Abstract number: 167

PREDICTION OF FUTURE TRIPS USING FUZZY LOGIC BASED TRIP GENERATION MODEL

Sarada Pulugurta [1], Madhu Errampalli [2], K.Ravinder[3]
[1] *(Trainee Scientist, Academy of Scientific and Innovative Research, CSIR-CRRI, sarada.pulugurta09@gmail.com)*
[2]*(Principal Scientist, Transport Planning Division, CSIR-Central Road Research Institute, errampalli.madhu@gmail.com)*
[3]*(Principal Scientist, Transport Planning Division, CSIR-Central Road Research Institute,* krrcrri@gmail.com)

ABSTRACT: Travel Demand Forecasting is an essential to design and develop any new transportation system. The first stage of conventional four stage travel demand forecasting is trip generation. Based on household interview conducted in Port Blair city, appropriate influencing variables were selected and accordingly linear regression model was developed using SPSS software. Limitation of traditional linear regression model is that they do not take into account subjectivity, imprecision, ambiguity and vagueness of human minds hence there was a need to develop an accurate and realistic model which could take into account linguistic variables and also does not require the precision of crisp input. Considering this aspect fuzzy logic model was developed using MATLAB software and trips produced and attracted to all zones were predicted using built in functions which uses subtractive clustering technique. Results obtained from both traditional linear regression and fuzzy logic models were compared using statistical parameters namely R^2 and Root Mean Square Error (RMSE) values. It was observed that the results obtained from fuzzy logic model gave better prediction accuracy in comparison to the traditional regression model. Thus it can be concluded that the fuzzy logic models were better able to capture and incorporate the human knowledge and reasoning into trip generation modeling. Further, developed fuzzy logic based trip generation models are applied to predict future trips for year 2021 and 2031.

Keywords- Fuzzy Logic; Regression analysis; trip generation; subtractive clustering

1 .Introduction

The basic objective of transportation planning and management is to match transportation supply with travel demand. Travel demand forecasting is an essential tool for computing future demand. Conventional four stage modeling is used for travel demand forecasting. The four stage modeling involves trip generation, trip distribution, mode choice and traffic assignment. The trip generation aims at predicting the total number of trips generated and attracted to each zone of the study area. Trip Distribution estimates number of trips between origin and destination zone. Mode choice analysis deals with choice of mode to reach respective destination through public, private modes and last stage involves assignment of traffic on routes selected based on shortest path. This study focuses on first stage of demand modeling i.e trip generation. Traditional methods of trip generation are linear regression model, cross classification etc which estimate number of trips based on certain influencing parameters. The limitations of traditional models are: it does not take into account subjectivity, ambiguity, imprecision and vagueness involved in those influencing variables. Moreover, the relationship may not be in linear form under the influence of so many variables which in turn affects the output of trips. Thus to overcome limitation of traditional model, application of one of the artificial intelligence technique like fuzzy logic seems to be most appropriate to consider ambiguity, imprecision and vagueness in influencing variables considered . Advantages behind using fuzzy logic model are firstly linguistic variables are taken into account of socio-economic variables like income, age, vehicle ownership etc. and secondly fuzzy logic takes into account of highly non-linear relationship between output and input compared to traditional models as fuzzy logic uses rule base (IF…THEN…rules) which has a close resemblance with human knowledge and behavior considering linguistic terms and are capable of handling complicated real life situation.

Considering this, the objectives of the present study is formulated to develop trip generation model using fuzzy logic technique. For this purpose, Port Blair City in India has been considered for data collecting through household survey. Subsequently fuzzy logic based models are compared with the traditional models. Using the developed fuzzy logic based generation model trips have been forecasted for year 2021 and 2031. The paper is organized as follows. Section 1 introduces the concepts and gives a brief detail about the limitations in modeling technique in trip generation analysis. A brief literature review forms the background for the study in Section 2. Section 3 describes the study area and the data collection procedure. Section 4 lays down the methodology adopted in the study in relation to variable selection and model development. Section 5 gives the results of model formulation followed by detailed analysis of the results. In Section 6, the results from developed fuzzy logic model for the study area. Finally, Section 7 gives the conclusions that can be drawn from this study.

2. Literature Review

This section provides the details about the different studies conducted in developing trip generation models to forecast the future trips. The study conducted by Joseph and Makinde (2009) through household survey in Ado Ekiti township located in Nigeria. The study area was divided into three zones and the zonal parameters considered in their study were income, number of autos, age of workers and family size. Regression analysis was carried out in SPSS for home based other trips, home based work trips and non home based trips and it was observed from the results that home based trips for other purposes were large in comparison with home based work trips and non home based trips. It was also observed that maximum numbers of trips were made by people of age group between 31-50 in all three zones. However the model developed for all three zones had low R^2 values of 0.370, 0.378, and 0.475 respectively which indicate that traditional linear regression model was not capable of predicting estimated results close to the observed trips.

In order to incorporate both subjective and objective knowledge, fuzzy logic concept was introduced by Zadeh (1965). According to classical set theory a set is collection of objects having similar characteristics. In this theory an element either completely belongs to set or it does not so each member takes the value of either 0 or 1. Fuzzy set unlike crisp set assigns degree of membership for each of its elements which lie in the interval [0,1]. The belongingness of element to a set is partial and not complete. Thus a fuzzy set is able to capture vagueness of human mind as boundary of set is not a crisp one rather it is a vague one. Zadeh (1965) successfully introduced concept of fuzzy logic which is based on generation of vague algorithm using vague input to produce vague output. Pappis and Mamdani (1977) were the first to apply fuzzy logic in field of transportation engineering on fuzzy controlled traffic signals that showed path for researchers to explore more in depth in this field. Xu and Chan (1993) applied fuzzy logic technique to estimate origin-destination matrix as it was observed that there was heavy volume of bicycle due to which traffic counts were found to be poor. Wang and Mendel (1992) developed procedure to generate fuzzy rule base from numerical data which involved fuzzification of input and output and generating a rule base followed by combining it with linguistic rules of human expert and finally defuzzifying the combined rule base to generate crisp output. Kalic' and Teodorovic' (1997) analyzed the problem of trip generation using fuzzy logic. The procedure proposed by Wang and Mendel (1992) was adopted for fuzzy rule base generation. Trips were also estimated from multiple linear regression and artificial neural network and it was observed that the fuzzy logic approach gave the best results. Aggarwal (2012) developed fuzzy logic model based on three variables namely age, income and vehicle ownership using C++ language by developing a rule base comprising a total of twelve rules and triangular type of membership functions. Aggarwal (2012) developed fuzzy model based on Sugeno type of fuzzy inference system and compared results with traditional model of linear regression and cross classification and found greater variation between observed and estimated trips in case of fuzzy logic.

Based on the above review of different studies, it can be inferred that the traditional models in terms of linear relationships suffers from the high prediction accuracy and fuzzy logic technique outperforms the predictions under different scenarios in prediction of trips for base year and future years. Considering this, the present study proposes to develop fuzzy logic trip generation model and the data collection for this purpose is demonstrated in the next section.

3. Data Collection

Household survey was conducted in the city of Port Blair and its surrounding area in July 2011(CRRI, 2013). Port Blair, capital city of Andaman & Nicobar Islands, India has total population of 1.5 lakhs as per Census 2011. Total numbers of households are 31506 out of which data from 2627 were collected. Total numbers of one way household trips were observed to be 70206 per day. Per Capita Trip Rate (PCTR) for one way was found to be 0.55. The study area was divided into 24 zones out of which 18 zones were internal and 6 zones were considered as external. The six external zones are Dollygunj, New Pahargaon, Garacharma, Prothrapur, Austinabad and

Brookshabad. A preliminary analysis of household data was carried out in order to understand the travel behaviour of the commuters in the Port Blair city. Fig.1 shows distribution of trips by purpose.

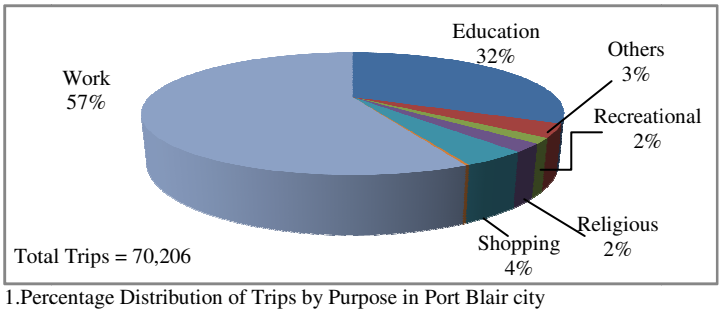

Fig 1.Percentage Distribution of Trips by Purpose in Port Blair city

From Fig. 1 it can be observed that maximum numbers of trips are work trips comprising of 57% followed by education i.e 32%. Shopping trips constitute 4% and religious and recreational trips are 2% each of total trips. Other trips constitute 3% of total trips. All types of trips are considered to develop model to predict total trips and the methodology adopted for this purpose is discussed in the next section.

4. Methodology

4.1 Trip Generation Using Linear Regression Analysis:

Trip generation generally comprises of trip productions and trip attractions. In the present study production trips are classified based on three purposes i.e work, education and other. Trips in each purpose are further classified based on vehicle ownership as having no vehicle (NV), car (C), two wheeler (TW) and cycle (Cy). Attraction trips are also classified based on purpose namely work, education and other. Correlation analysis was carried out to find the most influencing variables to develop relation with trip production and attractions. Out of all zonal parameters collected for productions and attractions and the selected influential parameters are mentioned below:-

Trip Productions:

Work Purpose:

(i) Population (P), (ii) Number of main workers (MW), (iii) Number of vehicle per household (VehPH), (iv) Average income per household (IncPH), (v) Number of earner per household (EPH), (vi) Number of households (HH) and (vii) Average age per household (AgePH)

Education Purpose:

(i) P, (ii) Number of literates (Lit), (iii) VehPH, (iv) IncPH, (v) Number of students per households (StuPH), (vi) HH and (vii) AgePH

Other Purpose:

(i) P, (ii) Number of literates (Lit), (iii) VehPH, (iv) IncPH, (v) HH and (vi) AgePH

Trip Attractions (for all purposes):

(i) P, (ii) Total Area of Zone (TA), (iii) Commercial Area of Zone (CA) and
(iv) Residential Area of Zone (RA)

A typical multiple linear function of the following form to estimate production trips having work purpose with no vehicle ownership is given below:-

$TP_{NV}^{W} = a_0 + (a_1 * \text{population}) + (a_2 * \text{number of main workers}) + (a_3 * \text{no. of vehicles per household}) + (a_4 * \text{average income per household}) + (a_5 * \text{number of earner per household}) + (a_6 * \text{number of households}) + (a_7 * \text{average age per household})$ (1)

TP_{NV}^{W} variable is trip production for work purpose by no vehicle group.

a_0 = constant term to be estimated.

a_i = coefficient of independent variables (i = 1, 2,3……n) to be estimated.

Similarly for all the groups and purposes, the multiple linear regression equations are considered.

4.2 Trip Generation Using Fuzzy Logic

Fuzzy logic takes into account human perception that brings imprecision and vagueness in traditional models and also uses rule base (IF…THEN…rules) which has a close resemblance with human knowledge and behavior as they use linguistic terms. The disadvantage of fuzzy logic method is formulation of rules requires expert knowledge and is quite tedious. Thus Fuzzy logic model is developed by dividing data set into groups called clustering. The process of clustering is done by three methods:-

- **Grid partitioning**: The input space is partitioned in number of grids where the fuzzy rules share membership functions for each input variable. However, this technique is also marred by the curse of dimensionality and hence clustering techniques are preferred.
- **Subtractive Clustering**: This technique assumes each data point to be a potential cluster centre and calculates a measure of the potential them based on the density of surrounding data points. The algorithm selects the data point with the highest potential as the first cluster centre and then delimits the potential of data points near the first cluster centre. This is continued for the next higher potential data points till all the data points are apportioned to one or the other cluster
- **Fuzzy c-means Clustering**: In this technique each data point belongs to a cluster with some degree of membership. The algorithm starts with an initial guess for the cluster centers, which are most likely incorrect. Then every data point is assigned a membership grade for each cluster. By iteratively updating the cluster centers and the membership grades for each data point, the correct mean location of each cluster is identified. This iteration is based on minimizing an objective function that represents the distance from any given data point to a cluster centre weighted by that data point‘s membership grade.

In this study Subtractive clustering is adopted as it is considered better than fuzzy c- means clustering based on study conducted by Bataineh et al (2011). Several trials for radius between 0.1 to 0.5 were made to get optimum solution and observed that 0.1 radius gave best results. MATLAB 2009 software was used to develop present fuzzy logic model and Genfis2 function which is based on subtractive clustering technique was adopted. In order to ensure reproducibility of results obtained from the model random numbers are generated using Mersenne Twister algorithm. Membership function considered for all input variables were gauss type and a typical membership function for the variable of population are shown in Fig.2. For the case of production trips with vehicle ownership as no vehicle for work purpose car input parameters considered were seven and output parameter was one as shown in Fig.3. Sugeno type of fuzzy inference system was generated for this purpose.

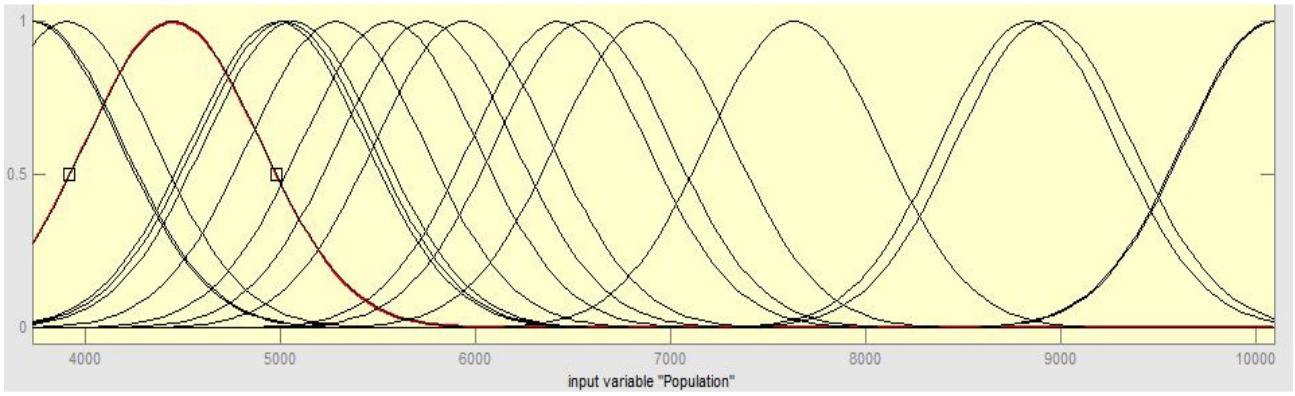

Fig 2. Typical Membership Functions of Input Variable Population for Work Purpose having No Vehicle Ownership Group

Twenty rules were formed for this exercise and some of rules are shown below:

Rule 1 IF (P is P-cluster1) AND (MW is MW-cluster1) AND (VehPH is VehPH-cluster1) AND (IncPH is IncPH-cluster1) AND (AgePH is AgePH-cluster1) AND (EPH is EPH-cluster1) AND (HH is HH-cluster1) THEN (TP^{w}_{nv} is TP-cluster1)

…..

…..

Rule 20: IF (P is P-cluster20) AND (MW is MW-cluster20) AND (VehPH is VehPH-cluster20) AND (IncPH is IncPH-cluster20) AND (AgePH is AgePH-cluster20) AND (EPH is EPH-cluster20) AND (HH is HH-cluster20) THEN (TP^{W}_{nv} is TP-cluster20)

Subsequently defuzzfication was carried out using weighted average method to obtain crisp value of output in terms of number of trips.

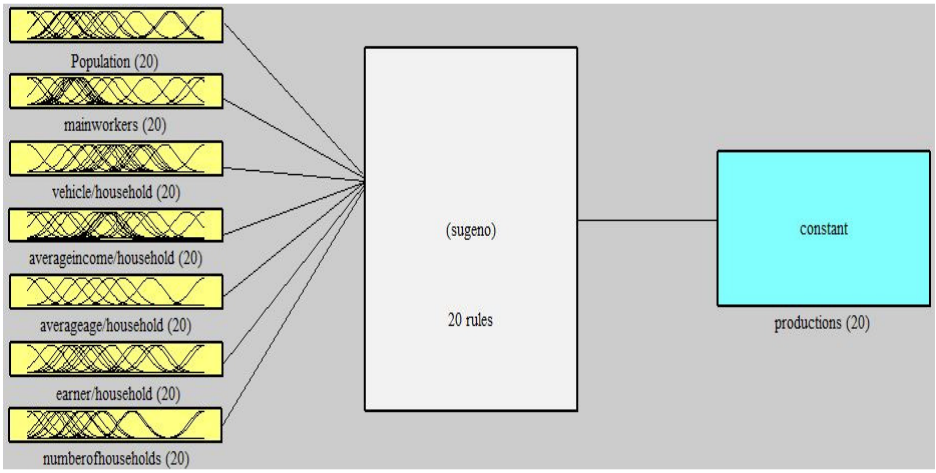

Fig 3 Fuzzy Inference System (FIS) structure for Trip Production for work purpose-no vehicle group.

5. Development of Trip Generation model

In this section, development of trip generation models through Linear Regression analysis and Fuzzy Logic model is discussed and comparison of both traditional linear regression and fuzzy logic model also carried out.

5.1 Traditional Linear regression models

Regression analysis was carried out in SPSS software and results are obtained as follows:-

Trip Productions (TP):

TP^{W}_{NV}=3507-0.025*P+0.046*MW-1972.683*VehPH+0.030*IncPH-66.452*AgePH+ 314.266*EPH+0.417*HH (2)

(6.951) (-0.379) (0.311) (-12.907) (4.329) (-6.346) (1.531) (3.477) R^2=0.96

TP^{W}_{C}=-1109+0.099*P-0.153*MW+666.367*VehPH+0.006*IncPH+2.046*AgePH+445.401*EPH-0.095*HH (3)

(-1.802) (1.248) (-0.853) (3.574) (0.670) (0.160) (1.769) (-0.650) R^2=0.78

TP^{W}_{TW}=-2507-0.023*P-0.072*MW+803.915*VehPH-0.020*IncPH+66.684*AgePH-650.797*EPH+0.862*HH (4)

(-3.206) (-0.218) (-0.296) (3.203) (-1.716) (3.874) (-1.920) (4.375) R^2=0.79

TP^{W}_{CY}=-62-0.010*P+1.121*MW+0.016*VehPH+35.4*IncPH-0.001*AgePH+12.323*EPH+0.027*HH (5)

(-1.192) (-1.517) (1.035) (1.072) (1.608) (1.953) (0.578) (2.190) R^2=0.37

TP^{E}_{NV}=1892-0.088*P+0.095*Lit-23.994*AgePH-1131.268*VehPH+0.010*IncPH-138.951*StdPH+0.366*HH (6)

(3.561) (-1.420) (1.218) (-2.551) (-7.247) (1.615) (-0.821) (3.662) R^2=0.91

TP^{E}_{C}=-801.616-0.036*P+0.057*Lit-1.731*AgePH+282.960*VehPH+0.010*IncPH+353.220*StdPH+0.021*HH (7)

(-1.865) (-0.731) (0.898) (-0.228) (2.243) (1.863) (2.581) (0.262) R^2=0.62

TP^{E}_{TW} =1539-0.275*P+0.281*Lit+18.744*AgePH+146.632*VehPH+0.004*IncPH+194.868*StdPH+0.726*HH (8)

(-2.575) (-3.791) (3.211) (1.733) (0.836) (0.538) (1.024) (6.458) R^2=0.84

TP^{E}_{CY} = 69+0.002*P-0.007*Lit+1.359*AgePH+27.789*VehPH-0.003*IncPH-46.320*StdPH+0.011*HH (9)

(0.466) (0.128) (-0.306) (0.522) (0.643) (-1.857) (-0.989) (0.411) R^2=0.25

TP^{O}_{NV}=1878-0.002*P-0.010*Lit-37.677*AgePH-978.885*VehPH+0.022*IncPH+0.194*HH (10)

(2.710) (-0.023) (-0.085) (-2.596) (-4.435) (2.253) (1.259) R^2=0.67

TP^{O}_{C}=0.005*Lit-2.423*AgePH+28.990*VehPH+0.004*IncPH+0.008*HH (11)

(0.408) (-1.994) (0.501) (1.236) (0.174) R^2=0.47

TP^{O}_{TW} = -0.099*P+0.122*Lit-1.950*AgePH-255.414*VehPH+0.006*IncPH+0.197*HH (12)

(-1.231) (1.211) (-0.611) (-1.699) (0.752) (1.565) R^2 =0.54

TP^{O}_{CY} = -0.005*P+0.005*Lit+0.151*AgePH+5.268*VehPH+0.007*HH (13)

(-2.615) (1.856) (1.899) (1.374) (2.319) R^2 =0.52

Trip Attractions (TA):

TA^{W}= -0.077*P-0.399*A+1482.214*CA-8.712*RA (14)

(-0.936) (-0.239) (8.664) (-1.123) R^2=0.88

TA^{E} = 0.026*P-0.030*A+508.632*CA-1.281*RA (15)

(0.385) (-0.022) (3.638) (-0.202 R^2=0.70

TA^{O} = -0.028*P-0.041*A+268.698*CA-1.186*RA (16)

(-1.192) (-0.088) (5.559) (-0.546) R^2=0.71

Note: Values in the paranthesis represents t-values of constants and coeefficents

Subscripts NV, C, TW and CY represents No Vehicle, Car, Two Wheeler and Cycle groups of vehicle ownership

Superscripts W, E and O represents Work, Education and Other Purposes

5.2 Fuzzy logic models:

Fuzzy logic models have developed using MATLAB software as described in Section 4.2 for both trip productions and attractions. Traditional Linear regression model was compared with fuzzy logic model using two statistical means of comparison namely:-

1) Root Mean Square Error (RMSE)
2) R^2 value

1) Root Mean Square Error (RMSE): Root mean square error is defined as square root of average of deviations from the observed values.

$$RMSE = \left(\sum_{i=1}^{n}(Oi - Ei)^2/n\right)^{1/2} \quad (17)$$

O_i= i^{th} observed trips
E_i =i^{th} estimated trips
n = number of observations

2) R^2value: the observed and estimated values are plotted and R^2 value of such plots indicates efficiency of model in terms of how close the estimated values are to observed values.

The formulated fuzzy logic model was used to estimate the results in terms of R^2 and RMSE values. It can be observed from the results that R^2 values are in range of 0.81 to 0.99 for different groups in case of trip productions. In case of trips attractions, R^2 values are 0.97, 0.99 and 0.99 for work, education and other purpose trips respectively. The RMSE lies in range of 2.9 to 74.5, 18.6 to 30, 11 to 62 and 4 to 4.2 for no vehicle, car, two wheeler and cycle groups respectively in case of trip productions. For trips attractions, RMSE was observed to be 345.6, 51.3 and 33.6 for work, education and other purpose trips respectively. Fig.4 shows the comparison of results from fuzzy logic models with linear regression models and it was found that the incorporation of fuzzy logic in trip generation models caused an improvement in R^2 from 0.84 to 0.96 for productions and from 0.801 to 0.987 in case of attractions. Though variables adopted in this study are observed to be continuous in nature and regression model perform better while dealing with continuous variables and regression models are quite flexible and can mimic complex IF..THEN rules but in this study for the explanatory variables adopted by us obtained from the primary and secondary sources it was observed that higher accuracy was obtained from fuzzy logic model as compared to the

traditional linear regression model as can be observed from the value of R^2 and RMSE values obtained from both models.

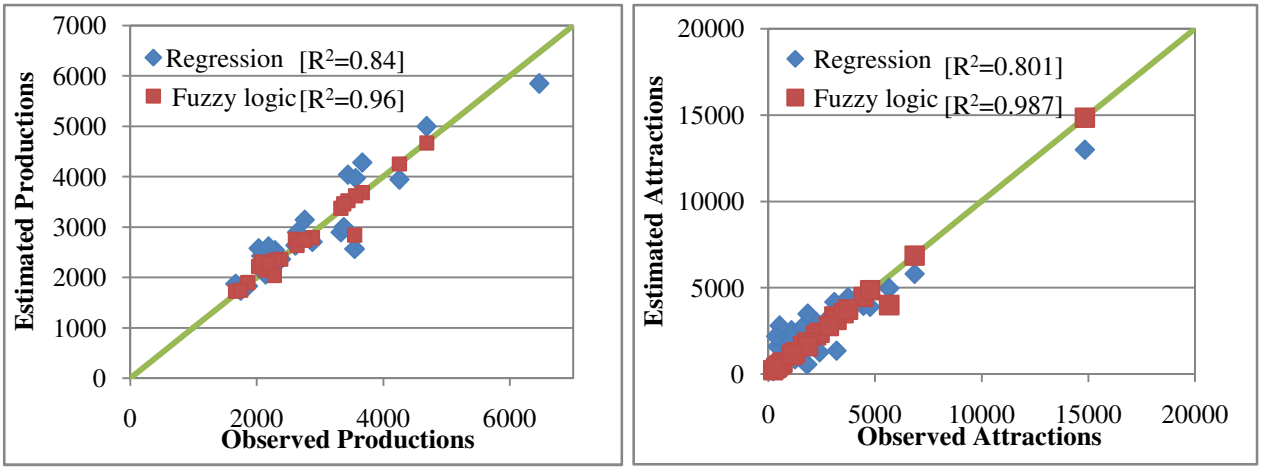

Fig 4.Comparison of Observed and Estimated Productions and Attractions from Linear Regression and Fuzzy Logic Model

6. Forecasting of Trips for the Years 2021 and 2031

Developed fuzzy logic model was applied for forecasting the demand for years 2021 and 2031. Out of all the variables listed in Section 4.1, the population data over the years is available and limited data for the other variables namely number of households, literates, vehicles per household etc. is also available. Using the growth factors for these variables, growth factors for other variables have been appropriately considered. From this exercise, the total trips productions and attractions for the year 2021 and 2031 are estimated from the fuzzy logic models as shown in Fig. 5 below :-

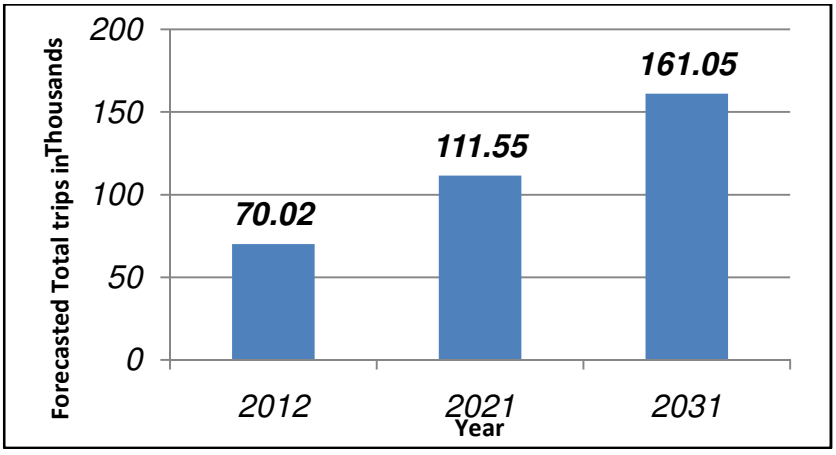

Fig 5 Forecasted Trips for the Years 2021 and 2031

From the Fig. 5, it can be observed that the trips are going to be increased enormously as the population and other economic parameters are set to grow in the Port Blair city. The trips in 2021 will be around 1.1 lakh and in 2031, a total of about 1.75 lakh trips will be generated. The increase in total trips for year 2021 is about 60% with respect to base year and for year 2031 is about 52% with respect to 2021. Considering the estimated and forecasted number of trips certain transport policies, for Port Blair city, could be framed in such a way to as to reduce congestion on the present and future road network.

7. Conclusions

From analysis of the present study with respect to fuzzy logic models, it can be clearly seen that after application of fuzzy logic to the trip generation model, the accuracy of results have improved to large extent. The R^2 value are better in all fuzzy logic models when compared with R^2 values obtained from traditional models and

RMSE values have decreased drastically in all fuzzy logic models when compared with traditional models. The value of R^2 and RMSE both indicate that fuzzy logic models are more efficient in comparison with traditional linear regression models as fuzzy logic models are able to estimate trips more close to observed trips.
Thus from this study the following inferences can be drawn:-

- Fuzzy logic models are able to capture vagueness of input variables and hence predict trips close to the observed values.
- Fuzzy logic models have more efficiency when compared to traditional linear regression models.
- For the Port Blair study the influencing variables for production model for work purpose influencing variables are average age per households, average income per households, average vehicles per households and number of households and for education purpose are average age per households, number of vehicles per households, and number of households and for other purpose are average age per households, number of vehicles per households and average income per households.
- For attraction model the influencing variables are commercial area and population for work and education purpose and commercial area alone in case of other purpose.
- Total growth in trips for 2021 from base year 2011 is about 60 % and total % growth for the year 2031 from 2021 is about 52%.

Prediction through fuzzy logic model indicates that the limitations of this study are that decision tree approach has not been considered which is proposed to take up as future study.

Acknowledgement

Authors are grateful to the Director, CSIR-Central Road Research Institute for allowing us to publish this paper. Authors also acknowledge the help rendered by staff of CRRI during the data collection.

References

Aggarwal, A. (2012). Application of fuzzy Logic to Model Trip generation Phase of sequential travel demand analysis. M. tech project thesis submitted to department of civil engineering, NIT Rourkela.

Bataineh, K. M., Naji, M., & Saqer, M. (2011). A comparison study between various fuzzy clustering algorithms. *Jordan Journal of Mechanical and Industrial Engineering*, 5, pp.335-343.

CRRI (2013). Traffic and Transportation studies for Port Blair city, Final Report Vol. I & II, Central Road Research Institute (CRRI), New Delhi, submitted to A&N Administration. Jan 2013.

Joseph, O. & Makinde, O.(2009). Regression model of household trip generation of Ado-Ekiti township in Nigeria, *European Journal of Scientific Research*, Vol. 28, No.1,132:140.

Kalic, Â , M., & Teodorovic, Â . D.(2003). Trip Distribution modeling using fuzzy logic and genetic algorithm, *Transportation Planning and Technology,* June 2003, Vol. 26, No. 3, 213:238.

Papis . C., & Mamdani . E.(1977). A fuzzy controller for a traffic junction IEEE Transaction on systems, *Man and Cybernetics SMC-7,* 707:717.

Wang, L-X., & Mendel, J. (1992a). Generating fuzzy rules by learning from examples, *IEEE Transactions on systems, Man and Cybernetics*, Vol. 22*,* 1414:1427.

Xu,W., & Chan ,Y.(1993). Estimating origin–destination matrix with fuzzy weights, Part1: *Transportation Planning and Technology 17,* 127:144.

Zadeh, L.A. (1965). Fuzzy sets*, Information and Control*, *Vol-8,* 338:353.