IRAQI Academic Scientific Journals

Iraqi Journal of Civil Engineering

Available online at https://www.iasj.net/

Journal homepage: https://www.iasj.net/iasj?func=issues&jId=141&uiLanguage=en

Developing a Modal Split Model Using Fuzzy Inference System in Ramadi City

Omaima A. Yousif [1*], Adil N. Abed[2], Hamid A. Awad[2]

[1] *MSc Student, Department of Civil Engineering, College of Engineering, University Of Anbar. Ramadi, Anbar, Iraq.*

[2] *Department of Civil Engineering, College of Engineering, University Of Anbar. Ramadi, Anbar, Iraq*

ARTICLE INFO

Article history:
Received 06 /10 / 2021
Received in revised form 11 /12 / 2021
Accepted 12 /12 / 2021
Available online 08 /02 / 2022

Keywords:
Modal split model
Fuzzy inference system (FIS)
Traffic analysis zones (TAZ)
Land use
Socioeconomic characteristics

ABSTRACT

Several different deterministic and probabilistic mathematical approaches have been used to develop modal split models. The data collected by a questionnaire survey approach is frequently associated with subjectivity, imprecision, and ambiguity. additionally, several linguistic terms are used to express some of the transportation planning variables. This can be solved by modeling mode choosing behavior with artificial intelligence techniques such as fuzzy logic. In this research, Ramadi city in Iraq has been selected as a study area. For the purpose of obtaining data, the study area was divided into traffic analysis zones (TAZ). The total number of traffic zones was set as 28 traffic zones, 22 were internal traffic zones and 6 external traffic zones. Field surveys and questionnaires are used to collect data on traffic, land use, and socioeconomic characteristics factors (age, gender, vehicle ownership, family income, trip purpose, trip origin and destination, trip time, waiting duration, duration inside mode, trip origin and destination, trip cost, and type of mode used for transport). The results showed that the modal split models based on the fuzzy inference system can deal with linguistic variables as well as address uncertainty and subjectivity and they gave very good prediction accuracy for future prediction. Fuzzy inference system proved that all factors affected the mode choice with a very strong correlation coefficient (R) equal to 93.1 for general trips but when the results were compared with multiple linear regression model found that the correlation coefficient (R) equal to 28.9 for general trips and the most influential factors on the mode choice are car ownership, age and trip cost. Thus, it can be concluded that fuzzy logic models were more capable of capturing and integrating human knowledge in mode selection behavior. In addition, this study will help decision-makers to plan transportation policies for Ramadi city.

1. Introduction

Transportation planning can be defined as the movement of people, goods, and services from one location to another in a safe and efficient condition. Transportation planning is critical for designing cities, supporting economic activity, encouraging community interaction, and improving quality of life. It is also necessary for long-term development and maintaining safe access for all people at various levels (Pulugurta, Arun, &

* *Corresponding author.* E-mail address: oma19e1017@uoanbar.edu.iq

Errampalli, 2013).. Travel demand is described as the number of people or vehicles that may be expected to travel on a specified section of a transportation system in a given unit of time under a specific set of land use, socioeconomic, and environmental conditions. The location and intensity of land use, the socio-economic characteristics of individuals living in the area, and the distance, cost, and quality of available transportation options are all factors that influence urban travel demand (Chakroborty & Das, 2017). In a typical individual's daily decision-making process, four types of decisions are made: trip generation, distribution, mode choice, and route assignment, which are the main stages in the travel demand and are of great importance in transportation planning. Modal split (mode choice) analysis is the third step in the four-step transportation forecasting model after trip generation and distribution but before route assignment. Mode choice is one of the most critical stages in the process of transport planning, it relates to the type of mode that travelers use to reach their destination and has a significant effect on the decision-making process. Mode choice models deal very closely with the behavior of human choice -making and therefore continue to draw researchers to further investigate the decision-making process of commuters (Minal & Sekhar, 2014).Traditional models such as binary and multinomial logit models, etc. are used for mode choice analysis, in these models, the input parameters should have crisp values in order to be accurately measured, which involve time and resources. Furthermore, the mode chosen by the traveler requires human approximations that are not represented by these models. This can be solved by modeling mode choosing behavior with artificial intelligence techniques such as fuzzy logic. Fuzzy logic allows for a great deal of flexibility in reasoning, allowing for the accounting of inaccuracies, uncertainties, and a lack of information. The construction of Fuzzy rules is based on human knowledge and experience and uses the conditional IF-Then rule similar to human behavior and knowledge where linguistic terms are used (Georgieva, 2016). The aim of this study to analyze the factors affecting the selection mode choice in different zones of Ramadi city using the fuzzy logic inference. Many researchers have developed mode choice prediction models worldwide. Obaid and Hamad, (2019), developed a multinomial logit model to predict the mode choice for the University City of Sharjah in the United Arab Emirates. The results of the research showed that travel time and trip maker characteristics are the most important factors that affect the choice (Obaid & Hamad, 2020). Jonii, (2015), tried to develop a model to understand the behavior of trip makers and try to convert them to public transportation using the logarithm model and SPSS statistical package. The results showed that the main factors influencing the mode choice are the cost and travel time (Mohammed & Joni, 2015).. Kedia et al., (2015): developed a model for selecting the mode of urban transportation for educational trips in India using fuzzy logic inference. The study found that people with high and middle incomes use the school bus more, while those with low incomes use walking and rickshaw more (Schmöcker, Quddus, Noland, & Bell, 2008). Pulugurta et al., (2015): developed a four-stage travel demand prediction model using the fuzzy logic technique and then compared the results with the traditional multinomial logit model. The results of this study showed that using fuzzy logic in demand modeling produced more accurate results than using traditional models(Pulugurta, Madhu, & Kayitha, 2015). Pulugurta, (2013): suggested using fuzzy logic to explain mode choice and comparing the results to the classic Multinomial logit model (MNL). The study concluded that travel time inside the mode and travel costs are the most influential factors on mode choice, and the findings of the fuzzy logic model were shown to have a greater prediction accuracy than the results of the traditional model. The fuzzy logic model was 70 percent accurate in predicting all modes, whereas the multinomial logit model was only 40 percent accurate in predicting two modes, the bus and the two-wheeled bus (Bai & Wang, 2006).

2. Methodology

2.1 Study Area

The study area is Ramadi city, which is located in the southeast of Anbar Governorate, and it is the political and administrative center in it. The basic design area of the city is estimated at about 143335 km^2 and it consists of thirty residential neighborhoods distributed over the area of the basic design map of the city. For the purpose of study and data collection, it was divided into traffic analysis zones as shown in Figure 1, which illustrate the location map and traffic zones of Ramadi City. A city is located between latitudes (33^o-27^o) and (33^o-23^o) N and longitudes (43^o-20^o) and (43^o-12^o) E. Ramadi city is about 100 kilometers west of Baghdad, 160 kilometers east of Haditha, 270 kilometers north of Al-Qaim, and 60 kilometers away from Heet, 300 kilometers west of Al-Rutba (Awad, Mohammed, & Mahmood, 2010). The study area is characterized by an important geographical location due to its association with neighboring areas, and this location led to the emergence of continuous transportation throughout the year

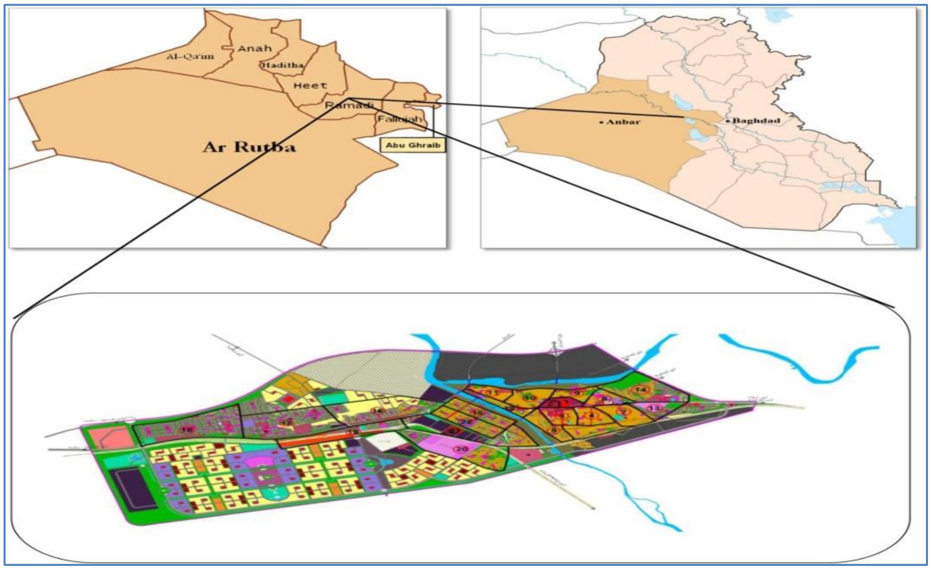


Figure 1: Location Map of Study Area (Ramadi City)

2.2 Study Area Boundary and Traffic Zones

For the purpose of obtaining data on the future travel demand and to enable us to link information about travel activity, travel and transportation between the study area zones, the study area was divided into traffic analysis zones (TAZ). The total number of traffic zones was set as 28 traffic zones, 22 were internal traffic zones and 6 external traffic zones as shown in Figure 2. The study was conducted on 26 zones only because zone 19 and zone 21 represent the expansion area and industrial zone respectively, because there are no household in these zones. The imaginary line representing the boundary of study area is termed as Cordon Line and the survey done inside the area covered by cordon line to study travel pattern to large extent is known as Cordon Line Survey. The area inside cordon line is studied extensively whereas the area outside cordon line is studied in a lesser degree of detail**.**

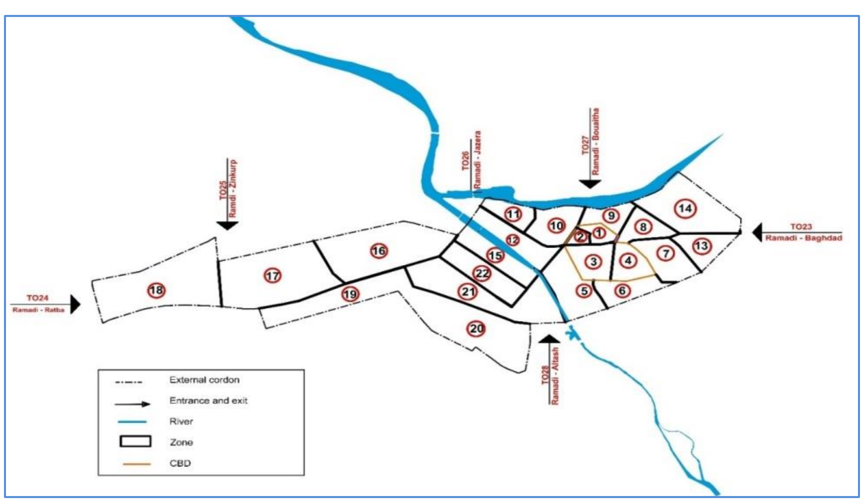


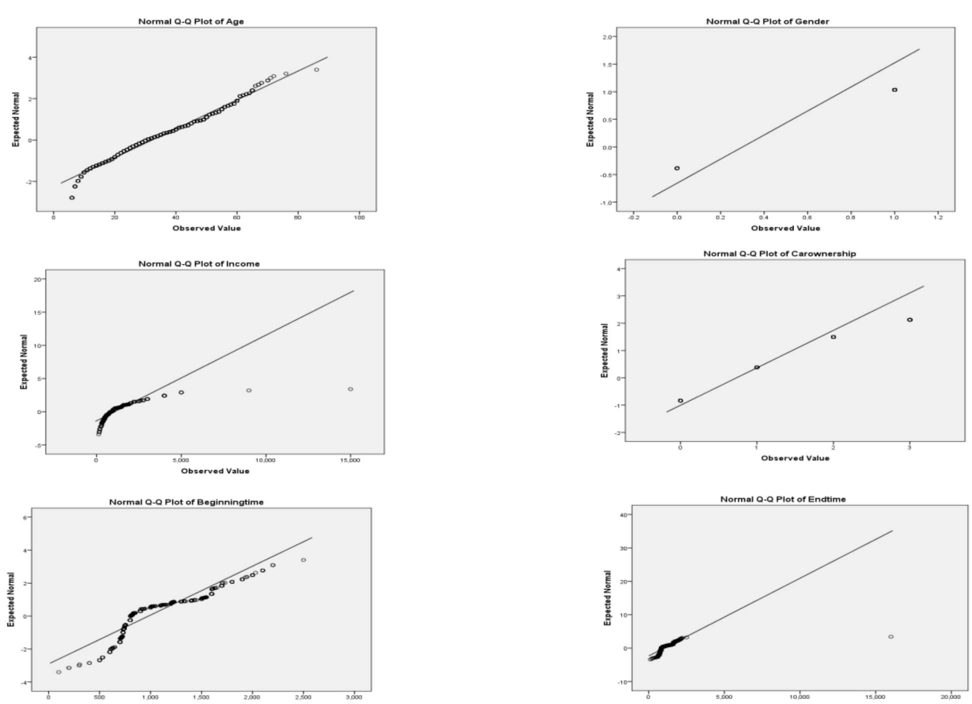
Figure 2: Map of Zones of Study Area

2.3 Data Collection

The aim of the field survey for this study is to collect data by taking a random sample representing a group of housing units that are used to represent the largest number of residents with specific characteristics. The field survey was conducted in this study through direct home interviews with family members in their places of residence within the traffic zones of Ramadi city. Data can be collected in many ways, and one of the most important and most common methods is surveys. The household interview questionnaires are characterized by generating a high response rate between (70-80%) and providing accurate data.
Sample size was chosen at random, calculated using the appropriate statistical procedure, then distributed to the household. The size of the sample to be interviewed depends on the total population, the required degree of accuracy, and the density of the population in the study area, The statistical formula used to calculate the sample size is as follows, which can be found in most statistical textbooks (Bluman, 2012)

$$\text{n}= X^2 \times N \times P(1-P) \div \{ME^2\} \times (N-1) + \{X^2 \times\ P \times (1-P)\} \quad \ldots\ldots\ldots..(\mathbf{1})$$
$$\text{n}= 3.84 \times 73215 \times 0.25\ \div \{(0.000625 \times 73215) + (3.84 \times 0.25)\}$$
$$\text{n} = 1504.4$$
$$\text{n} = 1505\ \ Sample$$

where n is sample size; N is population size; p is population proportion (equal to 0.5); X^2 is the Chi- square at 1 degree of freedom for the specified confidence level. This computation is based on a normal Gaussian distribution with more than 30 samples, according to the assumption. And according to margin of error (2.5%) 95 % confidence level, and 50 percent response distribution, the population size in this study is (73215) households. As a result, the sample size is 1505 samples. According to the variables that were adopted in building the model, the questionnaire includes a question about the following items; first, characteristics of trip makers such as age, gender, income per household, car ownership per household, and family size. Second Trip characteristics such as: Trip purpose (work, shopping, education, recreation, other), trip origin and destination, and the beginning and end time of trip. Third; mode characteristics such as: type of mode used. the cost of trip, waiting time, time inside mode. The modes of travel available were classified as follow: private vehicle, private bus (includes school buses and starex), public bus, government bus, Taxi, Motorcycle, bicycle, and walking.



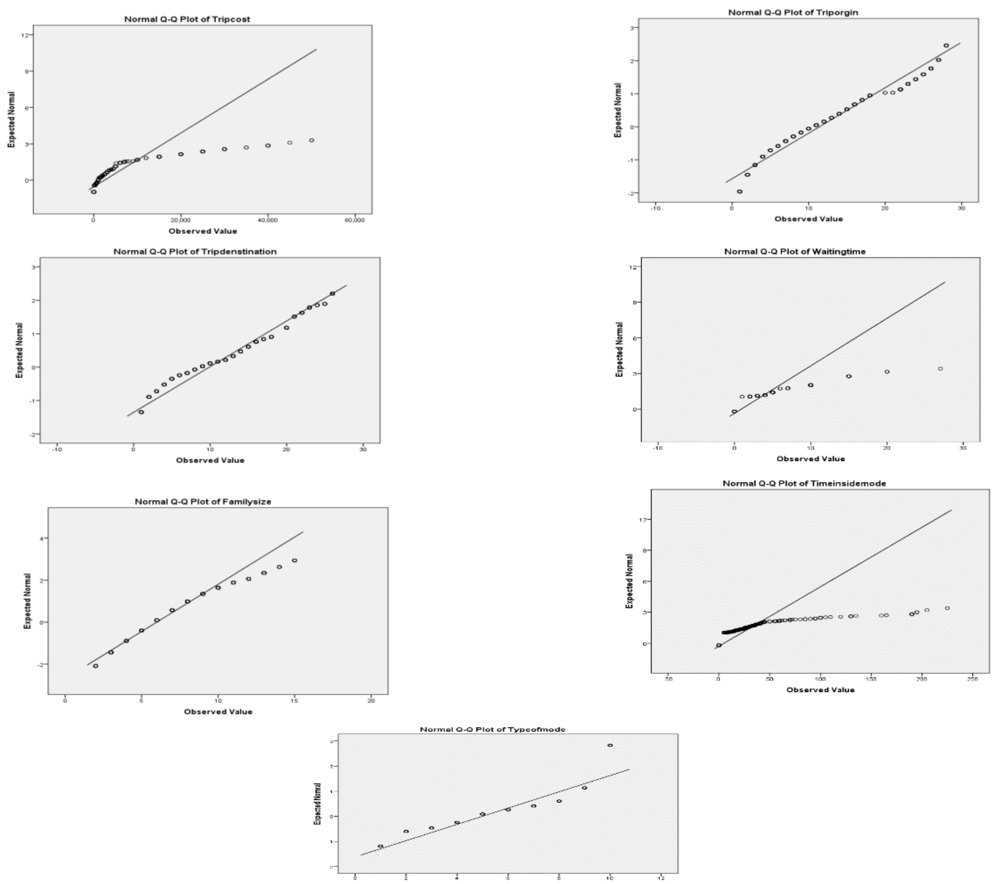


Figure 3: Normal Test for All Data

2.4 Modal Split Modelling Using Fuzzy Inference System (FIS)

The idea of fuzzy logic was invented in 1965 by the Azerbaijani scientist Lutfi Zadeh from the University of California, Lutfi developed fuzzy logic to use as a better way to process data**.** Fuzzy logic differs from traditional classical logic, where classical logic takes constant values of either zero or one and the element is either an element in the set or not an element in it, unlike the fuzzy logic that takes values between zero and one and these values represent the degree of membership of the element to a particular group (Ross, 2005) . Fuzzy logic is a methodology or method for solving problems without the need for an identifier of their mathematical origin, because fuzzy logic can deal with numerical data and linguistic knowledge at the same time, unlike the crisp classical logic that deals with numerical data only. Fuzzy logic is characterized by its ability to solve complex, ambiguous problems. It creates systems close to the spirit of human thought by understanding the behavior of the system(Jang, 1993) . Fuzzy rules are created using human knowledge and experience and uses the conditional IF-THEN rule similar to human behavior and knowledge where linguistic terms are used. Fuzzy IF-THEN rule links an output or conclusion to a condition defined using linguistic variables and fuzzy sets. The IF section is mostly used to capture knowledge through elastic circumstances, while the THEN part can be used to provide a conclusion or output in linguistic variable form.

2.5 Membership Function Generation

A fuzzy set enables a member to have a partial degree of membership, which can be mapped into a function or a universe of membership values. Assume we have a fuzzy set A, and if an element x is a member of this fuzzy set A, we can denote this mapping as:

$$\mu A(x): X \rightarrow [0, 1]$$

Where $\mu A(x)$ = 1 if x is totally in A; $\mu A(x)$ = 0 if x is not in A; $0 < \mu A(x) < 1$ if x is partly in A. The membership function is represented or equaled by the degree of membership, or membership value $\mu A(x)$ for an element x of set A and has a range from 0 to 1. This graphical expression includes a variety of shapes, including triangular, trapezoidal, Gaussian, and so on. There are a variety of approaches for learning from numerical data to build membership functions for each variable. The k-means clustering approach is used to build the membership functions of input and output variables in this study. The primary idea behind this clustering technique is to choose initial cluster centers at random. These initial cluster centers are adjusted over a number of generations to fit the data clusters as closely as feasible(Seetharaman, Errampalli, Mukhopadhyay, & Gangopadhyay, 2009) . For all input variables and output variables, triangle membership functions are developed in this study.

2.6 Fuzzy Rule Generation

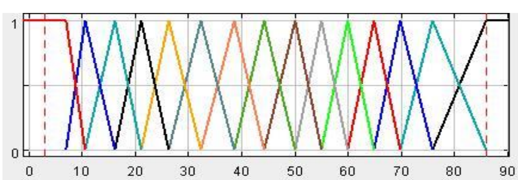
The data collected during home interview surveys is crisp data that needs to be fuzzified muddled through appropriate membership functions. It is assumed that the trip maker decides the usefulness of each mode based on the influencing variables and thus uses some simple rules from these variables. The fuzzy rules are derived either from expert knowledge or from numerical data, depending on the situation (Gou, Hou, Chen, Wang, & Luo, 2015). This work describes the generation of classification rules for a classification model that is challenging and complex due to the fact that it consists of eleven inputs and one output. Consequently. The fuzzy inference professional (FISPro) application is used to resolve the issue. FISPro (Fuzzy Inference System Professional) permits to generate fuzzy inference systems and to procedure them for reasoning purposes. This software was developed by (INRAE) National Research Institute for Agriculture, Food and Environment. The website of this software is (https://www.fispro.org/en). The fuzzy rules are produced using the Wang and Mendel (WM)approach, for each input and output, this method necessitates the use of specified fuzzy membership functions. It has the ability to generate rules from data on its own. Starting with the training set, it generates one rule for each data pair in the training.

3. Results

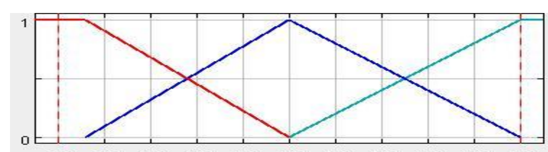
The mode choice model based on fuzzy logic is proposed to estimate the mode selection behaviour of individuals by looking at the variables affecting the mode choice, namely age, gender, family income, car ownership, family size, travel cost, start and end time of the trip, and origin and the destination of the trip. Figure 3 shows membership function for all input and output. The x-axis in these graphs shows the variables for each input and type of mode for output, while the y-axis is a membership function (MFs) that ranges from 0 to 1. The fuzzy rules for the general modal split model are shown in Table 1. These fuzzy rules are generated automatically as a result of data case learning. Where the X_1 represents the age variable, X_2 represents gender, X_3 represents income, X_4 represents vehicle ownership, X_5 represents the beginning time of the trip, X_6 represents the end time of the trip, X_7 represents the cost of the trip, X_8 represents trip origin, X_9 represents trip destination, X_{10} represents the waiting time, X_{11} represents the time spent inside the mode of transport, X_{12} represents the size of the family, and Y_p represents the dependent variable, which is the type of mode used. After generating membership functions, the performance of the modal split model for the whole trips is tasted. The relationship between the observed type of mode Yp and the fuzzified type of mode is shown in Figure 4. The observed and estimated mode type had a high correlation of about 93.1 percent. The determination coefficient R^2, as well as the root mean absolute error (MAE) and mean square error (RMSE), are shown in Table 2 to illustrate the amount of agreement of the type of mode values estimated by the fuzzified modal split model.

Table1- IF-Then Rules Generated by FIS for Whole Trips

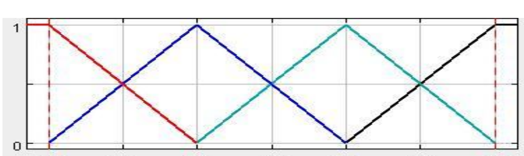
Rule	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	Y_P
1	D	A	A	A	E	A	A	A	A	A	A	K
2	E	A	A	A	E	A	B	A	H	A	A	G
3	E	A	C	C	F	B	A	A	A	A	A	K
4	H	A	C	C	E	A	A	A	H	A	A	B
5	H	A	C	C	I	C	A	A	A	A	A	K
6	H	A	B	B	E	A	A	A	B	A	A	B
7	B	A	B	B	E	A	A	A	C	A	A	G
8	C	A	A	A	E	A	A	A	D	A	A	G
9	H	A	A	A	G	C	C	A	F	B	B	F
10	H	A	D	C	E	A	A	A	E	A	A	B
11	L	A	D	C	G	C	A	A	A	A	A	K
12	D	A	D	C	E	A	B	A	H	A	A	G
13	C	A	D	C	H	B	A	A	A	A	A	K
14	F	A	B	A	E	A	C	A	G	A	A	C
15	B	A	B	A	E	A	A	A	A	A	A	K



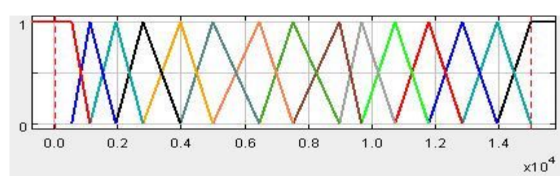
Age



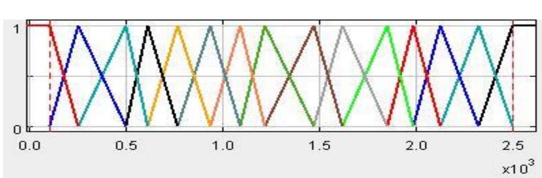
Gender



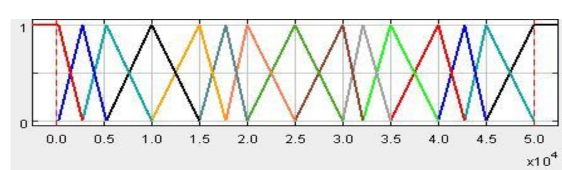
Car Ownership



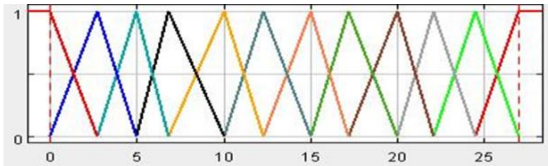
Income



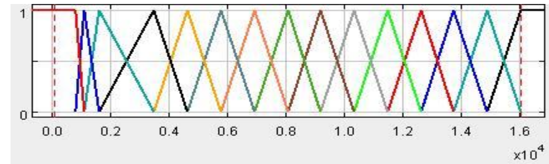
Beginning Time



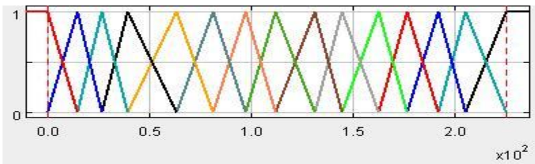
Trip Cost



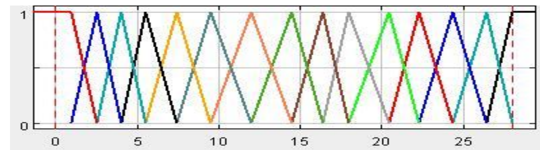
Waiting Time



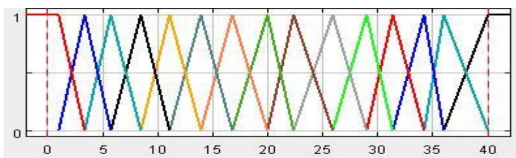
End Time



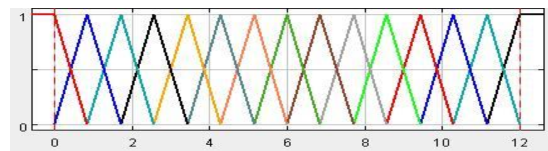
Time Inside Mode



Trip Origin



Trip Destination



Type of Mode

Figure 4: Membership Functions for General Modal Split Model

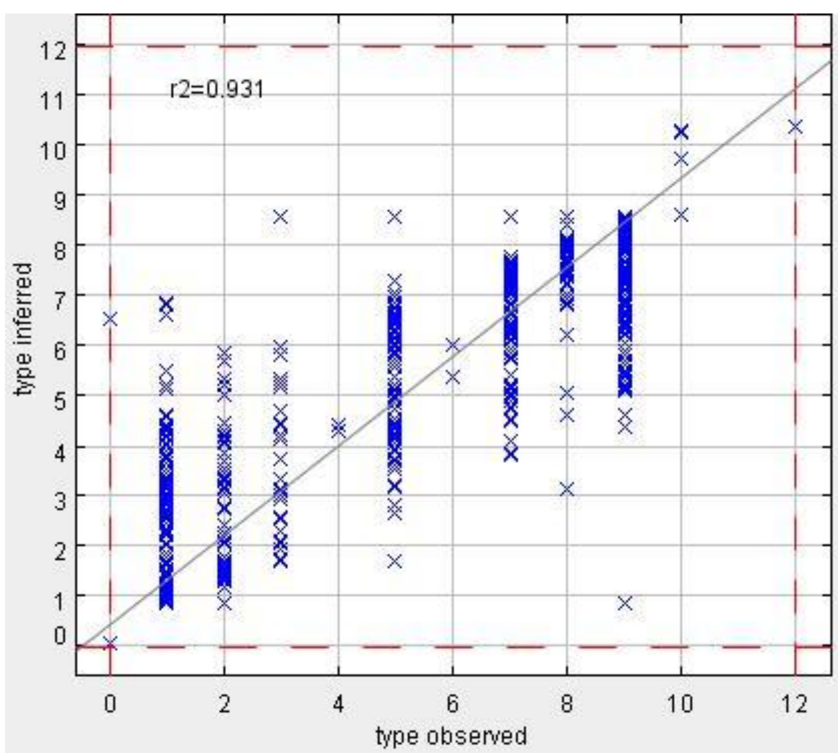
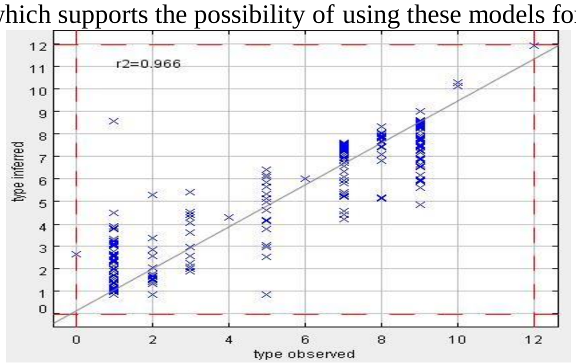


Figure 5: The Performance of General Modal Split Model by FIS

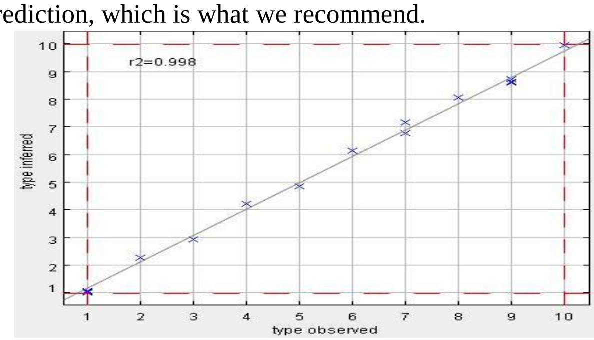
Table 2- The Improvement of Model Performance

Model	R^2 (%)	RMSE	MAE
Whole Trips	93.1	0.82	0.472

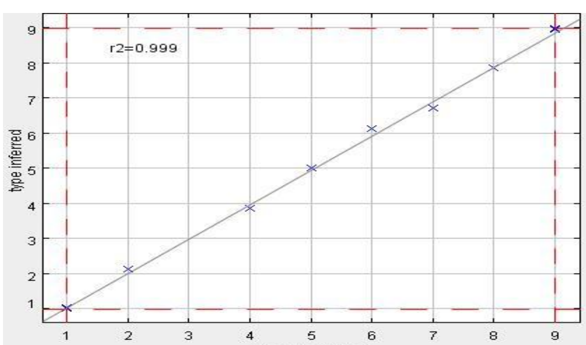
It is important to identify mode choice based on its purpose when developing transportation policy. Five purposes of the trip which are work, shopping, education, recreation, and others trips are considered in this research. Therefore, five modal split models are developed based on trip purpose using fuzzy inference system (FIS). These models are work, education, shopping, recreational, and others models. Figure 5 shows relation between the observed and estimated type of mode for five trip purpose. Table 3 shows the value of R^2, RMSE. It can be shown that the five modal split models achieve a high correlation. These models gave a strong correlation coefficient for the independent and dependent variables. Whereas, the correlation coefficient for work trips is 96.6%, for shopping trips 99.8%, for education trips 98.9%, for recreation trips99.9%, and 99.7% for other trips, which supports the possibility of using these models for future prediction, which is what we recommend.



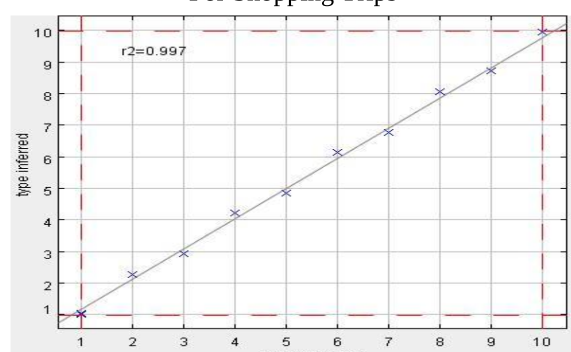
For Work Trips



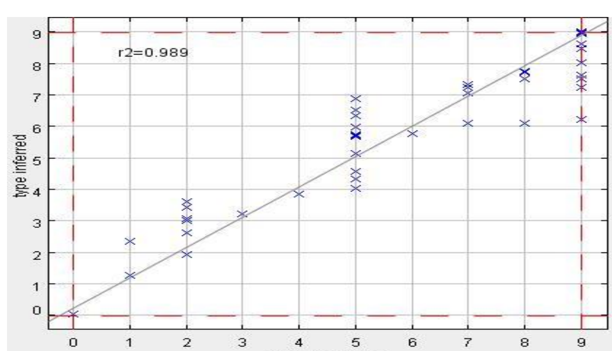
For Shopping Trips



For Recreation Trips



For Education Trips



or Other Trips

Figure 6: The Performance of Modal Split Model by FIS for Five Trip Purpose

Table 3- The Performance of Modal Split Model for Different Trip Purpose

Model	R^2 (%)	RMSE	MAE
Whole Trips	93.1	0.823	0.472
Work Trips	96.6	0.619	0.363
Shopping Trips	99.8	0.203	0.177
Education Trips	98.9	0.276	0.148
Recreation Trips	99.9	0.136	0.111
Other Trips	99.7	0.196	0.172

When these results were compared with the results of multiple linear regression using the SPSS statistical analysis program, the results showed a weak correlation between the variables as shown in Table 4, and it was found that the most influential factors on the mode choice are vehicle ownership, age and trip cost, unlike the fuzzy logic that proved that all factors affected on mode choice with very high correlation coefficients, which can be adopted in future prediction.

Table 4- Regression Results for The General Modal Split Model Using Regression Analysis

Model	R^2 (%)	RMSE	Std. Error
Whole Trips	0.537	0.286	2.606

The prediction accuracy of the fuzzy logic and regression model was estimated using the number of data that used for validation of these models. For both models 85% of data were selected to build the models, whereas 15% of data were selected as validation data. The prediction accuracy for fuzzy logic and regression model are represented in Table 5.

Table 5- Validation of Fuzzy Logic and Regression Model

Model	Prediction accuracy (%)	
	Fuzzy logic	Regression analysis
Whole Trips	88	37
Work Trips	79	29
Shopping Trips	83	31
Education Trips	73	30
Recreation Trips	77	27
Other Trips	68	24

Conclusions

1-The results showed that the fuzzy logic models gave a strong correlation coefficient for the independent and dependent variables. Whereas, the correlation coefficient for Whole trips was 93.1%, for work trips is 96.6%, for shopping trips is 99.8%, for education trips 98.9%, for recreation trips 99.9%, and 99.7% for other trips, which supports the possibility of using these models for future prediction, which is what we recommend.

2- The fuzzified modal split model for recreation trips is the most accurate and reliable model in estimation mode choice in Ramadi City.

3-The results of the analysis showed that the fuzzy logic overlapped the results of the multiple linear regression, as the fuzzy inference system proved that all factors affected the choice of the mode of transport, with very high correlation coefficients between the variables.

References

Awad, H. A., Mohammed, H. A., & Mahmood, W. M. (2010). Evaluation and improvement of traffic operation for Al-Zeoat intersection in Al-Ramadi city. *AJES, 3*(2), 46.

Bai, Y., & Wang, D. (2006). Fundamentals of fuzzy logic control—fuzzy sets, fuzzy rules and defuzzifications. In *Advanced fuzzy logic technologies in industrial applications* (pp. 17-36): Springer.

Bluman, A. G. (2012). *Elementary statistics: A step by step approach*: McGraw-Hill.

Chakroborty, P., & Das, A. (2017). *Principles of transportation engineering*: PHI Learning Pvt. Ltd.

Georgieva, P. (2016). Fuzzy rule-based systems for decision-making. *no. May, 2018*.

Gou, J., Hou, F., Chen, W., Wang, C., & Luo, W. (2015). Improving Wang–Mendel method performance in fuzzy rules generation using the fuzzy C-means clustering algorithm. *Neurocomputing, 151*, 1293-1304.

Jang, J.-S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics, 23*(3), 665-685.

Minal, & Sekhar, C. R. (2014). Mode Choice Analysis: The Data, the Models and Future Ahead. *International Journal for Traffic and Transport Engineering, 4*, 269-285.

Mohammed, A. A., & Joni, H. H. (2015). Evolution the transportation mode from private cars to publics by logit method: a case study in Baghdad. *age, 127*, 2.

Obaid, L., & Hamad, K. (2020). *Modelling Mode Choice at Sharjah University City, United Arab Emirates.* Paper presented at the MATEC Web of Conferences.

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.

Pulugurta, S., Madhu, E., & Kayitha, R. (2015). Fuzzy logic–based travel demand model to simulate public transport policies. *Journal of Urban Planning and Development, 141*(4), 04014044.

Ross, T. J. (2005). *Fuzzy logic with engineering applications*: John Wiley & Sons.

Schmöcker, J.-D., Quddus, M. A., Noland, R. B., & Bell, M. G. (2008). Mode choice of older and disabled people: a case study of shopping trips in London. *Journal of Transport Geography, 16*(4), 257-267.

Seetharaman, P., ERRAMPALLI, M., MUKHOPADHYAY, D., & GANGOPADHYAY, S. (2009). *Comparative Evaluation of Mode Choice Modelling by Logit and Fuzzy Logic.* Paper presented at the Proceedings of the Eastern Asia Society for Transportation Studies Vol. 7 (The 8th International Conference of Eastern Asia Society for Transportation Studies, 2009).