-Lane divided	15% 6-Lane divided, 85% 4-Lane divided

2.2 Traffic characteristics.

The composition of traffic on the three routes is determined for both morning and evening peak periods by conducting a Classified Vehicle Count (CVC) survey at selected CVC sections. The period from 9:00 AM to 12:00 PM was considered as the morning peak, and from 5:00 PM to 8:00 PM was considered as the evening peak. The peak-traffic periods are decided based on the traffic count survey that was conducted at CVC sections on all the three study routes. The proportions of various vehicle types on these routes are

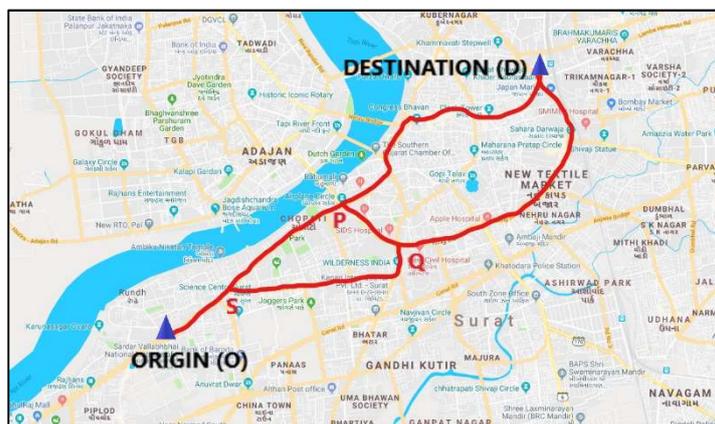


Figure 1: Study routes



Figure 2: (a) Portion of route A (CBD area) (b) Ring road portion of routes B and C

given in Table 2. It can be observed that two-wheelers constitute more than 50% of the traffic.

Table 2: Composition of various types of vehicles on the study routes (in percentages).

Route	2W	3W	4W	Bus	LCV	Truck	Total
A	62.09	20.35	16.29	0.70	0.53	0.04	100
B	57.68	17.66	22.28	0.48	1.65	0.25	100
C	54.73	16.96	25.93	0.40	1.76	0.22	100

Note: 2W: Two-Wheeler, 3W: Three-Wheeler, 4W: Four-Wheeler, LCV: Light Commercial Vehicle

3. Data Collection

A questionnaire was designed to obtain travellers' perceptions of route attributes and their preferences. A road-side interview survey was conducted in the origin area using a questionnaire form. The questionnaire was designed to capture travellers' rationale for the selection of a particular route and their priority among the various criteria. Minimizing travel time was considered the most important criterion affecting travellers' route choice in the literature (e.g. Abdel-Aty et al., 1997; Dubey et al., 2013). Also, directness, less congestion, and environment were among the other concerned route characteristics (e.g. Murat and Uludag, 2008; Ridwan, 2004). Based on this, in the present study, travel time (TT), traffic congestion level (TC), and environmental effects (EE) are considered as the factors affecting the choice of route. EE refers to disturbances and interruptions caused

by factors, such as road-side commercial activities, on-street parking, uncertain and abrupt movement of pedestrians, and air pollution.

Table 3: Attribute ratings.

Rating on scale	TC	EE
1	Very low	Very good
2	Low	Good
3	Medium	Medium
4	High	Poor
5	Very high	Very poor

The survey was conducted during morning and evening peak traffic periods in February and March 2015. Nearly 350 responses were obtained. Riders were asked for their perceptions about the considered attributes along each of the three routes between the selected O-D pair. TT was reported in minutes, and TC and EE were reported in terms of rating on a scale of 1 – 5. Table 3 shows the rating scale and the corresponding levels of TC and EE. A rating of ‘1’ for TC indicates a free-flow condition where riders can attain a travel time of the minimum possible value of that particular route, whereas a rating of ‘5’ indicates the near-jam condition with a ‘*Stop and Go*’ traffic. A rating of ‘1’ for EE indicates a clear route without any road-side disturbances or interruptions to the main traffic, whereas ‘5’ indicates the route is highly concentrated with road-side commercial activities and interruptions. Also, respondents were asked to rank their priorities for all the attributes and routes considered in the study. Route preference values for each route were obtained in percentages from each rider interviewed.

3.1 Observations from the survey.

It is identified that work trips constitute the major share of total trips. Figure 3(a) shows the purpose-wise distribution of trips, and 3(b) shows the observed route preferences of respondents. Route B is the most preferred route followed by the routes C and A. Among the total respondents, 69% preferred TT, 23% preferred TC, and 8% preferred EE. Thus, TT is found to be the most influential factor followed by TC and EE.

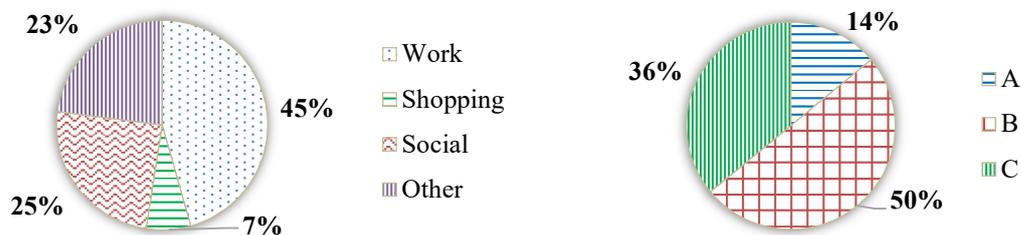


Figure 3: (a) Purpose-wise distribution of trips (b) Observed route preference

4. Fuzzy Rule-Based Model

4.1 Model framework.

The framework of the fuzzy rule based route choice model (FRB-RCM) follows a hierarchy. It comprises three comparative sub-models, namely, B-A sub-model, B-C sub-model, and C-A sub-model. Initially, B-A sub-model is run where the attributes are

compared between the routes A and B, and the riders opting A and B are separated into two different groups, named as *Group I* and *Group II*. These riders, when provided with route C may opt for C. Therefore, to check this possibility, these groups of riders are fed into next-level models i.e. B-C and C-A sub-models where the riders opting for C are filtered out. The detailing of the hierarchical process of the modelling framework is shown in Figure 4.

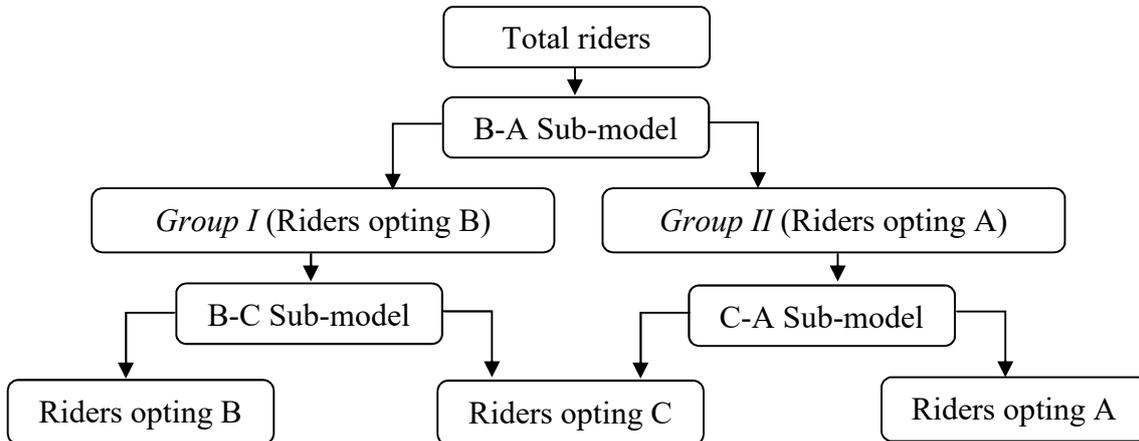


Figure 4: Hierarchy of the modelling framework

4.2 Model structure.

A Fuzzy Inference System (FIS) is a way of mapping an input space to an output space using the concept of fuzzy sets and control. FIS consists of three modules, called fuzzification, fuzzy inference engine, and defuzzification (Klir and Yuan, 1995). The structure of a typical fuzzy system is shown in Figure 5.

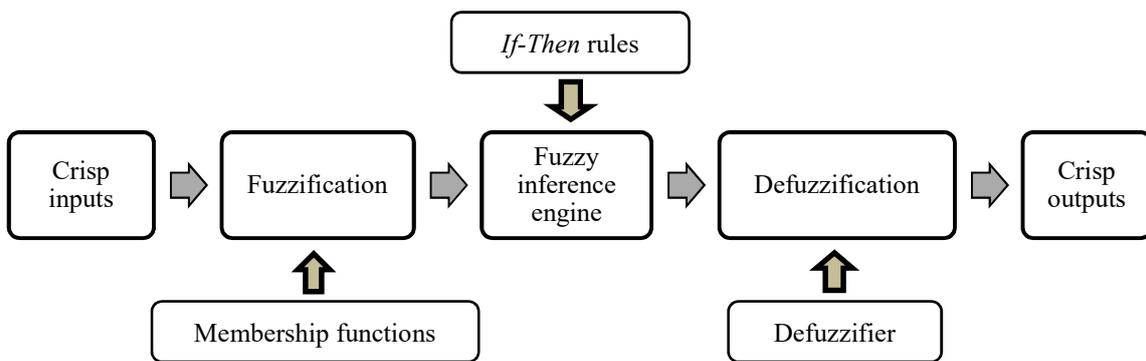


Figure 5: Structure of a fuzzy control system

In the fuzzification module, system inputs (crisp numbers) are transformed into fuzzy sets. This is done by applying a fuzzification function, i.e. membership function (MF). There are various types of MFs, such as triangular, trapezoidal, and gaussian MFs (Rajasekaran and Pai, 2003). Some studies used clustering methods, such as fuzzy c-mean clustering (e.g. Othayoth and Katti, 2017) or subtractive clustering (e.g. P.S. et al., 2017; Pulugurta et al., 2013), to arrive at the shapes of MFs. However, triangular MFs have been used in several studies (e.g. Dhulipala et al., 2017; Kedia et al., 2015, 2017; Pedrycz,

1994) and were found to be promising. Thus, triangular MFs are adopted to fuzzify the crisp inputs with five fuzzy sets for both input and output variables. After fuzzification, fuzzy inference system simulated the human reasoning process by making fuzzy inference on inputs and *If-Then* rules. There are two types of FIS, namely, *Sugeno* and *Mamdani* (Mamdani and Assilian, 1975). *Mamdani* FIS has been widely used, particularly for decision support applications, because of the intuitive and interpretable nature of the rule base (Jassbi et al., 2007). Thus, the *Mamdani* inference system is used in this study. A total of 125 *If-Then* rules are derived based on the observed data, to map the three inputs, each comprising five levels of MFs with the single output that comprised five levels of MFs.

The final stage is the defuzzification process, where the fuzzy output generated by the inference mechanism is transformed into numerical values for application purposes. There are several methods of defuzzification, such as Centre of gravity (i.e. Centroid), Centre of area, Maximum height, and Means of maxima. However, the most commonly used method is the Centroid method, because it considers the weighted average of the elements in the support set and analyses the combined shape of membership functions, which results in better results (Kedia et al., 2015; Kumar et al., 2013).

4.3 Fuzzy Rule Based Route Choice Model operations.

➤ Model inputs and outputs

The perceived values of the route attributes, TT in minutes, TC and EE on a rating scale of 1-5, are used as inputs to the model. For the comparative sub-models, inputs are the differences between the TTs, TCs, and EEs of the two routes being compared. For the B-A sub-model, the input value for the attribute TT is the time difference in minutes for the routes B and A (i.e. travel time in route B minus travel time in route A). The input value for TC is the difference in the levels of route B and route A (i.e. traffic congestion level on route B minus traffic congestion level on route A). Similarly, the input value for EE is the difference in the levels of route B and route A. The output of the model is the preference of route B over route A by riders in percentage. These inputs and outputs are considered in the form of five fuzzy sets as given in Table 4.

Table 4: Fuzzy sets for the input and output variables.

<i>TT</i>	<i>TC level</i>	<i>EE</i>	<i>Preference of B over A</i>
Much Less (ML)	Much Less (ML)	Much Better (MB)	Very High (VH)
Less (LS)	Less (LS)	Better (BR)	High (H)
Equal (EQ)	Equal (EQ)	Equal (EQ)	Medium (M)
Greater (GR)	Greater (GR)	Poor (PR)	Low (L)
Much Greater (MG)	Much Greater (MG)	Much Poor (MP)	Very Low (VL)

➤ Membership functions

In the input data, the TT values are ranging from -45 to 45 minutes, TC from -4 to +4 levels, and EE from -3 to +3 levels. The output values are in the range of 0 to 100. Five triangular MFs are framed for each input and output variable and ranges of MFs are given in Table 5. The ranges of MFs are decided based on the C-means clustering and

calibration. The shapes of MFs for input parameters are shown in Figures 6(a-c), and for the output parameter, in Figure 6d.

Table 5: Ranges of Membership Functions.

<i>Differences in TT (minutes)</i>		<i>Differences in TC (levels)</i>		<i>Differences in EE (levels)</i>		<i>Preference of B over A (percentage)</i>	
<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>
ML	[-45 -45 -23]	ML	[-5 -5 -2.5]	MB	[-4 -4 -1.3]	VL	[0 0 25]
LS	[-25 -13 -3]	LS	[-2.7 -1.5 0]	BR	[-2 -1 0]	L	[15 30 45]
EQ	[-5 0 5]	EQ	[-1 0 1]	EQ	[-1 0 1]	M	[35 50 65]
GR	[3 13 22]	GR	[0.5 1.5 2.5]	PR	[0.5 1.5 2.5]	H	[55 70 85]
MG	[20.2 45 45]	MG	[2.3 5 5]	MP	[2 3 3]	VH	[75 100 100]

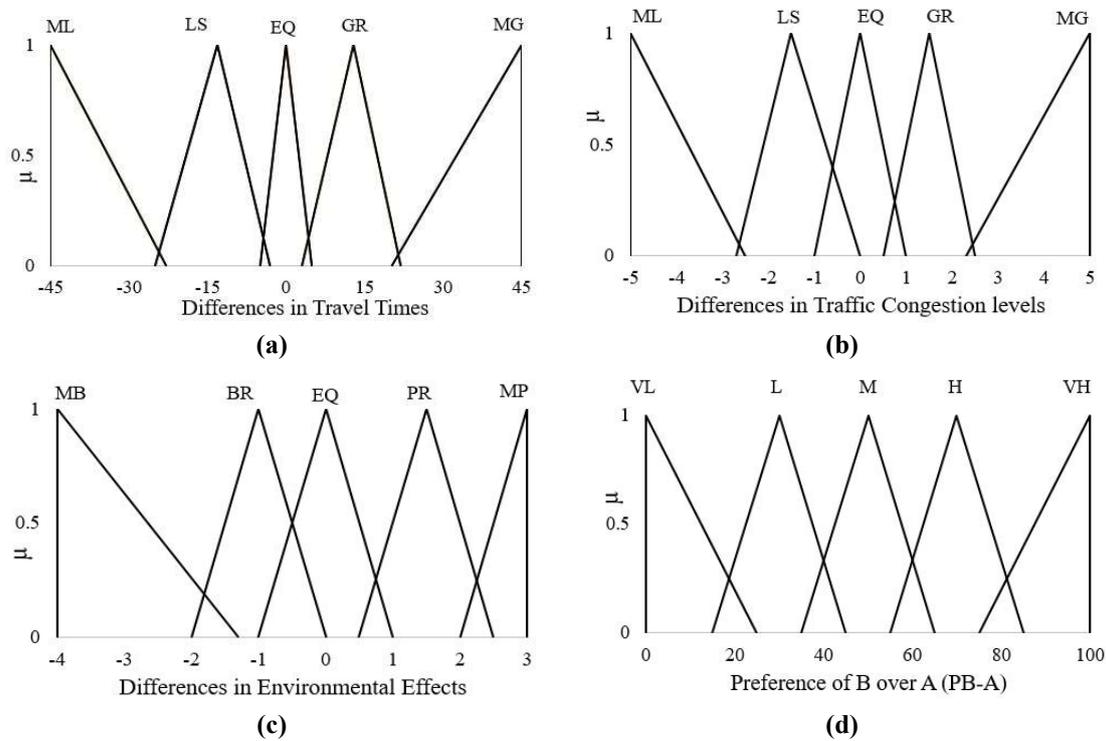


Figure 6: MFs of (a) TT, (b) TC, (c) EE, and (d) PB-A

➤ Framing of If-Then rules

If-Then rules, which represent the human decision-making process, are framed using general user experience and common sense. A total of 125 rules ($5 \times 5 \times 5$) are framed using an ‘AND’ operator. The process of framing rules is shown in Figure 7. A typical rule is ‘*If travel time of Route B is Much Less than Route A, and traffic congestion of Route B is Much Less than Route A, and Environmental Effect of Route B is Much Better than Route A, then Preference of Route B over Route A is Very High*’.

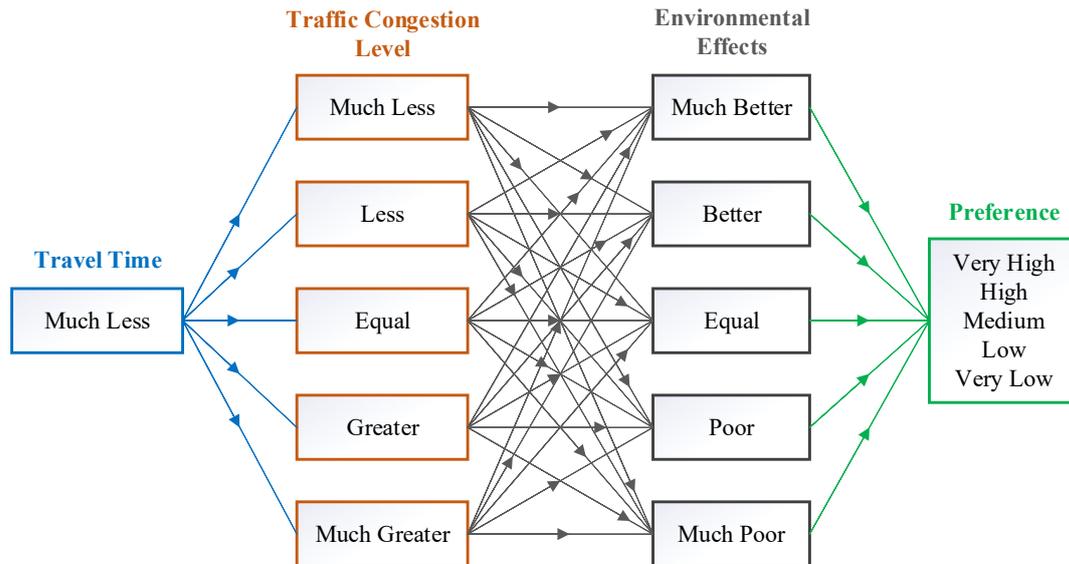


Figure 7: Schematic diagram of formulation of *If-Then* rules

➤ FRB-RCM results

The output of the B-A sub model is the preference of route B over route A. The cut-off value is taken as 50%, where the riders opt for route B if the model percentage preference value is greater than 50 and route A if it is less than 50. Thus, the riders are distributed among *Group I* and *Group II* based on the percentage preference score. Therefore, the output of the model is the distribution of riders among B (*Group I*) and A (*Group II*) as given in Table 6.

Table 6: Summary of FRB-RCM results.

The output of B-A Sub-model					
Route	Number of responses		Responses (%)		
Group I	201		80.4		
Group II	49		19.6		
Total	250		100		
The output of B-C sub-model			The output of C-A sub-model		
Route	Riders (#)	Riders (%)	Route	Riders (#)	Riders (%)
B	123	61.2	A	26	53.06
C	78	38.8	C	23	46.94
Total	201*	100	Total	49*	100

Numbers; * From B-A sub-model

Group I riders are fed into B-C sub-model to check for riders opting route C. The ranges of triangular MFs formed for the input variables of B-C sub-model are shown in the Appendix I. The output of the model is the preference of route B over route C as given in Table 6. *Group II* riders are fed into the C-A sub-model to separate riders opting for route C. The category-wise ranges and their triangular MFs of the input variables are provided in Appendix I. The output of the model is the preference of route C over route A as shown in Table 6.

Based on the above three sub-models, the riders are distributed among the three routes as shown in Table 7. The modelled and observed percentages of riders opting for three

routes are quite close. The prediction accuracy of the model for calibration data is shown in the cross-classification table (Table 8). The model is able to predict riders' choices of the route with an accuracy of 90.8%.

Table 7: Distribution of riders.

Route	Modelled		Observed
	Riders (nos.)	Riders (%)	Riders (%)
A	26	10.4	14.8
B	123	49.2	47.6
C	101	40.4	37.6
Total	250	100	100

Table 8: Cross-classification table.

Observed no. of responses	Predicted no. of responses				
	Route	A	B	C	Total
A	25	0	12	37	
B	0	116	3	119	
C	1	7	86	94	
Total	26	123	101	250	

Prediction accuracy = 90.8%

➤ Validation of the model

From the data of 350 riders, nearly 30% is used for validation of the model. Table 9 shows the cross-classification of the validation results. Clearly, the prediction accuracy for the validation data is as good as calibration. The developed fuzzy logic-based model is statistically validated with the help of the Chi-Square test, as shown in Table 10. O_i represents values obtained from the survey and E_i represents the values estimated by the model. The Chi-square statistic value is less than the critical value corresponding to the 95% confidence level. Therefore, we failed to reject the null hypothesis (H_0), and hence the model is statistically acceptable.

Table 9: Cross-classification table.

Observed no. of responses	Predicted no. of responses				
	Route	A	B	C	Total
A	9	2	1	12	
B	0	59	0	59	
C	0	5	27	32	
Total	9	66	28	103	

Prediction accuracy = 92.23%

Table 10: Statistical validation.

Chi-square test†			
Route	Survey (O_i)	Model (E_i)	$(O_i - E_i)^2 / E_i$
A	26	37	4.65
B	123	119	0.13
C	101	94	0.48
Chi-square			5.26
Critical value for sig. 0.05, at df = 2			7.37

† Null hypothesis (H_0): There is no difference between observed and modelled frequencies of routes.

5. Multinomial logit Model

As the routes B and C are partially overlapped sharing similar route characteristics, they can be nested under the same branch. Hence, initially, a Nested Logit (NL) model was estimated by nesting the routes B and C under one branch (i.e. Ring road, RR) and keeping the route A (i.e. via Central Business District, CBD) separate. However, the inclusive value (IV) parameter obtained for the branch RR was found to be 1.031, which implied that the model became equivalent to Multinomial Logit (MNL) model (Hensher et al., 2015). The IV parameter of the branch CBD was less than 0, indicating that an increase in the utility of route A would diminish the probability of choosing the route A, which is illogical. Therefore, it appeared that the MNL model is more suitable to the data

rather than the NL model. Although there is an overlapping portion for routes B and C, there is a substantial difference in the characteristics of the uncommon portions of these routes, such as commercial activities, which makes them comparable, making the MNL model appropriate for the present problem. The results of the NL model are not reported here, as it is not preferred.

5.1 Model Structure

The MNL model, which takes the form as in equation (1), is used to model the route choice behaviour. The explanatory variables used in the model are TT (minutes) and TC (in terms of rating), while the response variable is the route chosen. EE was not considered in the final model, as it did not turn out to be statistically significant.

$$P(i) = \frac{e^{U_i}}{\sum_{i=1}^n e^{U_i}} \quad (1)$$

$P(i)$ is the probability of the choice of a route (i),
 U_i is the utility function for the i^{th} route and is linearly regressive as

$$U_i = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \quad (2)$$

$i = 1, 2, 3$ routes; x_1 : TT, x_2 : TC, x_3 : EE

The estimated MNL model is specified as follows:

$$U_A = b_{0A} + b_1(TT_A) + b_{2A}(TC_A) \quad (3)$$

$$U_B = b_{0B} + b_1(TT_B) + b_2(TC_B) \quad (4)$$

$$U_C = b_1(TT_C) + b_2(TC_C) \quad (5)$$

where b_{0A}, b_{0B} : Alternative specific constants for route A, route B respectively

b_1 : Coefficient of TT; b_{2A} : Alternative specific coefficient of TC for route A;

b_2 : Coefficient of TC for routes B and C

The parameter values obtained from the model are summarized in Table 11. The coefficient for TT is considered as generic. The coefficient for TC is made alternative specific by considering the coefficient of TC for route A separately, due to its distinguishing characteristics. The TC parameter for route A is not statistically significant. This implies that riders choosing route A made their choice irrespective of congestion and environment conditions prevailing on that route. It is apparent that the choice of respondents preferring route A is based not only on the factors considered in the model but also on various other reasons, such as the shorter travel distance and en-route activities on route A. The sign of the coefficient of TT is negative, which indicates that the probability of choosing a particular route increases if the travel time on that route decreases. The sign of the coefficient of TC for routes B and C is negative, which indicates that the probability of choosing each of those routes decreases with the increase in the traffic congestion level on those routes. Although not statistically significant, the sign of the coefficient of TC for route A is positive; this is perhaps because riders preferred to choose route A irrespective of the congestion on that route. The estimated model

improved over the base (i.e. constants only) model, in terms of the log-likelihood value. Also, the Pseudo R-squared value (0.529) indicates a decent model fit, as the Pseudo R-squared value between 0.2 and 0.4 is equivalent to an R-squared value between 0.5 and 0.8 for the ordinary least square regression (Domencich and McFadden, 1975; Kedia et al., 2019; Kusumastuti and Nicholson, 2017). The estimated model is found to have a prediction accuracy of 78.4% for the calibration data, as shown in Table 12.

Table 11: Results of the MNL Model.

<i>Parameter</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>P-value</i>
b_{0A}	7.262***	2.305	0.002
TT	-0.396***	0.065	<0.001
TC_A	0.030	0.036	0.402
b_{0B}	0.715***	0.195	<0.001
TC_{BC}	-1.420***	0.189	<0.001
<i>Model statistics</i>			
Log-Likelihood function			-116.470
Log-Likelihood of the 'constants only' model			-247.769
Pseudo R-squared			0.529
Adj. R-squared			0.525

Note: *** → significant at the 1% significance level.

Table 12: Cross-classification table.

<i>Observed no. of responses</i>	<i>Predicted no. of responses</i>				
	<i>Route</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>Total</i>
<i>A</i>		30	0	3	33
<i>B</i>		0	89	24	113
<i>C</i>		1	26	77	104
<i>Total</i>		31	115	104	250

Prediction accuracy = 78.4%

6. Comparison of Fuzzy and Logit Models

Fuzzy rule-based model dealt the problem with its fuzzy inference engine, i.e. with the logical “*If-Then*” rules. The model is able to learn the vagueness in the statements of respondents and is able to predict the behaviour of riders appropriately. It is observed that the behaviour of riders is different for different routes; their priorities for attributes are changing with routes. Moreover, despite the condition of the route, riders tend to prefer a particular route due to several other reasons. This behaviour is well captured by the fuzzy rule-based model with its logical rules. Also, the subjective levels of attributes TC and EE are handled well with the MFs of the fuzzy system. MNL model, on the other hand, is found to model the situation well in the absence of the attribute EE. It is able to understand the behaviour of riders with respect to the attribute ‘traffic congestion’; however, it lacks statistical significance. That is, the typical behaviour of riders (i.e. prefer to choose the route with minimum travel time and minimum congestion) is well understood by the model whereas the unusual behaviour (i.e. prefer to choose the route even though it has higher travel time and higher congestion) remains unexplained. Fuzzy-based model outperformed the MNL model in terms of the prediction accuracy. Also, the fuzzy-based model can be considered as a step over the MNL model due to its behavioural understanding, while Logit models are statistically stronger than fuzzy-based models.

7. Conclusions

Route choice in the urban traffic situation plays an important role in the allocation of trips over a road network, as it can help planners optimize the system. Therefore, the

analysis of travellers' perceptions of traffic attributes, such as travel time and traffic congestion, is essential to model the route choice and shift behaviour. In the present study, the route choice among three alternative routes between an origin and a destination in an urban area is analyzed by considering travel time, traffic congestion level, and environmental effects as the explanatory factors. Both the fuzzy rule-based and multinomial logit models are developed to study the riders' behaviour of route choice. It is found that travel time is the most influential factor among the three attributes and the environmental effect is the least. The fuzzy-based model showed good prediction accuracy considering all the three attributes. The *If-Then* rules of the fuzzy model are able to replicate the behaviour of riders reasonably well. The multinomial logit model also performed reasonably well in terms of the prediction accuracy but is not able to explain the effect of the qualitative attribute environmental effect, which is less preferred by riders. Also, the model is unable to capture the riders' peculiar behaviour towards congestion on route A. Thus, even though both the models performed well, the fuzzy-based model can be considered better due to its ability to capture the uncertain behaviour. The developed models can be used as tools for determining the most preferred routes in the network for various traffic conditions. This helps in obtaining the link volumes, which can be used in network assignment.

Some limitations of the study are: the fuzzy rule based model did not account for the weights of attributes, which, if considered, may replicate the scenario better. The data used for the study are based on riders' perceptions of the attributes of three routes and their preferences for each route. There is a possibility that riders might not have experienced one or the other route, resulting in a biased preference. Hybrid techniques, consistent with fuzzy models, such as neuro-fuzzy may help understand the choice behavior in a better way. Other models in the logit family, such as random parameter logit might well help examine the route choice behaviour.

References

- Abdel-Aty, M. A., Kitamura, R., & Jovanis, P. P. (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transportation Research Part C: Emerging Technologies*, 5(1), 39–50. [https://doi.org/10.1016/S0968-090X\(96\)00023-X](https://doi.org/10.1016/S0968-090X(96)00023-X)
- Arslan, T., & Khisty, C. J. (2005). A rational reasoning method from fuzzy perceptions in route choice. *Fuzzy Sets and Systems*, 150(3), 419–435. <https://doi.org/10.1016/j.fss.2004.03.021>
- Arslan, T., & Khisty, J. (2006). A rational approach to handling fuzzy perceptions in route choice. *European Journal of Operational Research*, 168(2), 571–583. <https://doi.org/10.1016/j.ejor.2004.04.043>
- Bekhor, S., Ben-Akiva, M., & Scott Ramming, M. (2002). Adaptation of Logit Kernel to Route Choice Situation. *Transportation Research Record: Journal of the Transportation Research Board*, 1805, 78–85. <https://doi.org/10.3141/1805-10>
- Ben-Akiva, M., Bergman, M., Daly, A., & Ramaswamy, R. (1984). Modelling Inter Urban Route Choice Behaviour. In *Ninth International Symposium on Transportation & Traffic Theory* (pp. 299–330). VNU Science press.
- de Palma, A., & Picard, N. (2005). Route choice decision under travel time uncertainty. *Transportation Research Part A: Policy and Practice*, 39(4), 295–324. <https://doi.org/10.1016/j.tr.2004.10.001>

- Dhulipala, S., Kedia, A. S., P.S., S., & Katti, B. K. (2017). Building A Neuro-Fuzzy Based Route Choice Model in Metropolitan Context: Surat City in India. *Transportation Research Procedia*, 25, 3203–3219. <https://doi.org/10.1016/j.trpro.2017.05.363>
- Domencich, T. A., & McFadden, D. (1975). *Urban travel demand-a behavioral analysis*.
- Dubey, S. K., Mishra, D., Arkatkar, S. S., Singh, A. P., & Sarkar, A. K. (2013). Route Choice Modelling Using Fuzzy logic and Adaptive Neuro-fuzzy. *Modern Traffic and Transportation Engineering Research (MTTER)*, 11–19.
- Gärling, T., Kwan, M.-P., & Golledge, R. G. (1994). Computational-process modelling of household activity scheduling. *Transportation Research Part B: Methodological*, 28(5), 355–364. [https://doi.org/10.1016/0191-2615\(94\)90034-5](https://doi.org/10.1016/0191-2615(94)90034-5)
- Hawas Yaser E. (2004). Development and Calibration of Route Choice Utility Models: Neuro-Fuzzy Approach. *Journal of Transportation Engineering*, 130(2), 171–182. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2004\)130:2\(171\)](https://doi.org/10.1061/(ASCE)0733-947X(2004)130:2(171))
- Henn, V. (2000). Fuzzy route choice model for traffic assignment. *Fuzzy Sets and Systems*, 116(1), 77–101. [https://doi.org/10.1016/S0165-0114\(99\)00039-1](https://doi.org/10.1016/S0165-0114(99)00039-1)
- Jassbi, J. J., Alavi, S. H., Serra, P. J., & Ribeiro, R. A. (2007). Transformation of a Mamdani FIS to First Order Sugeno FIS. *FUZZ-IEEE*, 1–6.
- Kedia, A., Kusumastuti, D., & Nicholson, A. (2019, September 30). Effect of online shopping on consumers' shopping travel and goods' last-mile travel: A case study in Christchurch. *Australasian Transport Research Forum 2019 Proceedings*. 41st Australasian Transport Research Forum, Canberra.
- Kedia, A. S., Dhulipala, S., Salini, P. S., Mannilthodi, J., & Katti, B. K. (2017). Transit Shift Response Analysis Through Fuzzy Rule Based-Choice Model: A Case Study of Indian Metropolitan City. *Transportation in Developing Economies*, 3(1), 8. <https://doi.org/10.1007/s40890-017-0038-9>
- Kedia, A. S., Saw, K. B., & Katti, B. K. (2015). Fuzzy logic approach in mode choice modelling for education trips: A case study of Indian metropolitan city. *Transport*, 30(3), 286–293. <https://doi.org/10.3846/16484142.2015.1081279>
- Klir, G. J., & Yuan, B. (1995). Fuzzy sets and fuzzy logic: Theory and applications. *Upper Saddle River*, 563.
- Kumar, M., Sarkar, P., & Madhu, E. (2013). Development of fuzzy logic based mode choice model considering various public transport policy options. *International Journal for Traffic and Transport Engineering*, 3(4), 408–425.
- Kusumastuti, D., & Nicholson, A. (2017). *Mixed-Use Urban Planning and Development*. <https://ir.canterbury.ac.nz/handle/10092/15417>
- Lotan, T., & Koutsopoulos, H. N. (1993). Models for route choice behavior in the presence of information using concepts from fuzzy set theory and approximate reasoning. *Transportation*, 20(2), 129–155. <https://doi.org/10.1007/BF01307056>
- Luce, R. D. (2005). *Individual Choice Behavior: A Theoretical Analysis*. Courier Corporation.
- Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7(1), 1–13. [https://doi.org/10.1016/S0020-7373\(75\)80002-2](https://doi.org/10.1016/S0020-7373(75)80002-2)
- McFadden, D. (1986). The Choice Theory Approach to Market Research. *Marketing Science*, 5(4), 275.
- Mizutani, K., & Akiyama, T. (2000). A logit model for modal choice with a fuzzy logic utility function. *Traffic and Transportation Studies*, 311.
- Murat, Y. S., & Uludag, N. (2008). Route choice modelling in urban transportation networks using fuzzy logic and logistic regression methods. *JSIR Vol.67(1) [January 2008]*. <http://nopr.niscair.res.in/handle/123456789/731>
- Othayoth, D., & Katti, B. K. (2017). Modelling Trip Distribution Using Fuzzy Logic Approach. *Transportation in Developing Economies*, 3(2), 15. <https://doi.org/10.1007/s40890-017-0036-y>

- Pedrycz, W. (1994). Why triangular membership functions? *Fuzzy Sets and Systems*, 64(1), 21–30. [https://doi.org/10.1016/0165-0114\(94\)90003-5](https://doi.org/10.1016/0165-0114(94)90003-5)
- Prashker, J. N., & Bekhor, S. (2004). Route Choice Models Used in the Stochastic User Equilibrium Problem: A Review. *Transport Reviews*, 24(4), 437–463. <https://doi.org/10.1080/0144164042000181707>
- Prato, C. G. (2009). Route choice modeling: Past, present and future research directions. *Journal of Choice Modelling*, 2(1), 65–100. [https://doi.org/10.1016/S1755-5345\(13\)70005-8](https://doi.org/10.1016/S1755-5345(13)70005-8)
- P.S., S., Kedia, A., Dhulipala, S., Saw, K., & Katti, B. K. (2017). Spatial distribution of urban trips in recently expanded Surat city through Fuzzy Logic with various clustering Techniques: A case study of typical metropolitan city in India. *Transportation Research Procedia*, 25, 2396–2407. <https://doi.org/10.1016/j.trpro.2017.05.245>
- Pulugurta, S., Arun, A., & Errampalli, M. (2013). Use of artificial intelligence for mode choice analysis and comparison with traditional multinomial logit model. *Procedia-Social and Behavioral Sciences*, 104, 583–592.
- Rajasekaran, S., & Pai, G. V. (2003). *Neural networks, fuzzy logic and genetic algorithm: Synthesis and applications (with cd)*. PHI Learning Pvt. Ltd.
- Ridwan, M. (2004). Fuzzy preference based traffic assignment problem. *Transportation Research Part C: Emerging Technologies*, 12(3), 209–233. <https://doi.org/10.1016/j.trc.2004.07.003>
- Teodorovic, D., & Kikuchi, S. (1990). Transportation route choice model using fuzzy inference technique. [1990] *Proceedings. First International Symposium on Uncertainty Modeling and Analysis*, 140–145. <https://doi.org/10.1109/ISUMA.1990.151240>
- Teodorović, D., Vukanović, S., & Obradović, K. (1998). Modeling route choice with advanced traveler information by Fuzzy Logic. *Transportation Planning and Technology*, 22(1), 1–25. <https://doi.org/10.1080/03081069808717617>
- Wang, L.-X., & Mendel, J. M. (1992). Generating fuzzy rules by learning from examples. *Systems, Man and Cybernetics, IEEE Transactions On*, 22(6), 1414–1427.

Appendix I

Table 13: Ranges of MFs for B-C Sub-model.

<i>Differences in TT (minutes)</i>		<i>Differences in TC (levels)</i>		<i>Differences in EE (levels)</i>		<i>Preference of B over C (percentage)</i>	
<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>
ML	[-45 -45 -12]	ML	[-5 -5 -1.5]	MB	[-4 -4 -1.3]	VL	[0 0 25]
LS	[-15 -8 0]	LS	[-2 -1 0]	BR	[-2 -1 0]	L	[15 30 45]
EQ	[-5 0 5]	EQ	[-1 0 1]	EQ	[-1 0 1]	M	[35 50 65]
GR	[3 10 15]	GR	[0.5 1 1.5]	PR	[0.5 1.5 2.5]	H	[55 70 85]
MG	[12 45 45]	MG	[1.3 5 5]	MP	[2 3 3]	VH	[75 100 100]

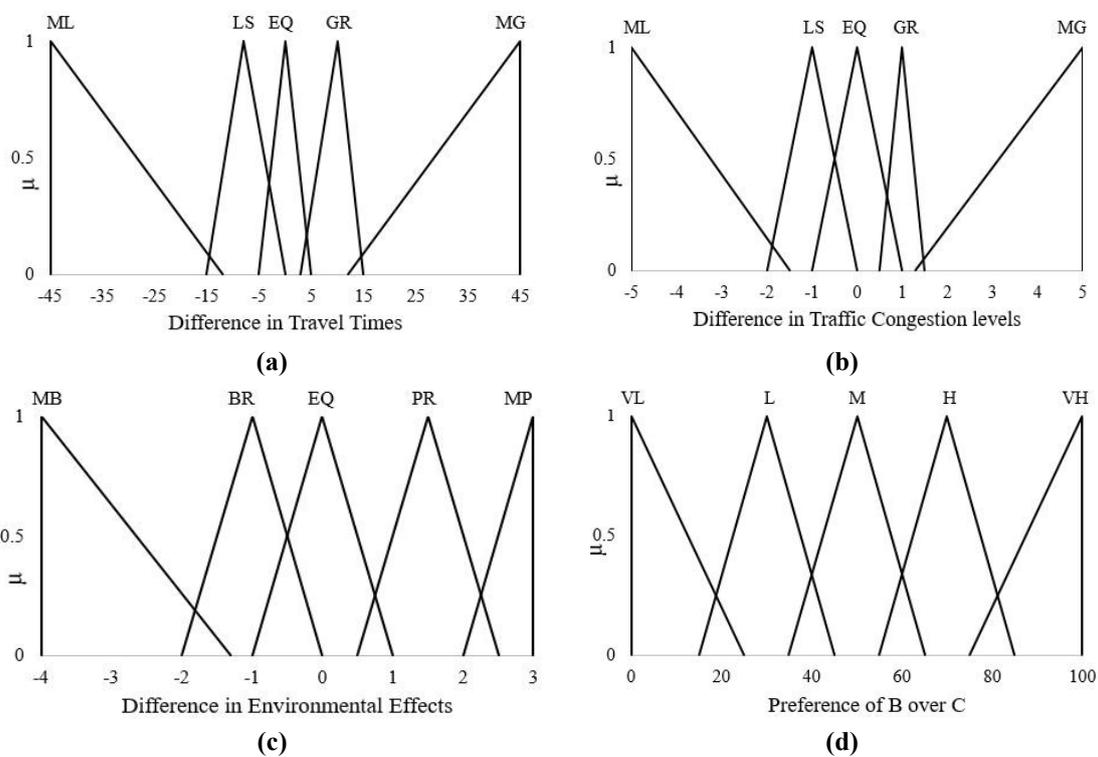


Figure 8: MFs of (a) TT, (b) TC, (c) EE, and (d) PB-C

Table 14: Ranges of MFs for C-A Sub-model.

<i>Differences in TT (minutes)</i>		<i>Differences in TC (levels)</i>		<i>Differences in EE (levels)</i>		<i>Preference of C over A (percentage)</i>	
<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>	<i>MF</i>	<i>Range</i>
ML	[-45 -45 -17]	ML	[-5 -5 -1.5]	MB	[-4 -4 -1.3]	VL	[0 0 25]
LS	[-20 -15 5]	LS	[-2 -1 0]	BR	[-2 -1 0]	L	[15 30 45]
EQ	[-7 2 3]	EQ	[-1 0 1]	EQ	[-1 0 1]	M	[35 50 65]
GR	[1 10 17]	GR	[0.5 1.5 2.5]	PR	[0.5 1.5 2.5]	H	[55 70 85]
MG	[15 40 40]	MG	[2.3 5 5]	MP	[2 3 3]	VH	[75 100 100]

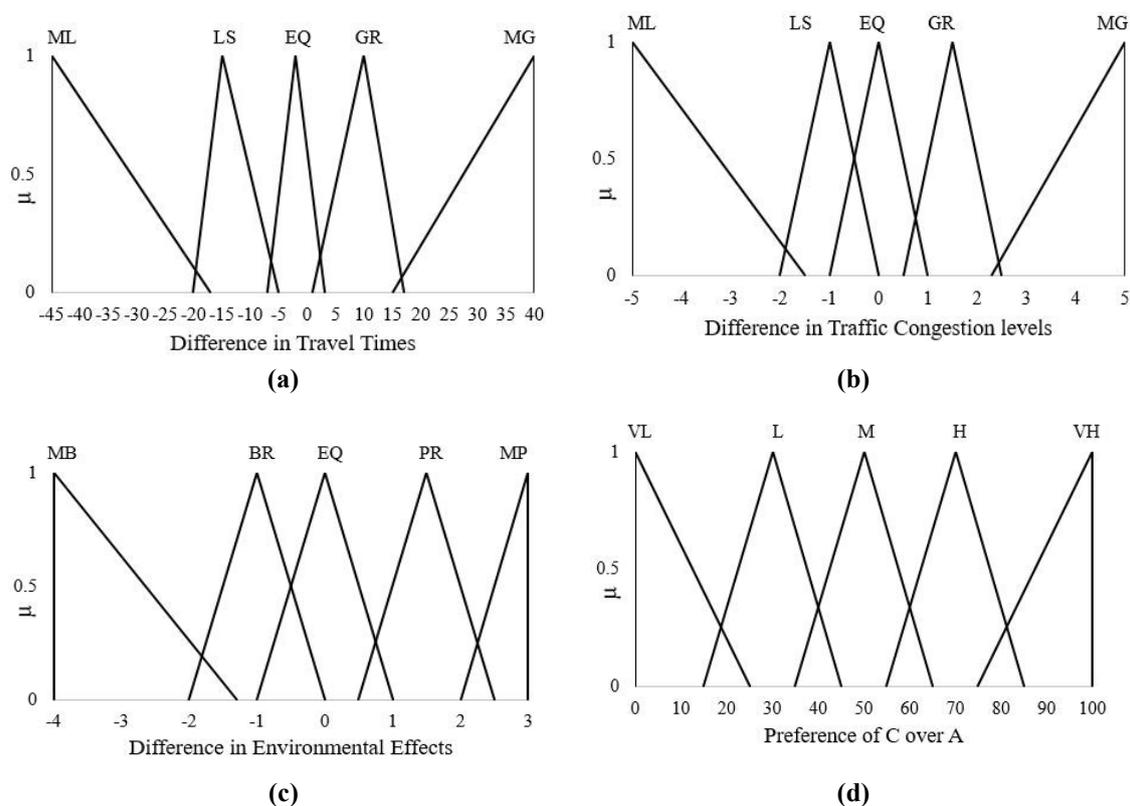


Figure 9: MFs of (a) TT, (b) TC, (c) EE, and (d) PC-A