ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization, Volume 2, Special Issue 1, December 2013

Proceedings of International Conference on Energy and Environment-2013 (ICEE 2013)

On 12th to 14th December Organized by

Department of Civil Engineering and Mechanical Engineering of Rajiv Gandhi Institute of Technology, Kottayam, Kerala, India

PREDICTION OF TRIP ATTRACTION BASED ON COMMERCIAL LAND USE CHARACTERISTICS

Pretina George, Gymmy Joseph Kattor, Abdul Malik K.V.
M Tech scholar, Department of Civil Engineering, Kottayam, Kerala India
Professor, Department of Civil Engineering, Kottayam, Kerala India
District Town Planner, Malappuram, Kerala India

ABSTRACT

This paper presents an analysis of trip attraction characteristics of the commercial nodes in medium sized towns of Kerala. The study aims to provide a trip attraction model using multiple regression, that is able to predict the trip attracted to any commercial nodes in the medium sized towns in Kerala. Using questionnaire survey, characteristics of the eight selected commercial nodes from the three medium sized towns Tirur, Perinthalmanna, and Ponnani in Kerala are found out. Socioeconomic surveys are conducted for the selected towns for getting the origin-destination data. Based on these surveyed data, correlation and regression analysis are performed. The study showed that the multiple regression model with the independent variables, the number of employees and percentage of office in the commercial node (Adjusted R^2 = 0.9997) gives the better estimate of trip attraction. This model should be useful for estimating the trips attracted to a new or existing commercial center in any medium sized towns in Kerala, to assess the traffic impact of the commercial center on the geometric design of roadways in the surrounding area.

1. INTRODUCTION

Transportation planning process relies on travel demand forecasting which is essential for the design of transportation facilities and services, and also for planning, investment, and policy development. In this process, planners develop information about the impacts of implementing alternative courses of action involving transportation services, such as new highways, bus route changes, or parking restrictions. Detailed data on current travel pattern and traffic volumes are needed for developing travel forecasting/prediction models. One of the most common approaches used by Transportation Engineers to forecast transportation demand is the four-step process. The four-step process is divided into four phases: trip generation, trip distribution, modal split, and traffic assignment. The first step of travel demand forecasting is the trip generations which include trip production and trip attraction. Trip generation serves as an input to modeling transportation requirements and traffic flows influenced by the development. Trip generation aims at predicting the total number of trips generated and attracted to each zone of the study area.

Several studies in western countries have already been done that can be related to analysis of trip attraction characteristics of commercial land use. Most of the studies were supporting the ability of regression analysis to describe the future travel patterns that will result from forecast urban growth. The data collection of the independent variables was mainly done using the questionnaire survey. The Floor area is the most promising factor which affects the trip attraction [4]-[5]. Numbers of employees, parking lots, number of stores are some other factors which were found to affect the trip attraction [6]-[7]. Trips from the household travel survey were taken as the dependent variables which were collected by conducting household survey or from previously implemented origin destination surveys [8].

Commercial areas in a city can take up about 5% of a city‟s land. Majority of the trips are attracted to this commercial area. Trip attraction is important to the traffic engineer and urban planner in considering the impact of new development such as office complex, shopping center and residential development.

Even though a number of trip generation models are developed using multiple regression analysis to develop the prediction equations for the trip generated by various types of land use in different countries, it is not possible to implement these models in all places as the models are developed based on the local conditions of the study area. This study develop a suitable trip attraction model using multiple regression analysis, to forecast the future trips that will be attracted to a commercial node of certain characteristics for a medium sized town in Kerala. So, this model will be beneficial in finding the impact of any development in the commercial nodes on the traffic and thereby, proper planning of transportation facilities and services can be effectively done by the urban planners and traffic engineers.

2. METHODOLOGY

This study uses a combination of analytical and descriptive methods. As a first step in conducting a trip generation model analysis, commercial nodes are selected from the medium sized towns in Kerala using the land use data of the corresponding selected medium sized towns. Then the selected towns are divided into appropriate traffic analysis zones in order to simplify the travel pattern. Origin–Destination (O-D) travel data is collected from the socio economic surveys that are conducted for the selected towns so that the O-D matrix can be prepared which shows the trips that are attracted towards each commercial zone from different traffic zones.

A number of relevant independent variables that influence the trip attraction towards the commercial nodes are identified. And the data related to the identified independent variables from each commercial node are collected by conducting questionnaire survey.

After data collection, the characteristics of each of the selected commercial nodes are analyzed. The correlation and regression analysis are done using the identified variables and the best regression model is formulated for predicting the trips attracted towards the commercial nodes in medium sized towns in Kerala. Then a check for the validation of the formulated multiple regression model is done to verify whether the formulated model is valid or not.

3. STUDY AREA

The study area is selected in such a way as to satisfy the criteria that the commercial nodes should be taken from a minimum of three medium sized towns in Kerala having the latest origin destination data.

The selected medium sized towns for this study include Tirur, Perinthalmanna, and Ponnani from Malappuram district in Kerala. The map showing the position of towns in Malappuram district is shown in fig 1.

The commercial nodes are selected from the towns by analyzing the concentration of commercial land use and the category of commercial nodes in the town. A total of eight commercial nodes are selected from these towns. The details of the selected commercial nodes are given in the table I.

Figure 1. Location Of Selected Medium Sized Towns

Table 1. Summary of Selected Commercial Nodes

Name of commercial node	Town
City Junction	Tirur
Central Junction	Tirur
Thazhepalam Junction	Tirur
Main Junction	Perinthalmanna
Bypass Junction	Perinthalmanna
Hospital Junction	Perinthalmanna
Chamravattam Junction	Ponnani
Chandapadi	Ponnani

4. DATA COLLECTION

4.1 Dependent Variables

The selected towns are divided into zones such that the trips attracted to the selected commercial nodes from these zones can be taken into account. The dependent variables selected for the multiple regression modeling is the number of trips attracted towards the commercial nodes from the divided zones of the selected towns, 'y'. The trips attracted to the corresponding commercial node is obtained from the socio- economic survey which includes the origin-destination (O-D) travel data that was conducted for the selected towns by the Town and Country planning office, Malappuram in the year 2013. A 10% sample of the trips is taken as the dependent variable in this study.

In the case of socio economic survey, 10% sample households are randomly selected from each town and survey is conducted to get the household and travel data. From the socio-economic survey data, the origin destination details of the members of the household on a working day are collected. The details of the total trips attracted towards the selected commercial node from the divided zones of the selected towns are given in the figure2.

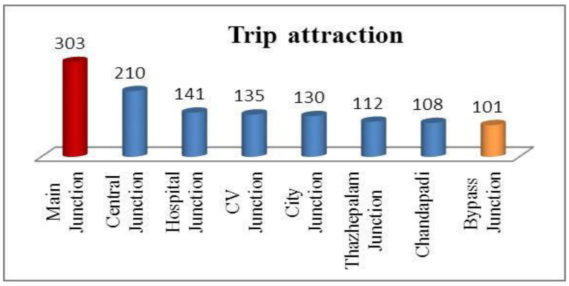

Figure 2. Trips Attracted to the Selected Commercial Nodes

4.2 Independent Variables

Independent variables are identified by studying the characteristics of commercial nodes and also by analyzing the past studies which showed that a number of independent variable influence the trip attraction.

The selected independent variables are listed below:
1.Total number of commercial establishments
2.Type of building.
3.Type of commodity sold.
4.Commercial floor area in the commercial node, m2.
6.Percentage of commercial establishments having one, two or more than two number of floors.
7.Total number of employees in the commercial node.
8.Total number of parking lots in the commercial node.
9.Average width of the major corridor, m
10.Year of operation.

The selected independent variables are collected from the eight commercial nodes by conducting questionnaire survey using a properly designed questionnaire. The entire commercial node is taken into consideration for the data collection of the independent variable.

5. STATISTICAL ANALYSIS OF DATA

Multiple linear regression and correlation techniques involve the seeking of a functional relationship between two or more related variables. Multiple linear regression attempts to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data.

5.1 Model formulation

For model formulation, four commercial node data are used. These include, two nodes from Perinthalmanna town (Main junction, Bypass junction), one node from Tirur town (Central junction) and one node from Ponnani town (Chamravattam junction) is taken. Correlation and regression analysis is performed using Microsoft excel for model formulation.

5.2 Correlation analysis

A "correlation coefficient" is a value that indicates whether there is a linear relationship between two variables. The absolute value of the correlation coefficient will be in the range 0 to 1. A value of 0 indicates that there is no relationship whereas a value of 1 indicates that there is a perfect correlation and the two variables vary together. The sign of the correlation coefficient will be negative if there is an inverse relationship between the variables.

As a first step of model formulation, for obtaining the correlation among the selected variables, correlation analysis is performed using Microsoft excel. The Table III shows the correlation of each independent variable on the dependent variable, along with their notations. Sign (**) indicates a strong correlation ($r < 0.9$), whereas the sign (*) indicates a moderate correlation ($0.5< r <0.9$), where r is the correlation coefficient.

The highest correlated independent variable with dependent variable that is the number of trips attracted to commercial node is the number of employees, x1 (0.935), the number of commercial establishments, x2 (0.931) and percentage of shops where food items are sold, x7 (0.917). Also, the independent variables such as the percentage of shops, x6 (0.742) and total commercial area, x3 (0.645) are moderately correlated with the dependent variable. All the above independent variables are positively correlated with the dependent variable, y. This means that as these independent variables increases, the trip attracted to the commercial node increases.

The independent variables such as the number of parking

TABLE III
CORRELATION RESULT

SL No.	Independent variables	Correlation Test values
1	Total number of employees (x_1)	0.935 **
2	Total number of commercial establishments in the commercial node (x_2)	0.931**
3	Commercial area, m^2 (x_3)	0.645*
4	Average width of the major corridor in meter (x_4)	0.34
5	Number of parking lots (x_5)	-0.697*
6	Percentage of number of shops in the commercial node (x_6)	0.742*
7	Percentage of number of offices in the commercial node(x_7)	-0.645*
8	Percentage of commercial establishments with only one floor (x_8)	0.378
9	Percentage of commercial establishments with two floors. (x_9)	-0.158
10	Percentage of commercial establishments with more than two floors (x_{10})	-0.642*
11	Percentage of commercial establishments with year of operation between 0 to 5 years (x_{11})	-0.133
12	Percentage of commercial establishments with year of operation between 5 to 10 years (x_{12})	-0.651*
13	Percentage of commercial establishments with year of operation between 10 to 20 years (x_{13})	0.139
14	Percentage of commercial establishments with year of operation above 20 years (x_{14})	0.368
15	Percentage of shops where stationary items are sold (x_{15})	-0.02
16	Percentage of shops where food items are sold (x_{16})	0.917**
17	Percentage of shops where textile items are sold (x_{17})	-0.617*

lots, x_5 (- 0.7), percentage of office, x_7 (- 0.65) and the percentage of commercial establishments with year of operation between 5 to 10 years x_{12}, (- 0.65) are negatively and moderately correlated with the trip attraction. Trip attraction decreases with the increase in percentage of the above factors which are negatively correlated with trip.

Only the above mentioned ten independent variables which are moderately and strongly correlated with the trip attraction are considered for the further analysis.

5.3 Multicollinearity

Multicollinearity which is a statistical phenomenon is a problem in multiple regression that develops when one or more of the independent variables is highly correlated with one or more of the other independent variables. If one independent variable is a perfect linear combination of the other independent variables; that is, if it is regressed on the other independent variables and the resulting R^2 = 1.0, then the matrix of inter correlations among the independent variables is singular and there exists no unique solution for the regression coefficients. Correlation matrix is used to detect the possible multicollinearity. It is assumed that, multicollinearity will be diagnosed to be present if the absolute value of the correlation between the independent variables (off-diagonal elements) are larger than 0.4, otherwise there is no multicollinearity or collinearity. Most of the independent variables are correlated in this case and the combinations of such independent variables are avoided for the regression analysis.

5.4 Regression Analysis

After correlation analysis and considering the multicollinearity, regression analysis is performed with the suitable subsets. The model with high R, R^2 and Adjusted R^2 value, minimum standard error of estimate, low significance F value and low p value for the coefficients of independent variables and y intercept is selected as the best model using the regression analysis. The best model obtained is the multiple linear regression model with the independent variables, number of employees (x_1) and percentage of number of office in the commercial node (x_7). The predictive equation from the multiple regression model is

$y = 93.579 + (0.169 \times x_1) - (5.653 \times x_7)$
where,

y = 10% trips attracted to the commercial node

x1 = Total number of employees in the commercial node

x7 = Percentage of number of office in the commercial node.

The first term in the prediction equation (93.579) is a constant that represents the predicted criterion value when both predictors equal zero. The values of 0.169 and -5.653 represent regression weights or regression coefficients of the selected independent variables, number of employees and percentage of office respectively.

Trip will increase on average by, 0.169 trips per day for each one additional number of employees and trips will decrease on average by, 5.653 trips for each 1 % increase in the percentage of office in the commercial node. Regression model details are given in the table IV. R is a measure of the correlation between the observed value and the predicted value of the criterion variable. For this model it has a correlation of 0.99997. This shows that there is a high correlation between the observed and predicted value.

Coefficient of determination or R-square obtained for this model is 0.99993, this shows that, 99.993% of the variation in the trips is explained by the variation in number of employees in the commercial node and percentage of offices in that node, while 0.007% is influenced by other factors.

Adjusted R Square value takes into account the number of variables in the model and the number of observations (participants) our model is based on. This Adjusted R Square value gives the most useful measure of the success of the model. Adjusted R^2 value obtained is 0.99979. This means that 99.979% of the variation in trip is explained by the variation in number of employees in the commercial node and percentage of offices in that node, taking into account the sample size and number of independent variables.

The standard error of the model, which is the standard deviation of residuals, indicates the degree of variation on the data about the regression line established that is, the error we would expect between the predicted and actual dependent variable. The standard error of the regression model is 1.28824. This means that, the expected error for trip attraction predicted is off by 1.28824 trips. The error is comparatively a smaller one.

Significance F value indicates the probability that the regression output could have been obtained by chance. A small Significance of F confirms the validity of the regression output. For this model, significance of F is 0.0083, so there is only a 0.83% chance that the regression output was merely a chance occurrence. This shows there is a linear relationship between all of the x variables considered together and y.

Table IV REGRESSION MODEL DETAILS

Multiple R	R^2	Adjusted R^2	Standard Error	Significance F
0.99997	0.99993	0.99979	1.28824	0.0083

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	93.579	2.725	34.345	0.019	58.958	128.2
Number of employees, x_1	0.169	0.002	92.051	0.007	0.146	0.192
Percentage of office,x_7	-5.653	0.132	-42.828	0.015	-7.33	-3.976

B. Model validation

Model validation involves testing the model's predictive capabilities. Travel models need to be able to replicate observed conditions within reason before being used to produce future-year forecasts. For model validation, the data from the four commercial nodes are used. These data are different from that which we used for model formation. These include Chandapadi from Ponnani town, Hospital junction from Perinthalmanna town, City junction and Thazhepalam junction from Tirur town.

To test the validity of the formulated model, the following three approaches are used.

1) Comparison of observed and predicted trip attraction values:

The Figure 3 shows the observed and predicted trip values that are obtained for the four commercial nodes that are used for model validation. Here we can see that both lines for the observed and predicted trips are coinciding without much difference.

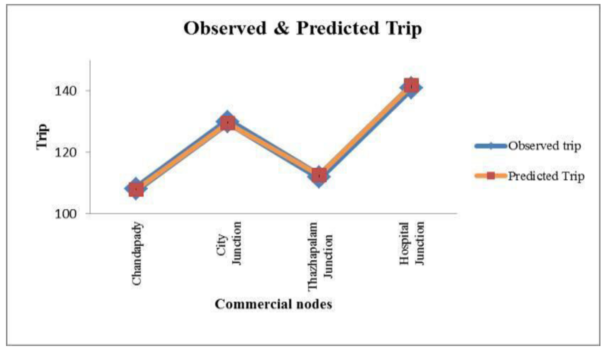

Figure 2. Observed and predicted trip

During the model formulation, the absolute difference obtained is near to 1. Absolute difference between the observed and predicted trip value obtained during the validation process is also less than 1. This shows the validity of the formulated model.

2) Correlation

During the model validation process, to determine the degree to which observed and predicted values are related, determination of correlation is very important. The most commonly used measure of correlation is the coefficient of determination R^2, which describes the amount of variation in the dependent variable which is explained by the regression equation. R^2 predicted obtained in the model validation process is 0.99845. This high R2 value shows the high correlation between the observed and predicted trips hence the model is valid.

3) Variance

Statistical measures can be calculated which measure the variance between observed and estimated values. The most common measure for validation purposes is the standard error of estimates and the standard error of the estimate obtained during the model validation is 1.0555. During model formulation the standard error of estimate is 1.288. As the standard error of estimate obtained for model validation is almost same as that during model formulation, the formulated regression model is valid.

As the above three test are satisfied, the formulated model is valid and can be effectively used for the prediction trip attraction towards the commercial nodes in medium sized towns in Kerala.

6. CONCLUSIONS AND RECOMMENDATIONS

The study developed a trip attraction model for the commercial nodes in medium sized towns in Kerala by using multiple linear regression analysis. Similar to the related studies, this study also found that the number of employees is highly correlated to the trip attraction. Other factors which found to be correlated to the trip attraction are the number of commercial establishments, total commercial area, percentage of shops where food items are sold, percentage of office and shop in the commercial node and the percentage of commercial establishments with year of operation between 5 to 10 years.

This paper provides important information related to the trip attraction characteristics of commercial land use in medium sized towns in Kerala, thereby giving a vast idea about the factors affecting trip attraction towards commercial land use.

The main limitations in this study are the restricted number of commercial nodes and the assumption in regression analysis. The accuracy of the regression model can be ensured by including more representative commercial nodes from different medium sized towns in Kerala. Since, there are no comprehensive studies regarding trip attraction of commercial centers in Kerala, this study will help the future researcher. Also more factors that affect trip attraction can be studied for further research purpose.

REFERENCES

[1] R.J. Paquette, N.J. Ashford, and P.H. Wright, Transportation Engineering Planning and Design, John Wiley and Sons, Inc., New York, 1981.

[2] AASHTO and FHWA, "Quick-Response Urban Travel Estimation Techniques and Transferable Parameters," Users, Guide, National Research Council, Washington, D.C., 1978.

[3] Institute of Transportation Engineers, "Transportation impact analyses for site development: an ITE proposed recommended practice," Washington, DC, Institute of Transportation Engineers, 2005.

[4] A.M. Fillone and M.R. Tecson, "Trip Attraction of Mixed-Use Development in Metropolitan Manila,"Proc., Eastern Asia Society for Transp. Studies, Vol.4, October, 860-868, 2003.

[5] B.S. Waloejo, Surjono, and H. Sulistio, (2012), "The Influence of Trip Attraction on the Road"s Level of Service (LOS) at Traditional Market Land Use," J. Applied Envir. and Biological Sci., February,pp. 92-96.

[6] J.D. Innes, M.C. Ircha, and D.A. Badoe, "Factors affecting automobile shopping trip destinations," J. Urb. Plan. and Develop., Dec., Vol. 116, No. 3,1990.

[7] SCAG Weekend Travel demand Model, Technical Memorandum no.8 - Weekend Trip Attraction Model, Southern California association of Government, June 30, 2008.

[8] A.K. Al-Taei and A.M. Taher, "Trip Attraction Development Statistical Model in Dohuk City Residential Area," Al-Rafidain Engineering, vol.14, no.2, 2006.