SSRG International Journal of Civil Engineering
ISSN: 2348-8352/ https://doi.org/10.14445/23488352/IJCE-V13I7P118

Volume 13 Issue 7, 276-296, July 2026

Original Article

Household-Based Trip Generation Modelling Using Multi-Model Statistical Approaches: Evidence from a Tier-2 Indian City

Adnya S. Manjarekar[1], Anand V. Shivapur[2], Vilas V. Karjinni[3]

[1]*Department of Civil Engineering, Visvesvaraya Technological University, Belagavi, Karnataka, & Department of Civil Engineering, Sanjay Ghodawat University, Maharashtra, India.*
[2]*Department of Civil Engineering, Visvesvaraya Technological University, Belagavi, Karnataka, India.*
[3]*Shree Warana Vibhag Shikshan Mandal & Professor, TKIET, Warananagar, Maharashtra, India.*

[1]*Corresponding Author : adnya8111@gmail.com*

Received: 01 May 2026 Revised: 25 June 2026 Accepted: 03 July 2026 Published: 29 July 2026

Abstract - *Trip generation modelling is a fundamental part of transportation planning, especially in fast-growing Tier 2 cities where there is a limited amount of data available. This study introduces and compares household trip generation models for Sangli–Miraj–Kupwad Municipal Corporation (SMKMC), India, based on the household survey data collected from 1,294 households. A comparative modelling framework: Multiple Linear Regression (MLR), Generalized Linear Model (GLM), Generalized Additive Model (GAM), and Quantile Regression was done in R to test linear, nonlinear, and distributional behavior of travel. The most important factors contributing to household trip generation were found to be household size and household vehicle ownership. Based on the results of the performance evaluation, the Generalized Additive Model (GAM) showed the best performance (AIC = 4329.25; RMSE = 1.26), which is highly capable of capturing the nonlinear relationship while keeping the model interpretation. Quantile regression additionally showed behavioural differences between the trip generation levels. The results from the sensitivity analysis indicated that household size is more influential on the generation of trips by households than vehicle ownership, and the results of the scenario analyses showed that the proposed model is applicable to forecast future travel demand under different scenarios of urban growth. The results suggest that the combination of interpretable statistical models with nonlinear modelling offers a solid and practical way to model trip making on the household level in data-limited urban settings. Though the study has been limited to considering a single year data set and selected socio-economic data, the proposed framework gives valuable support to the forecasting of travel demand and transportation planning in Tier-2 cities of the Indian scenario.*

Keywords - *Household size, Multiple Linear Regression, Trip generation, Urban transportation planning, MLR.*

1. Introduction

To planners and engineers, note "pulse is not a mere academic study, but the pillar of transportation planning in the present day," and the science, which aims at quantifying it, is called trip generation modelling - the initial and most conclusive step in predicting travel demand [1]. Since the inception of the four-step travel forecasting framework in the mid-20th century, trip generation has been recognized as the initial and essential phase, influencing subsequent stages such as "trip distribution, mode choice, and traffic assignment." [2].

Older methods of trip generation, developed decades ago based on a small number of surveys and fixed assumptions, do not work well in the modern, dynamic, and data-driven world [3-6]. Urban mobility has become representative of complexity layers - of income distribution and work patterns, land use transformation, and digital life, which the older models cannot handle [7, 8].

Across time, researchers have tried to close this gap. In Jericho City, [9] in Palestine was developed based on multiple linear Regression. The analysis showed that the family size, employment, and education were strong predictors of trip-making- almost 70 percent of the household travel behavior was explained. A decade later, took this tradition to Hyderabad, India, where modern computation with the use of R software was used to analyse survey data in an efficient manner and visualize the correlations between socio-economic and spatial variables [10]. The similarity between the two studies, in spite of the geographical and time gap, was that they all aimed at shedding light on the patterns of daily movement.

But with the growth of cities, the growth of data, there is a question that steams silently in the air: can linear Regression, with its simplicity, nevertheless, tell the whole story of urban travel? Some argue it is not the most precise due to newer techniques such as artificial neural networks and spatial regressions [7, 11]. Some argue that the value of Regression is that it is transparent - it can explain why mobility occurs, not just how [12]. This debate is more than statistical; it reflects the larger struggle to balance innovation with interpretability in urban analytics [13-18].

The use of regression and machine learning methods for trip generation modelling has been successfully tested in recent studies. But the empirical studies conducted so far have involved almost one model, certain notional outputs, or big cities; very little attention has been devoted to comparing a number of models based on the same data about households. Furthermore, studies addressing household trip generation in Indian bi-nuclei urban systems remain limited, leaving uncertainty regarding the most appropriate modelling approach for such urban contexts. This presents a challenge for the fast-growing Tier-2 cities in India, where data is scarce, but there is a pressing need for transparent and robust planning tools. Recognizing this, the proposed study takes a multi-model approach, using a dataset of 1,294 household samples from Sangli-Miraj-Kupwad Municipal Corporation (SMKMC). Unlike previous studies, which have used only one of the regression or machine learning models, the novelty of the present study is that four regression models were systematically compared, using a common household survey dataset. Furthermore, a review of recent trip generation studies is provided, and a comparative study is conducted to place the model proposed in this study into the perspective of the recent literature and to assess the best model available for the household trip generation modelling in a bi-nucleus urban system.

2. Literature Review

2.1. Chronological Review of Literature

The history of literature on trip generation modelling has passed through several phases of development, beginning with the pioneering ideas on the statistical and Regression approaches, passing through improving econometric and disaggregated models, and finally leading to the incorporation of computational and machine learning algorithms. The preliminary studies focused on the learning of the household and zonal features, the permanence of models, and model transferability. Gradually, more detailed discrete choice models, count data models, and combined land and socio-eco factors were included in the studies, enhancing the predictive power and relevance in different urban settings. Recent articles show an influx of the exploitation of big data, proxy data, and state-of-the-art AI, including neural networks, gradient boosting, and graph neural networks, which enable improved spatial-temporal modelling and interpretability in transportation planning. The detailed overviews are explained in Table 1.

Table 1. Chronological findings

Year Range	Research Direction	Description
1968–1980	Foundational Statistical and Regression Models	Early research focused on establishing the basis of trip generation modelling using linear Regression and category analysis. Research explored aggregation biases, structural relationships in trip generation, and the use of disaggregate Poisson models to overcome issues with linear models. These studies brought attention to household trip-making and non-negativity and discreteness in trip counts.
1989–1997	Advancements in Econometric and Disaggregate Models	Progress was made with the use of generalized linear models, ordered probit, and discrete choice models for trip generation. Focus on research to overcome overdispersion, temporal stability, and inclusion of behaviour theory. Models based on tours and activities were created, taking into account the chaining and scheduling of trips and research on the relationship between Accessibility and induced travel.
2004–2010	Temporal Stability and Model Performance Enhancement	The possibility of the use of multiple cross sections was investigated to enhance the temporal stability and prediction. Also, four-step model weaknesses, such as doing work on logit trip generation models and adding hierarchical and piecewise Regression, were emphasized. Time and place validation and transferability, as well as improved data aggregation, were important themes.
2011–2016	Integration of Socioeconomic and Land Use Variables	During this time, efforts were directed at improving models through the inclusion of socio-economic, household, and land-use attributes. Comparisons of alternative Regression and count data models were made to assess their performance. There was also a shift towards rule-based decision tree models and an improved categorization framework for databases to improve predictive accuracy and practicality in developed and developing environments.

2017–2019	Emerging Data Sources and Machine Learning Applications	The literature began to explore using big data, proxy data, and new machine learning neural networks and boosted trees for trip generation modelling. Attention was paid to integrating multiple data sources for hierarchical and computational graph-based models. There were also efforts to improve spatial transferability problems and benchmarking models in different cities.
2020–2025	Advanced Computational Techniques and Deep Learning	The recent studies center on using state-of-the-art machine learning techniques like CatBoost, gradient boosting decision trees, graph neural networks, and neuro-fuzzy systems. These are superior in trip generation predicting power and interpretation, especially at the local level. Theory-driven and data-driven models in urban and emerging areas, temporal stability, and policy-relevant applications are also discussed.

2.2. Descriptive Analysis

Literature on Trip generation modelling in terms of mathematical models involved, their application in transportation planning, and the historical development of the models throughout history. The studies reviewed use a wide range of approaches, including classical Regression and category analysis, and some of the most sophisticated machine learning and neural network techniques, and were implemented in different geographical settings, such as developed and developing nations. The comparative analysis brings out the merging of socio-economic and land use variables, time stability issues, and practical viability of models in practical transportation planning, which gives insight into the development and present state of trip generation modelling. Table 2 shows details about all studies.

Table 2. Summary of literature review

Study	Model Accuracy	Variable Integration	Methodological Evolution	Practical Applicability
[2]	Models with imputed car ownership outperform simpler models; transferability is limited.	Socio-demographic variables with car ownership modeled	Ordered response probit models for joint car ownership and trip generation	Applied in Nairobi and Dar-es-Salaam with a transferability focus
[3]	CatBoost outperforms ANN with near-perfect accuracy in Ilorin city	Household and trip characteristics integrated	Comparison of CatBoost and ANN models	Application in urban trip production and attraction modelling
[4]	Regression, ANN, and FES models show similar R2; marginal gains from ML	Socio-economic variables included	Linear Regression compared with ANN and fuzzy expert systems	Application in Lagos metropolis trip generation estimation
[5]	Linear Regression is competitive with ML models; spatial heterogeneity is important.	Socio-economic variables with spatial heterogeneity are considered	ML models benchmarked against linear Regression	State-specific models for context-sensitive planning
[7]	ANN highest accuracy for household data; Regression is best for dormitory data	Socio-economic and trip characteristics included	ANN, Poisson, negative binomial, and linear Regression compared	Integrated with the VISUM simulation for urban planning
[8]	Review highlights challenges in multimodal trip generation modelling	Land use and socio-economic variables emphasized	Discusses classical and emerging ML approaches	Proposes a framework for multimodal trip generation modelling
[11]	ANN outperforms regression models in diverse provincial contexts	Household characteristics and development status included	Compared linear, Poisson, negative binomial, and ANN models	Proposed ANN as an alternative for provincial trip generation
[16]	Multivariable regression models with acceptable R2; limited land use impact	Socio-economic and land use variables included	Regression models developed for mixed land use	Developed for a Palestinian city with transferability suggestions
[17]	MLR models with proxy data show strong predictive power	Electricity consumption, land use, and household data integrated	Use of high-frequency proxy data for updating variables	Applied in Sri Lanka for trip generation and attraction
[18]	Category analysis superior in performance for home-based work trips	Socio-economic variables used; traveler behavior is less incorporated	Regression, tobit, Poisson, ordered logit, category, multiple classification compared	Applied to the Seoul metropolitan area transport planning

[19]	Disaggregate analysis preferred over aggregate for trip generation	Socio-economic variables included	Statistical methods for sampling variation and aggregation	Early foundational work on trip generation methodology
[20]	Household-level models show temporal and spatial stability over 7 years	Socio-economic variables included	Disaggregate household models are preferred over zonal models	Applied in longitudinal household travel surveys
[21]	Combined generation-distribution model with microeconomic consistency	Accessibility and behavioral variables integrated	Combined gravity and generation models	Applied at the aggregate level with supply-demand feedbacks
[22]	Disaggregate Poisson models provide a statistical basis and avoid negative forecasts.	Household characteristics included	Poisson regression applied to shopping and personal trips	Applied in the Netherlands transportation study
[23]	GLM framework improves category analysis reliability	Land use and socio-economic variables were modeled with a negative binomial	Generalized linear modelling applied to category analysis	Demonstrated with British rail trip data; applicable to urban trips
[24]	Hierarchical tree-based models address OLS limitations	Socio-economic variables included	Tree-based models compared with OLS regression	Applied to trip generation forecasting
[25]	Joint estimation on multiple cross-sectional surveys improves forecast accuracy.	Socio-economic variables included	Alternative forecasting models with multi-period data	Applied in an urban area with multiple survey points
[26]	Negative binomial and ordered logit models outperform simple linear models	Individual-level data with socio-economic variables	Progression from linear to truncated normal, Poisson, negative binomial, ordered logit	Forecasting in the Toronto region with temporal validation
[27]	Ordered probit models better replicate trip patterns than linear and count models	Socio-economic variables included: trip purpose differentiation	Linear Regression, log-linear, negative binomial, and ordered probit were compared	Applied to the US NHTS
[28]	Panel analysis reveals time-series effects on trip generation by purpose	Socio-economic variables with time effects included	Panel data analysis for time-series trip generation	Applied to national household trip surveys in Korea
[29]	Cumulative logistic regression models show good predictive accuracy	Life cycle, area type, and accessibility variables improve models	Logistic Regression applied to improve temporal stability	Applied to the US survey data with a temporal transferability focus
[30]	The integrated intervening opportunities model outperforms the gravity model for transit trips.	Socio-demographic, socioeconomic, and supply variables included	Nonlinear generation-distribution model for transit trips	Applied in Montreal for morning peak transit trips
[31]	Joint models on multiple cross-sectional datasets yield smaller forecast errors.	Socio-economic variables included	Models estimated on pooled cross-sectional data	Applied in the Greater Toronto Area over multiple years
[32]	Decision tree models sensitive to demographics; validated transferability	Socio-economic and demographic variables at the individual level	Rule-based decision tree approach	Suitable for microsimulation frameworks
[33]	The multimodal trip generation model includes mode choice and local context.	Socio-economic and land use variables integrated	Site-level trip generation with mode choice modelling	Applied in the District of Columbia urban developments
[34]	Hierarchical flow network integrates multiple data sources effectively	Multi-layered socio-economic and behavioral data integrated	Introduced computational graph and back-propagation for trip generation	Demonstrated in the Beijing subnetwork for travel demand estimation
[35]	An adaptive neuro-fuzzy inference system outperforms traditional Regression.	Socio-economic variables included	Neuro-fuzzy system compared with multiple linear Regression	Developed for the Palestinian urban area trip generation

[36]	Random tree regression with LBSN data captures land use impact on trips	Land use and social network data integrated	Data-driven modelling with social network and cellular data	Case study in Nanjing, China, for urban and non-local travelers
[37]	Machine learning models identify GDP and population as key trip drivers	Socio-economic variables integrated	ML methods applied to time-series trip generation data	Forecasting for ferry transport demand between ports
[38]	Gradient boosting decision tree outperforms linear Regression	Socio-economic and urban planning variables included	Introduced GBDT for trip generation forecasting	Applied to the Beijing residential travel census data
[39]	OLS models identify employment and establishment area as key for service trips	Socio-economic and land use variables included	OLS regression applied to service trip generation	Application in the Medellin Metropolitan Area, Colombia
[40]	Multiple piecewise Regression fits data more accurately than previous methods.	Handles categorical, binary, and continuous variables flexibly	Piecewise regression is proposed for nonlinear relationships	Encouraged for trip generation and related modelling
[41]	ANN outperforms MLR with higher R2 and lower error metrics	Socio-economic and household variables included	ANN with logistic sigmoid transfer function compared to MLR	Applied in Baghdad city for trip production forecasting
[42]	Boosting trees best among ML models for population inflow estimation	Socio-economic, demographic, and land use variables analysed	Explainable ML models compared for nonlinear effects	Big data approach for travel demand modelling
[43]	Emerging machine learning models show promise in developing countries	Socio-economic and land use variables are increasingly integrated	Transition from regression and category analysis to ML	Comparative overview of data collection and modelling techniques

In the last 50 years, trip-generation studies have shifted away from simple aggregate OLS modelling to more elaborate count, econometric, and machine-learning models.

The importance of disaggregated household models and socio-economic predictors was put forward in early work. Since the 1980s-2000s, Poisson/Negative-Binomial and ordered models have been suggested by researchers to deal with count distributions and prevent inadmissible forecasts; hierarchical and tree-based models have been suggested as ways of addressing non-linearity and heterogeneity.

Since 2010, an increasing amount of research has been conducted to demonstrate that machine learning (ANN, GBDT, CatBoost, Random Forest) and hybrid methods can be particularly more predictive when large, high-frequency, or multi-source data (LBSN, electricity consumption, panel surveys) are available.

Simultaneously, comparative studies again and again reveal a pragmatic trade-off, where ML methods are more accurate but may lack interpretability as well as transferability across contexts; count/econometric models (Poisson/NegBin) are more appropriate to trip count distributions, but still need to be careful with overdispersion and spatial heterogeneity. Recent reviews also emphasize that Tier-2 city settings and data limitations are underrepresented: most of the state-of-the-art ML performance is observed in large cities, and smaller cities are less explored and, in many cases, have interpretable, transparent models to plan decisions.

2.3. Comparative Analysis of Existing Trip Generation Models

Table 3 shows a comparative analysis of recent trip generation studies that have been published in the last five years (2020-2025). Based on the review, it can be concluded that trip generation modelling is widely used, and conventional regression models, nonlinear regression techniques, and machine learning approaches are all used extensively.

The most consistent variables for all study areas were household size, income, vehicle ownership, employment, and land-use characteristics. While a few studies showed increased predictive accuracy for machine learning algorithms, most studies concentrated on a single modelling approach or particular trip types. The present study, however, evaluates four commonly used statistical regression techniques in a systematic manner to assess the suitability of the models for the household trip generation in the bi-nuclei urban system of Sangli–Miraj–Kupwad, using a common household survey data set.

Table 3. Comparative analysis of existing household trip generation models

Citation	Study Area	Model Used	Significant Variables	Key Findings	Comparison with the Present Study
[36]	Nanjing, China	Random Forest + OLS	Land-use characteristics, Points of Interest (POIs), public transport stations, and highway entrances	The best results in terms of prediction accuracy ($R^2 > 0.93$) were obtained from land-use characteristics.	The present study concentrates on socio-economic variables of the household, instead of land-use variables.
[38]	Beijing, China	GBDT + OLS	Population, demographic characteristics, occupation, income, transport ownership, land use	GBDT improved prediction performance by approximately 12% over OLS based on RMSE and MAE.	The present study compares statistical regression models instead of machine learning algorithms.
[40]	Cairo, Egypt	Multiple Piecewise Regression	Household size, household income, car ownership, employees, gross floor area	Multiple Piecewise Regression improved prediction accuracy compared with multiple regression and cross-classification models.	The present study compares several statistical regression models rather than a single nonlinear regression model.
[15]	Kumasi, Ghana	Multiple Linear Regression (MLR) + Ordinal Logistic Regression (OLR)	Age, household size, gender, trip purpose	Household size and age significantly influenced student household trip generation, while OLR showed better predictive capability.	The present study develops household trip generation models for the general population rather than student travel behaviour.
[44]	New York City, USA	Spatial Multi-Graph Attention Network (Spatial-MGAT)	POI density, socio-demographic characteristics, road network, transport facilities	Spatial-MGAT achieved the lowest prediction error among the compared models by effectively capturing spatial interactions.	The present study models household trip generation without incorporating spatial graph-based learning.
[10]	Hyderabad, India	Multiple Linear Regression (MLR)	Travel distance, trip cost, monthly income, journey duration, vehicle ownership	Travel distance, trip cost, income, journey duration, and vehicle ownership significantly influenced trip generation ($R^2 = 0.646$).	The present study extends conventional MLR by comparing multiple statistical regression models.
[16]	Jericho, Palestine (14 TAZs)	Multivariable Linear Regression	Household size, employed persons, students, income, vehicle ownership, age groups, land-use variables	The overall trip model achieved $R^2 = 0.691$, improving to $R^2 = 0.722$ after incorporating land-use variables, while the education trip model achieved $R^2 = 0.969$.	The present study compares multiple regression techniques and evaluates model performance to identify the most suitable trip generation model for an Indian bi-nuclei context.
[45]	Patna Urban Agglomeration, India (100 TAZs)	Linear Regression + Support Vector Machine (SVM)	Population (trip production), employment (trip attraction)	SVM with an RBF kernel achieved the highest prediction accuracy ($R^2 \approx 1.00$ for trip production and $R^2 = 0.95$ for trip attraction).	The present study compares statistical regression models, whereas this study evaluates machine learning algorithms.

[5]	United States (NHTS dataset)	Linear Regression + Random Forest + CatBoost	Household size, young children, household income, vehicle ownership, age, education, public transit use	Linear Regression remained competitive (weighted $R^2 = 0.321$) and outperformed other models in 23 of 51 states.	The present study develops household trip generation models using household survey data from a single urban region.
[46]	Sydney, Melbourne, Brisbane (Australia); Seattle and Chicago (USA)	Negative Binomial Regression + Bayesian Regression + Random Forest	Population, household size, household income, vehicle ownership, station density, land use, POIs, network connectivity	Random Forest achieved the highest prediction accuracy, while Negative Binomial Regression performed well for over-dispersed pedestrian trip generation data.	This study evaluates pedestrian trip generation across multiple cities, whereas the present study focuses on household trip generation in a bi-nuclei urban system.

2.4. Research Gap

The literature on trip generation modelling shows two simultaneously occurring initiatives: the growing use of advanced data-driven techniques, such as machine learning and spatial models, to enhance predictive performance and the need for more interpretable and operational models in data-limited urban settings. Although advanced models can perform better, they typically demand large, diverse datasets and are opaque, making them difficult to apply for planning in smaller and new urban areas. Conversely, legacy methods, based on Regression, while interpretable, suffer from their inability to represent complex relationships and variations in travel behavior. This leaves a key challenge for developing modelling approaches that meet the needs of simplicity, interpretability, and rigorous analysis. Despite these developments, there has been very little research to systematically compare various statistical modelling techniques for household trip generation across multiple cities in Tier-2 India, with a common dataset. The current research mostly uses only one modelling tool, and it is not clear whether linear modelling, nonlinear modelling, or distribution modelling is more appropriate for data-limited urban settings. This paper fills this gap by adopting a multi-modelling approach using Multiple Linear Regression (MLR), Generalized Linear Model (GLM), Generalized Additive Model (GAM), and Quantile Regression to understand the linear, nonlinear, and distributional features of trip generation behavior with local data.

3. Details of Study Area

The two cities of Sangli and Miraj, which are very historic cities of Maharashtra, have transformed into major metropolitan cities since they were established in the Middle Ages and ancient India. The princely state of Sangli (which was earlier under the Maratha Empire) was ruled by the Patwardhan family since 1782 [47]. Sangli was a significant city throughout the Maratha and British periods, famed for its agricultural wealth, notably turmeric, and the sugar industry [48]. The strategic position, large tract of agricultural land, and the prosperous industrial sector of the region have a demand for a thorough approach to the governance of municipalities [49]. Figure 1 shows the Sangli-Miraj-Kupwad Municipal Corporation (SMKMC) area that was established in 1998 to manage the rapidly developing urban agglomeration composed of the twin cities of Sangli and Miraj, as well as the neighboring town of Kupwad. In the 2011 census, Sangli-Miraj-Kupwad region had a population of 502,793, with an area of 118.18 square kilometres.

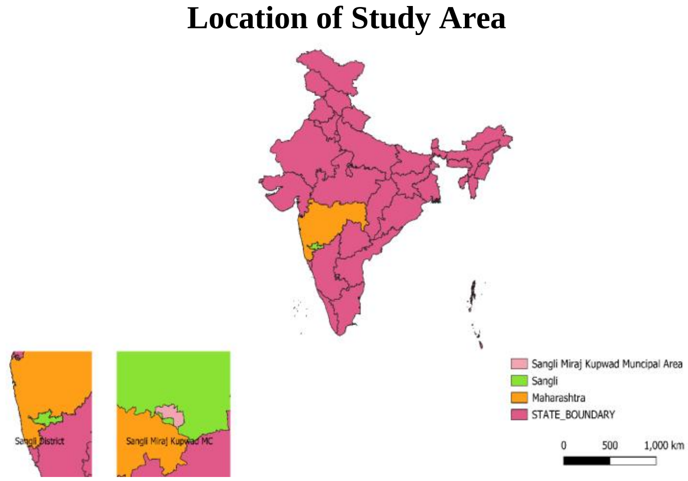

Fig. 1 Sangli miraj kupwad municipal corporation area

4. Methodology

The aim of the project is to create a trip generation model, which is founded on individual trips which start at the residential area of the study area. The required data was obtained through a survey of households by use of a questionnaire. The data obtained was then analysed statistically, that is, correlation and four model analysis, in R-software. The output of the model may be capable of forecasting the trips that will be created by similar residential land use in future. Figure 2 shows the methodology used in this study.

Fig. 2 Methodology

4.1. Data Collection

The dependent variable of the chosen study area is the number of trips that are generated in each residential zone. A detailed household survey was carried out to get the required data by conducting home-based interviews on 1,294 households. The survey was aimed at getting specific travel data of people living in these households to know what factors contributed to the generation of trips. As the source of most trips is usually the household, the home-based trips were singled out and examined. The gathered data was put to the purpose of creating an Origin-Destination (O-D) matrix, which allowed for evaluating the travel behavior of the study area. To achieve representative spatial and socio-economic coverage of the study area, a multistage stratified sampling approach with proportionate allocation was used. The Sangli – Miraj – Kupwad Municipal Corporation was first divided into 12 Traffic Analysis Zones (TAZs), and the sample size was proportionately divided among the zones according to the number of households in each zone. The zonal sample was then proportionately distributed across the 74 municipal wards, giving each ward the same size in terms of the number of households. The medium-sized household (3–4 people) accounted for 66.5 percent of the population sample and therefore was the largest household size in the study. The proportion of all households that owned at least one private vehicle was 84.5% of the total, with a high level of dependency on private transport for commuting to work. In terms of types of trips, the most important trips were also household trips (38.7% work trips; 18.5% shopping trips), indicating that the primary motives for travel were employment, education, and essential trips. Moreover, the private area is the largest occupation, and a two-wheeler is the most popular mode of transportation among the employees, both male & female. The following features indicate that the survey data are reasonably representative of the demographic and travel behavior differences that can be used to formulate, compare, and test valid models of household trip generation.

- The questionnaires were conducted by direct household interviews, hence ensuring that comprehensive and dependable information was collected by residents.
- The questionnaire was made in a structured format to gather information that pertains to both dependent and independent variables that affect the generation of a trip.

- Several independent variables were also contained in the questionnaire and they included: Age and gender of individuals, Occupation and monthly income, Type and nature of dwelling unit, Vehicle ownership, Approximate distance of travel and travel time, Cost of trips per day and frequency of travel, Mode of transport used and purpose of trip and Accessibility, satisfaction, and willingness to shift or pay more to have improved transport facilities.

The questionnaire gathered a large number of socio-economic and travel-related variables, but household size, vehicle ownership, and household income were considered for model development. These variables were chosen because of their theoretical importance in household trip generation studies, exploratory correlation analysis, and their statistical significance in the initial model estimation. The variables chosen gave the best explanatory power in terms of model simplicity, with no over- or under-identification of the model.

5. Data Analysis and Model Development

5.1. Multiple Linear Regression (MLR)

The data was initially tabulated and purged in Microsoft Excel, and then, extensive statistical analysis was performed with the help of the R software environment. R was selected due to its effectiveness in working with massive data sets, statistical calculations, and visualization of the relationship between variables.

The analysis was mainly aimed at analyzing the variables that have an impact on generating trips, where the number of trips (ntrip) was used as a dependent variable, and the independent variables included the household size (hhsize), number of cars owned (numcars), and income. Data analysis was initiated using Exploratory Data Analysis (EDA) to know the structure and distribution of the variables. R was used to create the measures of mean, median, range, and standard deviation of each variable with the help of the summary () and describe () functions. Frequency distributions and graphic representations (histograms and scatter plots) were developed to determine the trends, distribution, and possible outliers in the data. Then a correlation table was created to determine the strength and direction of the relationships between the main variables, in this case, the number of trips, the number of people in a household, and the number of cars. This step assisted in establishing the variables that were closely related to the trip generation behavior. After the exploratory analysis, a multiple linear regression model was constructed in order to measure the relationship between the independent and dependent variables. The R regression model was given as follows:

$$\text{ntrip} = \beta 0 + \beta 1(\text{hhsize}) + \beta 2(\text{numcars}) + \beta 3(\text{income}) + \varepsilon$$

β0 indicates the intercept, β1, β2, and β3 indicate the regression coefficients, and ε is the random error. The model was tested by the coefficient of determination (R2) and adjusted R2, which determines the level at which the explanatory variables can predict the number of trips generated. The given analytical framework allows grasping the effect of socio-economic and household factors on the number of generated trips in a certain region, such as family size, vehicle ownership, and income, etc. A regression model can therefore be used as a basis for trip generation modelling in transport planning and policy analysis.

5.1.1. Descriptive Statistics

Table 4 shows the descriptive statistics of all the variables in the dataset, which were obtained using the summary () command. The number of trips (ntrip) was between 1 and 10.5, the mean number of trips was 5.19 trips per household, and the median was 4.5, which indicates the household trips were moderate. The number of persons living in households (hhsize) ranged between 1 and 7, with a central tendency of 3.39, which implies that the majority of the households were small to medium-sized family households. Numcars, which was vehicle ownership, was between 0 and 5 vehicles, and the mean was 1.55, indicating that most of the households had one vehicle. There was a broad distribution of household income (income), ₹2,000 to ₹750,000 per month, with an average income of ₹37,341, which implies that the socio-economic status of households was very different. Other household attributes like the number of working members (numwork) and number of driving members (numdrive) also had means of 1.56 and 1.55, respectively, which are consistent with the mean household size and number of cars, which supports the correlation between employment, mobility, and frequency of travel.

Table 4. Descriptive statistics and frequencies output

Statistic	ntrip	hhsize	numtot4	num5to21	numadult	numwork	numdrive	income	numcars
Min.	1.000	1.000	0.000	0.000	0.000	0.000	0.000	200	0.000
1st Qu.	4.000	3.000	0.000	0.000	0.000	1.000	1.000	25000	1.000
Median	4.500	3.000	0.000	1.000	0.000	2.000	1.000	33000	1.000
Mean	5.193	3.395	0.299	0.887	0.178	1.564	1.554	37341	1.554
3rd Qu.	6.500	4.000	0.000	2.000	0.000	2.000	2.000	45000	2.000
Max.	10.500	7.000	2.000	4.000	2.000	4.000	5.000	750000	5.000

5.1.2. Graphical Representation

The graphical outputs created with the help of R Software that helpful to understand the features and interdependencies between the most important variables of the study. The scatter plot of Household Size vs. Number of Trips, as illustrated in Figure 3, indicates that a positive relationship exists between the Household Size and the Number of Trips. There is also a tendency to increase the number of trips generated by household size, which means that the larger the household, the higher the travel activity.

The distribution of points in various household sizes also indicates that there is variability in traveling behavior, which is indicative of the disparity in the socio-economic status, trip purpose, and vehicle accessibility among similar groups of households. The histogram of Income Distribution illustrated in Figure 4 indicates a very skewed rightward distribution where a significant percentage of households are between the lower to middle-income range, and only a percentage of households record very high-income levels. This is a skewed distribution common in the socio-economic data and points to a high level of income inequality among the research population. The majority of the concentration at lower income levels can also reflect on travel behavior, especially mode choice and frequency of travel.

These plots, combined with the descriptive and Regression analyses, can be used to indicate that household size and vehicle ownership are key factors in influencing the trip generation, whereas income, albeit not distributed equally, helps explain the differences in the travel behavior of households of various groups.

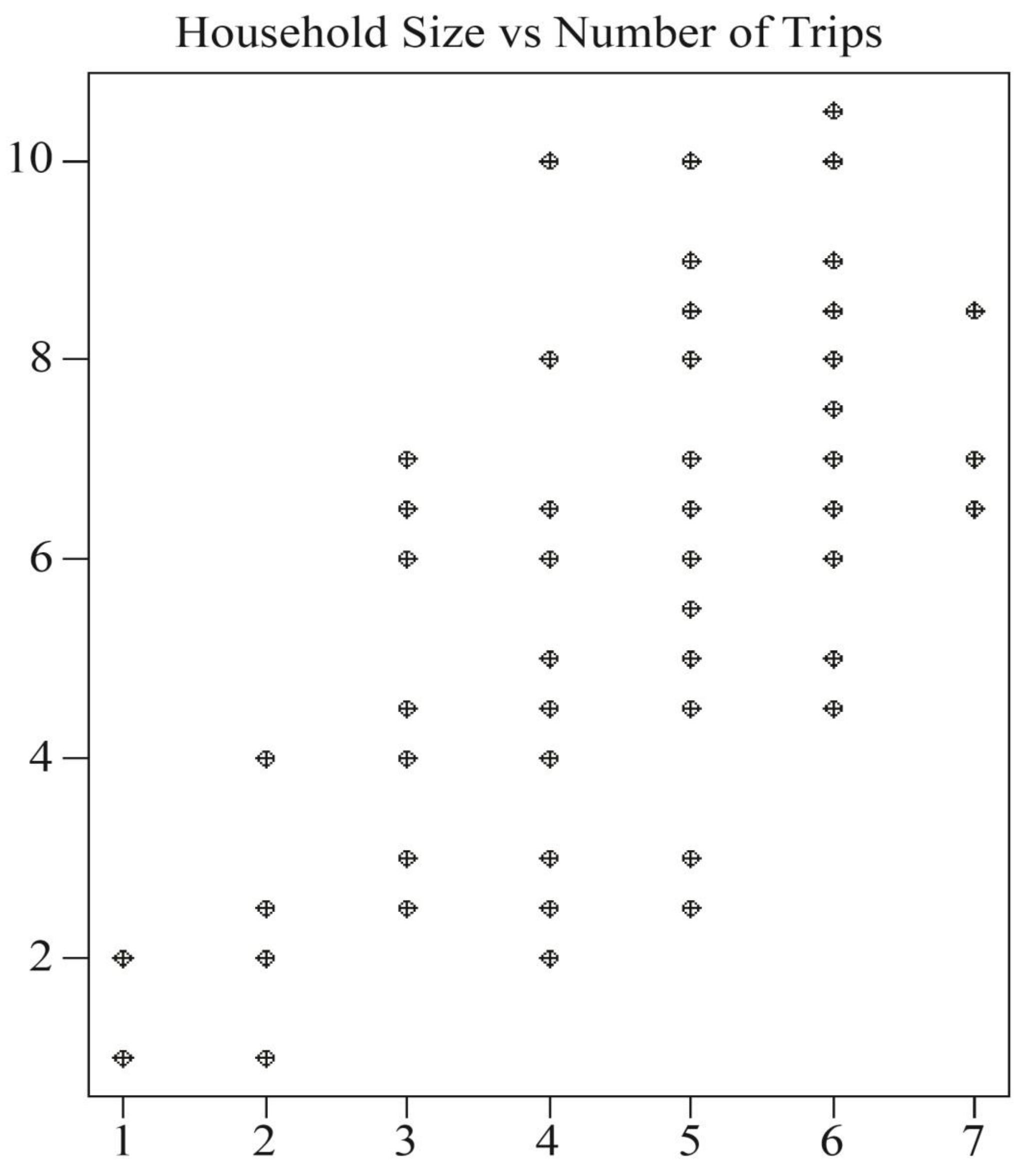

Fig. 3 Household size vs number of trips

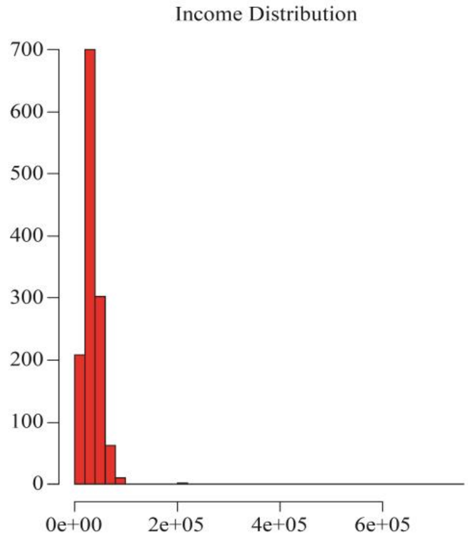

Fig. 4 Income distribution

5.1.3. Regression Model Development

In the current research, the multiple linear regression analysis was used to determine the statistical correlation between the number of trips generated (dependent variable) and the key household variables, including household size, number of cars owned, and household income (independent variables). Regression analysis is a powerful instrument in trip generation research because it enables researchers to measure the degree to which changes in socio-economic and demographic factors will affect travel.

The predictor variables data were collected using detailed household surveys, and the response variable, the total number of trips, was also recorded using the same data. A multiple regression model was created with the help of the R software environment to determine the contribution of each independent variable to trip generation. Table 5 shows the model equation, which is as follows:

Table 5. Regression coefficients

Explanatory Variable	Parameter Estimate	Standard Error (S.E.) of the estimate	T-statistics
Household Size	1.168280	0.038105	30.659
Number of Cars (β_{ncar})	0.491105	0.043999	11.162
Income	0.001657	0.001161	1.427
Constant	0.401905	0.136041	2.954

$$Yi=0.4019+1.1683* X_1 +0.4911* X_2 +0.001657* X_3$$

Where,
Yi = Number of Trips
X_1 = Household Size
X_2 = Number of Cars (motorized vehicles)
X_3 = Income

Interpretation of the coefficient of household size:

- For every 1 person increase in a household, the trips will increase by 1.16 for the household.
- Interpretation of the coefficient of the number of cars:
- For a unit increase in car ownership, the trips will increase by 0.49 for the household.

Interpretation of the coefficient of income:

For a unit increase in the income, the number of trips increases very marginally, which is about 1.64e-03. However, this is an incorrect way to interpret income, because income, being an important socio-demographic characteristic, will affect the number of trips for the household. The income value has to be scaled down to match the other values in the dataset.

$$Yi=0.4019+ (1.1683×3.39) + (0.4911×1.55) + (0.001657×37.341)$$

$$Yi=5.19$$

The model predicts an average of 5.19 trips per household, closely aligning with the observed mean from survey data.

5.1.4. Regression Statistics and Goodness of Fit

In order to test the effectiveness and consistency of the developed trip generation model, it is necessary to consider the regression statistics and goodness-of-fit indicators. Table 6 below gives the calculated regression statistics and explains their relevance in the model developed.

Table 6. Regression statistics

Goodness of fit measures:	
Regression Sums of Squares (SSR)	2406.084
Residual Sum of Squares (SSE)	2231.11
Total Sum of Squares (SST)	4637.2
R 2	0.5188
Adjusted R 2	0.5177
Number of Cases	1294

Evaluations of the developed regression model performance were done based on several goodness-of-fit measures. The Regression Sum of Squares (SSR) was 2406.084, with the Residual Sum of Squares (SSE) of 2231.11, and a Total Sum of Squares (SST) of 4637.20. These values imply that the model is able to explain a significant amount of the total change in the number of trips created. The calculated Coefficient of Determination (R^2) is 0.5188, which is about 51.88 per cent of the variability of trip generation that can be accounted for by the independent variables, which are the number of household size, the number of cars, and income. The adjusted R2 (0.5177), which reflects the number of predictors, shows that the model has a high level of explanatory importance, and it has no risk of over-fitting. The regression results indicate a statistically dependable yet fairly powerful model fit with a sample size of 1, 294 households. This means that the chosen household traits are a significant representation of trip generation behavior in the study region, which justifies the appropriateness of the multiple linear regression method applied in this analysis.

5.2. Generalized Linear Model (GLM)

5.2.1. Graphical Representation

Visual graphical analysis was performed to help investigate the relationship between the explanatory variables and trip generation. The scatter graph in Figure 5 of Household Size on Trip Numbers confirms that the relationship is positive, linear, and there is a positive association between trip generation and household size. However, the data points do have a little 'noise' in them, suggesting that there could be some variation and non-linearity in its relationship. It can be observed in the vehicle ownership histogram that most of the households own between 1 and 2 vehicles, and that the number of households with 2+ vehicles gradually declines. This distribution indicates the impact of motor vehicle ownership on trip generation. To sum up, the visualization validates the importance of household size and vehicle ownership for the trips generated.

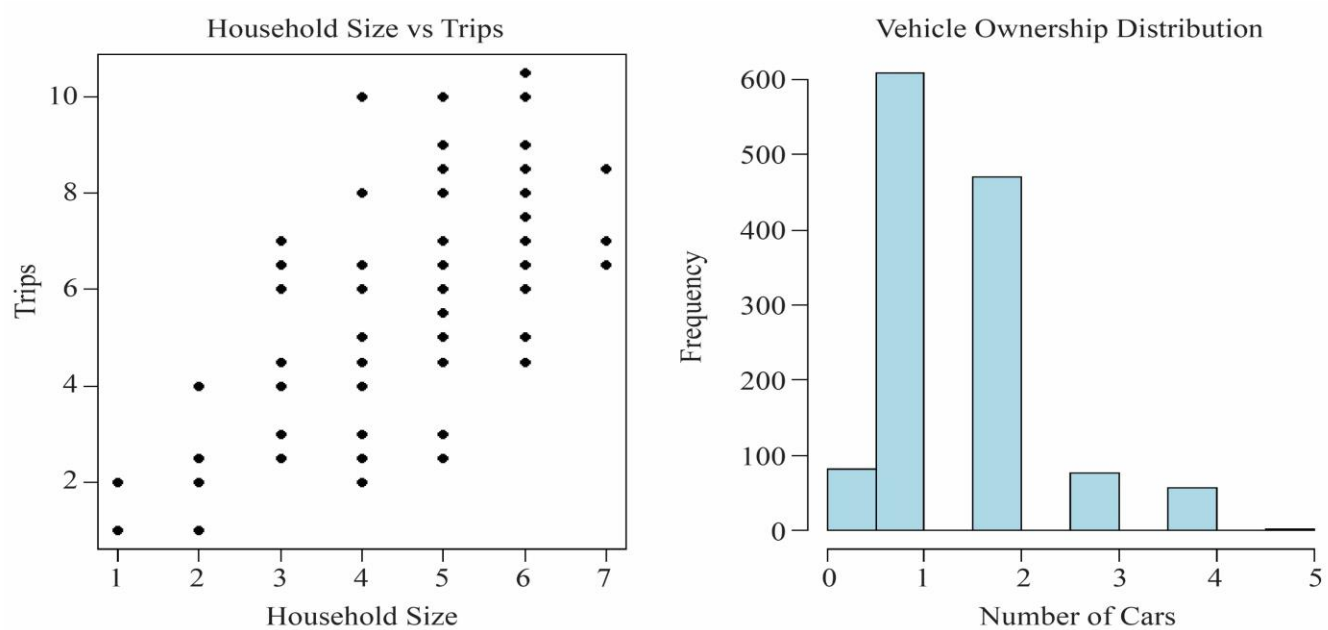

Fig. 5 Household size and number of cars

5.2.2. Regression Model Development

Trip generation was modelled using a Gaussian Generalised Linear Model (GLM) and an identity link function. This model is the same as multiple linear Regression, but can be expanded to more sophisticated ones. A functional form of the model:

$$ntrip=\beta0+\beta1(hhsize)+\beta2(numcars)+\epsilon$$

The model was fitted using the maximum likelihood estimation approach, and diagnostic tests were performed for the statistical significance of the model. To determine multicollinearity of explanatory variables, the Variance Inflation Factor (VIF) was calculated, which is very close to 1, indicating the absence of multicollinearity.

5.2.3. Results and Diagnostics

The results of the estimated GLM show that household size and vehicle ownership are significant factors influencing trip generation at the 1% level. Household size (1.17) implies that an increase in the number of family members is associated with an increase of trips by about 1.17 trips per day.

Similarly, the coefficient on ownership of own cars (0.50) indicates that for every additional car, the number of trips goes up by approximately 0.5 trips per day. The standardized coefficients also show that, compared to vehicle ownership (0.23), household size (0.62) is a better predictor of trip generation, and thus the most important. Model fit is satisfactory, there is a significant reduction in residual deviance, and the Akaike Information Criterion (AIC) is 4387.2. However, the Breusch-Pagan test indicates that there is evidence of heteroscedasticity ($p < 0.001$). To correct for this issue, robust standard errors were used to draw accurate statistical conclusions. The model was validated with a train-test split, resulting in an RMSE of 1.28 and an MAE of 1.05, suggesting reasonable model performance for travel behavior applications.

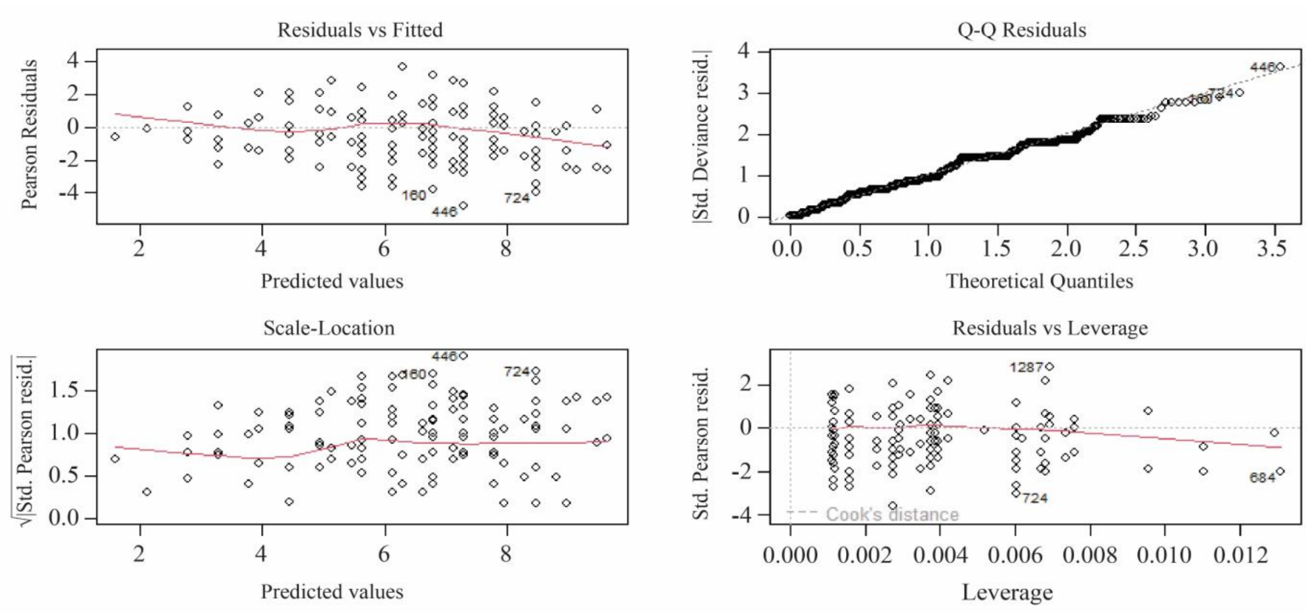

Fig. 6 Diagnostic plots for GLM

Figure 6 below shows diagnostic plots, which show that the residuals are mostly centered on zero, with a slight pattern that suggests a bit of non-linearity. The Q-Q plot is relatively normal with some departures at high values. The scale-location plot shows an increasing spread of residuals, so heteroscedasticity is present. There are no influential observations in the plot of residuals vs leverage. The model is good, with heteroscedasticity corrected for by using robust standard errors.

5.3. Generalized Additive Model (GAM)

5.3.1. Graphical Representation

Figure 7 shows the plot of household size and trip generation, which shows a positive but nonlinear relationship, implying the growth in trip generation is not linear over all household sizes. This finding suggests that it is important to adopt a flexible approach to model this relationship. There is a relatively linear positive effect of vehicle ownership on trip generation.

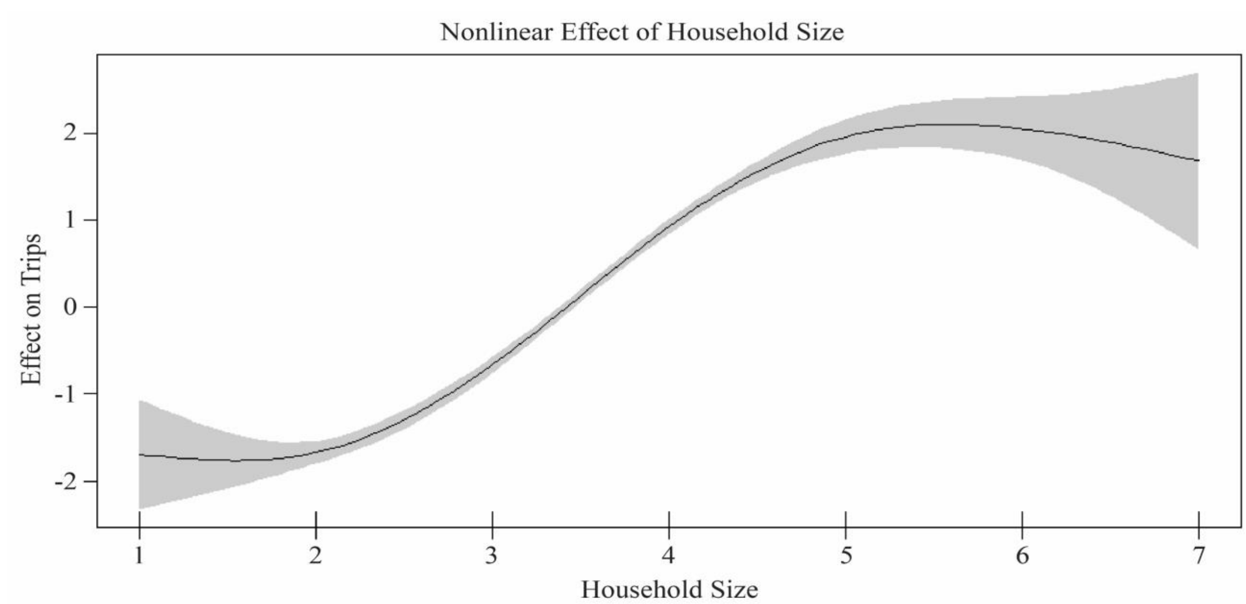

Fig. 7 GAM smooth effect of household size

5.3.2. Model Development

To model the potential nonlinear effects, a Generalized Additive Model (GAM) was used with the family set to Gaussian and link set to identity. This model features a smooth term for household size and a linear term for vehicle ownership:

$$ntrip=\beta0+s(hhsize)+\beta1(numcars)+\epsilon$$

The smooth function allows the model to model the possible nonlinearities between household size and trip generation.

5.3.3. Model Results and Interpretation

The GAM indicates that the smooth function of household size is highly significant ($p < 0.001$), with an effective degree of freedom (edf = 3.756), which suggests the trip generation process is indeed nonlinear. This does establish that the impact of household size varies at the household level to which it applies and is not necessarily linear. The number of vehicles remains a significant linear predictor (coefficient = 0.48), implying that household trips increase by almost 0.5 trips per household for each additional vehicle. The GAM provides a better model fit than the GLM. The drop in the AIC (GLM: AIC = 4387.2, GAM: AIC = 4329.2) indicates that this is an improved fit. Similarly, the prediction performance of the models is improved by decreasing the RMSE from 1.28 (GLM) to 1.26 (GAM) or the MAE from 1.05 to 1.04. The deviance explained by the model is around 54.1%, which further suggests that the GAM model is a better predictor than the linear model.

In summary, our findings show that travel behavior is indeed nonlinear in relation to household size, and the GAM offers a better and more realistic model of household travel.

5.4. Quantile Regression Model

5.4.1. Graphical Representation

Figure 8 quantile regression coefficient plots show the variation in the effects of explanatory variables across quantiles. The coefficient of household size increases (from low to high quantiles), suggesting that it has a greater effect on high-trip households. By contrast, the coefficient of vehicle ownership shows a decreasing trend, which implies that it affects households with low trip generation more. The shaded confidence bands show increasing variability at higher quantiles, suggesting that there is some degree of variability in travel behavior across different levels of travel demand.

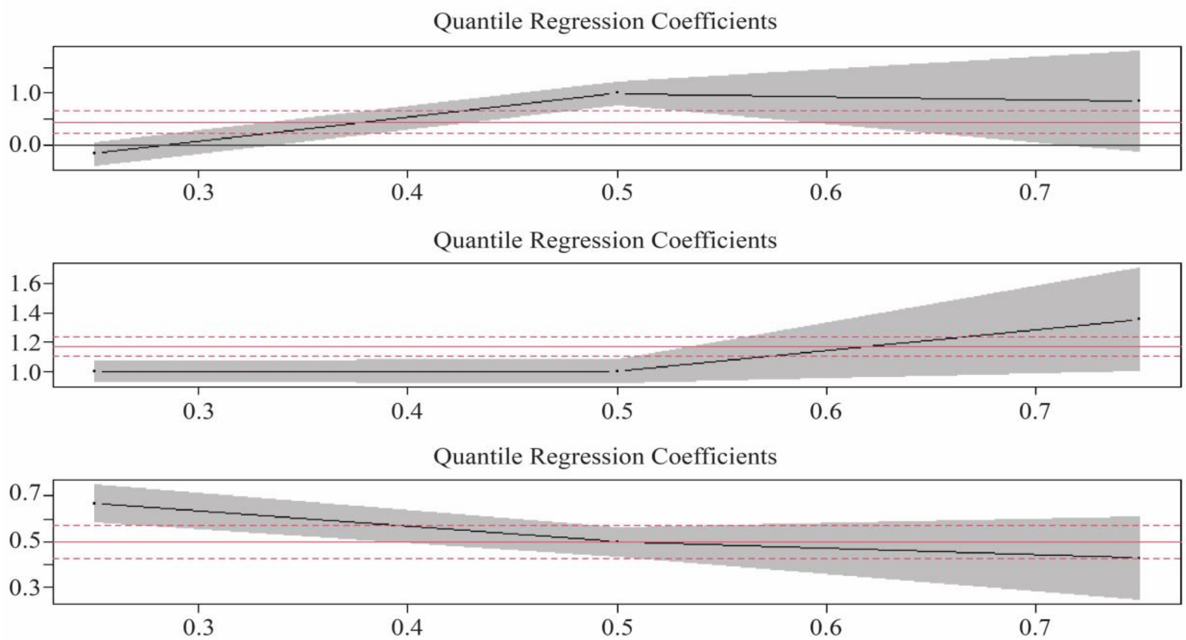

Fig. 8 Quantile regression coefficient variation

5.4.2. Model Development

Three quantile regression models were estimated (25th, 50th, 75th percentiles) to capture varying effects at varying levels of trips generated. The model will be:

$$ntrip=\beta0+\beta1(hhsize)+\beta2(numcars)$$

This is more specific than mean-based models because it allows estimation of the effects of explanatory variables over all percentile levels of the conditional distribution, and thus provides a more detailed understanding of the behaviour.

5.4.3. Model Results and Interpretation

The results reveal that the explanatory variables have quantile-specific effects. Household size is a progressively stronger influence, with the coefficient ranging from 1.00 at the 25th percentile up to 1.36 at the 75th percentile; in other words, the higher the trip-generating household size, the greater the influence. The coefficient for ownership of vehicles is decreasing, on the other hand, from 0.67 at the lower quantile to 0.43 at the higher quantile. This indicates that having a car is more important for households with lower travel demand. The median model's RMSE equals 1.32, and its MAE equals 1.05, indicating that the model has a reasonable fit to the data. However, the utility of this type of model is the finding of a heterogeneity in behaviour and not necessarily its prediction accuracy. In general, the results show that trip generation behavior is heterogeneous with respect to demand and cannot be explained by mean-based models such as GLM and GAM.

6. Comparative Analysis of Developed Models

A comparative evaluation of the four modelling approaches was conducted to assess their effectiveness in explaining household trip generation behavior.

Table 7. Comparative analysis of trip generation models

Model	Type	Key Variables	Nature of Relationship	AIC	RMSE	MAE	Key Findings	Role
MLR	Linear	hhsize, numcars	Linear	–	–	–	Basic relationship established	Baseline
GLM	Linear (Gaussian)	hhsize, numcars	Linear	4387.2	1.28	1.05	Statistically robust, significant variables	Refined baseline
GAM	Nonlinear	s(hhsize), numcars	Nonlinear (hhsize)	4329.2	1.26	1.04	Captures nonlinear behavior, best fit	Final Model
Quantile Regression	Distribution-based	hhsize, numcars	Varies across quantiles	–	1.32	1.05	Reveals behavioral heterogeneity	Supporting Model

Table 7 shows a comparison of the trip generation results of the four models. The results indicate that while MLR and GLM have a good model fit and significance of variables, the linear association fails to explain the complexity of travel behaviour.

The GAM outperforms MLR and GLM with a lower AIC and better accuracy, indicating the presence of nonlinear effects. Quantile regression, while not enhancing prediction performance, adds valuable insights about the impact of variables across different demand levels. In conclusion, the GAM is the most appropriate, with quantile regression offering insights into behaviour.

6.1. Sensitivity Analysis and Scenario Testing

To assess the effect of the significant explanatory variables on household trip generation, a sensitivity analysis was performed with the chosen Generalized Additive Model (GAM). The analysis was done with a representative baseline household, which assumed the median household size (3 persons) and median vehicle ownership (1 vehicle). The one-factor-at-a-time design was used; the one factor being changed as the other factors were kept constant, and the change in each factor was tested on its effect on the predicted household trips.

This household size sensitivity analysis (Table 8 and Figure 9) shows that household size is the major factor that influences household trip generation.

The nonlinear response observed in the GAM suggests that there are varying effects of household size within each of the household categories. The decline in the marginal increases for larger-sized households indicates saturation in trip-making behaviour, while the higher participation and travel demand increases for medium-sized households are reflected in the rapid growth in their predicted trips. This proves that the linear relationship between household size and trip generation could lead to under- or over-estimation of travel demand for some types of households, showing the benefit of the GAM over conventional regression models.

Table 8. Sensitivity analysis of household size on household trip generation

Household Size	Vehicle Ownership	Predicted Household Trips	Change in Predicted Trips	Percentage Change (%)
2	1	3.26	0.00	0.00
3	1	4.26	1.01	30.92
4	1	5.85	1.59	37.28
5	1	6.88	1.03	17.61
6	1	6.98	0.09	1.35
7	1	6.61	-0.37	-5.31

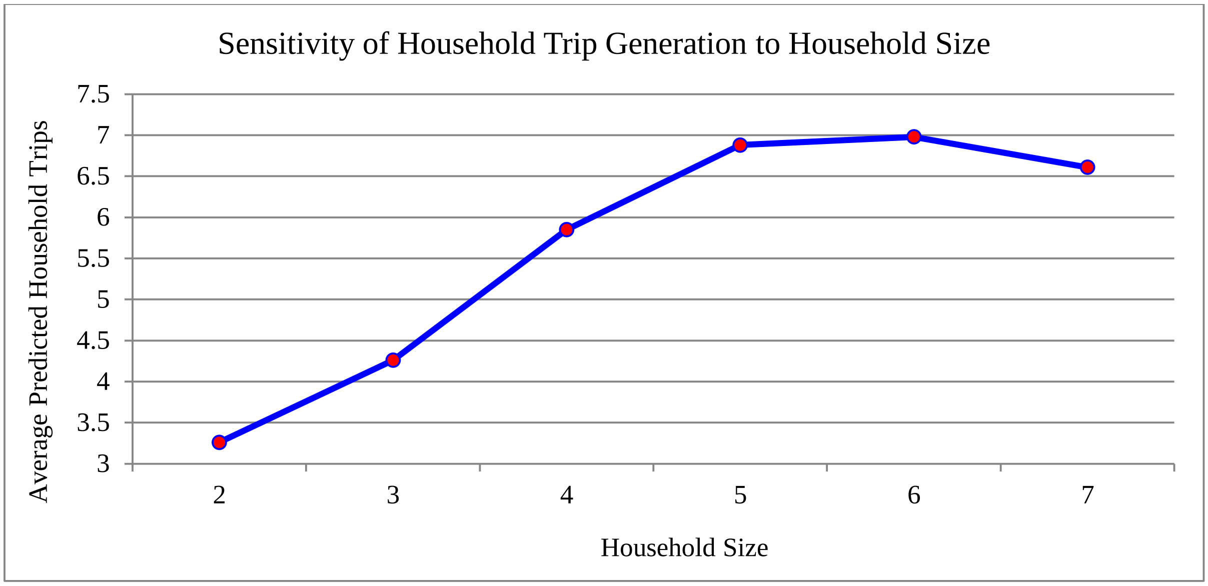

Fig. 9 Effect of household size on predicted household trip generation

The sensitivity analysis of vehicle ownership shows a steady positive relationship between vehicle ownership and household trips (Table 9 and Figure 10). The answer is not greatly different depending on the different vehicles people own, and it seems like each vehicle addition added in a similar amount of travel demand, unlike with household size. The result confirms the significance of private vehicle ownership as a measure of household mobility and Accessibility, and indicates that further expansion of private vehicle ownership will lead to a rise in travel demand, unless accompanied by effective public transport and travel demand management measures.

Table 9. Sensitivity analysis of vehicle ownership on household trip generation

Household Size	Vehicle Ownership	Predicted Household Trips	Change in Predicted Trips	Percentage Change (%)
3	0	3.78	0.00	0.00
3	1	4.26	0.48	12.72
3	2	4.74	0.48	11.28
3	3	5.23	0.48	10.14
3	4	5.71	0.48	9.21

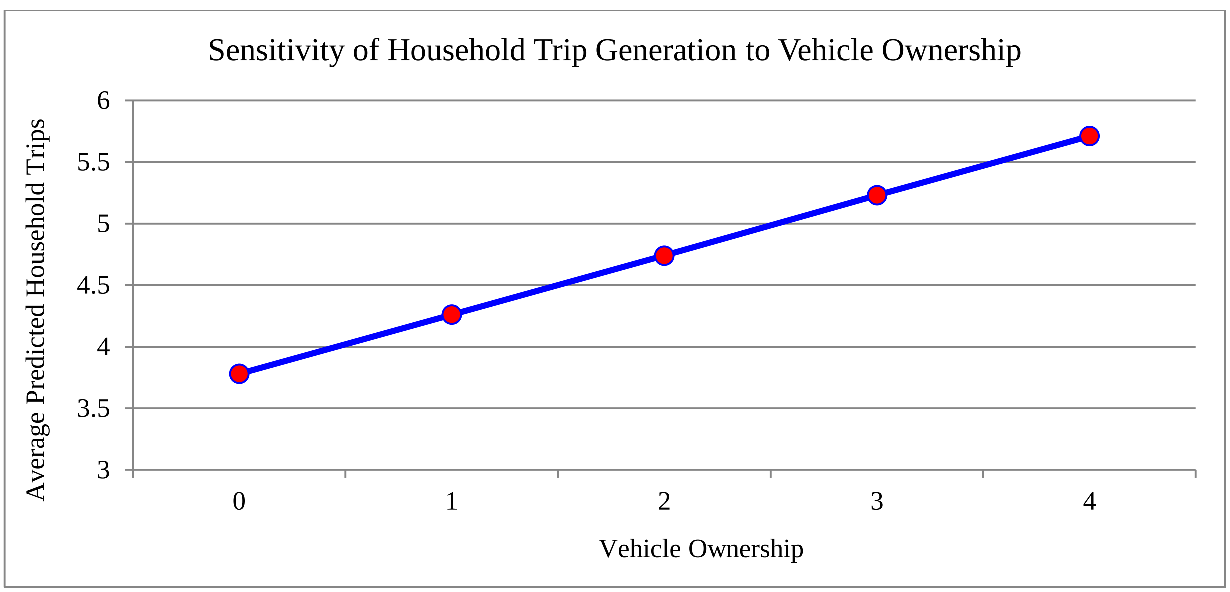

Fig. 10 Effect of vehicle ownership on predicted household trip generation

Another way to demonstrate the usefulness of the developed model for transportation planning is through scenario testing (Table 10 and Figure 11). Of the scenarios that were evaluated, the scenario of Combined Growth resulted in the largest projected increase in the number of household trips, revealing that the simultaneous growth in household size and vehicle ownership has a significantly larger impact on travel demand than either growth alone. By contrast, the Small Household Scenario saw the reduction of trips predicted, showing that the travel demand of households was sensitive to the changes. The results show that the proposed GAM is effective in predicting household trips and useful for decision-making support when assessing the future growth scenarios of the city, when forecasting the travel demand in the presence of shifting socio-economic conditions, and when providing evidence for transportation planning in bi-nuclei urban systems.

Table 10. Scenario testing using the Generalized Additive Model (GAM)

Planning Scenario	Household Size	Vehicle Ownership	Predicted Household Trips	Change in Predicted Trips	Percentage Change (%)
Baseline	3	1	4.26	0.00	0.00
Residential Growth	4	1	5.85	1.59	37.29
Higher Vehicle Ownership	3	2	4.74	0.48	11.26
Small Household Scenario	2	1	3.26	-1.01	-23.69
Combined Growth	4	2	6.33	2.07	48.55

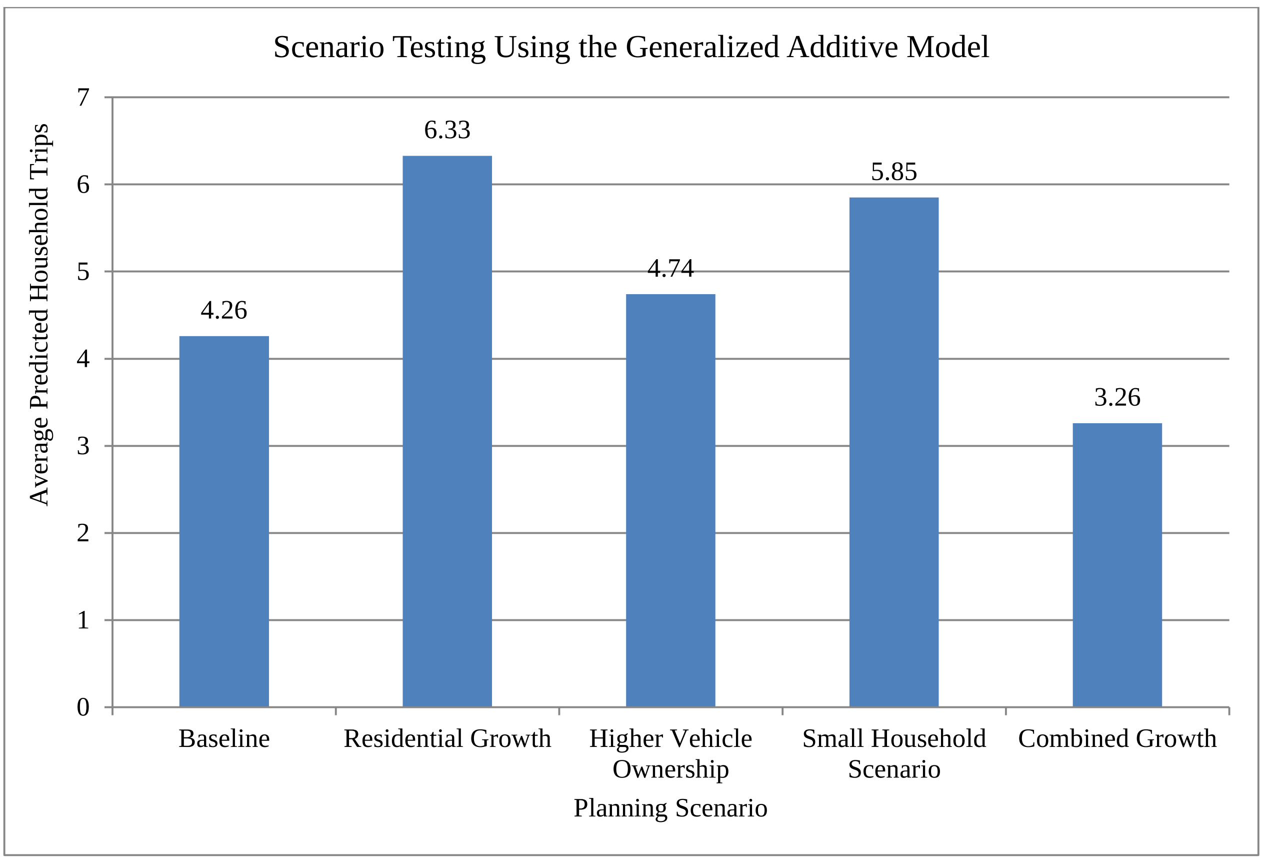

Fig. 11 Comparison of predicted household trips across planning scenarios

6.2. Policy Implications, Scenario Analysis, and Practical Recommendations

The household trip generation models developed are useful tools for evidence-based planning in bi-nuclei urban systems. Household size and vehicle ownership were found to be the most important variables influencing household trip production, and their consideration is important in transportation planning. To illustrate the applicability of the modelling framework proposed in this study, the policy implications, planning scenarios, and actionable recommendations resulting from the study findings are summarized in Table 11.

Table 11. Policy implications, scenario analysis, and actionable recommendations

Key Finding	Planning Scenario	Expected Impact	Policy Implication	Actionable Recommendation
The size of the household has a significant impact on the number of trips generated.	Development of new high-density residential areas resulting in larger household populations.	Growth in household trip production and travel demand.	Increased access requirements for road capacity, public transport, and local access.	Incorporate household growth projections into transport master plans and infrastructure investment decisions.
Vehicle ownership significantly influences household trip generation.	Increase in private vehicle ownership due to socio-economic development.	Higher private vehicle trips, traffic congestion, parking demand, and emissions.	Greater need for sustainable mobility policies and demand management.	Improve public transport services, strengthen parking management, and promote non-motorized transportation infrastructure.
The best predictive performance was obtained by the use of the Generalized Additive Model (GAM).	Adoption of advanced statistical models for travel demand forecasting by planning agencies.	More precise estimation of household trips and better forecasting.	Better-informed transportation planning and evidence-based policy making.	Adopt the GAM framework for household trip forecasting, traffic impact assessment, and long-term transportation planning.

7. Limitations and Future Research

Limitations in the present study should be noted while interpreting the results of the study. Firstly, the developed trip generation models are based on household survey data collected for the Sangli – Miraj – Kupwad Municipal Corporation, and hence their applicability to other urban areas is to be verified before their wide-scale implementation. Second, socio-economic and household characteristics are more often analysed in the study, while land-use characteristics, built-environment attributes, and emerging modes of transport like ride-sharing, micro-mobility, and telecommuting were not explicitly accounted for, as survey data for these are not available. Third, the study compares four statistical regression models and does not explore the more advanced machine learning techniques and deep learning. New mobility services, built environment data, and real-time mobility data should be introduced to enhance the predictive power of household trip generation models in the future. Finally, a comparative analysis of statistical, machine learning, and hybrid modelling techniques for various urban settings would further highlight their generalizability and applicability for transportation planning purposes.

8. Conclusion

This research shows that a multi-model approach can be applied to model household trip generation in SMKMC, with nonlinear and distributional models offering more insights than linear models. Household size and car ownership are the strongest trip generation factors. In the various models tested, the GAM offered the best overall performance, able to capture nonlinear relationships without reducing the interpretability of the model. Quantile regression also reveals differential impacts of variables at different trip levels, suggesting behavioural differences. The principal contribution of this study is a systematic comparison of four complementary statistical modelling approaches with a common household survey case data representing a Tier-2 Indian city, which can assist in identifying the most suitable and interpretable model for modelling household trip generation. The results indicate that linear models may underestimate travel patterns. The distributional and nonlinear models provide a better fit to urban travel patterns, particularly in rapidly developing Tier-2 cities. The results of the sensitivity analysis showed that a household's size has a higher impact on a household's trip generation than vehicle ownership, and the results of the scenario analysis showed how the proposed framework can be used in practice to evaluate future urban travel demand under various urban growth scenarios. Future research should include land-use, activity-based information, new modes of transport (ride sharing and micro mobility), and data from multiple years to enhance transferability, temporal stability, and policy usefulness. Study limitations include: only a single year of cross-sectional data available, which means that temporal variability and broader behavioural functions can only be inferred; and that very limited socio-economic data are available to further inform the interpretation of temporal variability. Additionally, the study did not include count-based models or spatial models. Further research can consider additional socio-economic and land-use variables, count-based and spatial models, and cross-validate the findings with more years of data to enhance the accuracy and the utility of the model. In general, it is concluded that the suggested modelling approach is a combination of model comparison, sensitivity analysis, and scenario testing that is more comprehensive in understanding household travel behaviour than a single modelling approach. Based on these findings, this paper presents an operational approach to provide a decision support structure for travel demand forecasting and transportation planning in rapidly urbanizing Tier-2 cities.

References

[1] B. G. Hutchinson, *Principles of Urban Transport Planning*, McGraw-Hill, New York, 1974. [Google Scholar] [Publisher Link]

[2] Andrew Bwambale, Charisma F. Choudhury, and Nobuhiro Sanko, "Car Trip Generation Models in the Developing World: Data Issues and Spatial Transferability," *Transportation in Developing Economies*, vol. 5, no. 2, pp. 1-15, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[3] Oreoluwa Temidayo Biala, "A Comparative Study of Catboost and Artificial Neural Networks in Enhancing Trip Generation Modelling for Ilorin City," *Journal of Civil Engineering, Science and Technology*, vol. 15, no. 1, pp. 18-29, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Olanrewaju Oluwafemi Akinfala et al., "A Soft Computing Approach to Trip Generation Estimation in Lagos Metropolis, Nigeria," *Journal of Civil Engineering, Science and Technology*, vol. 13, no. 1, pp. 6-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Saber Naseralavi et al., "Machine Learning Modeling of Household Trip Generation by State Using NHTS Data," *Urban Science*, vol. 9, no. 9, pp. 1-22, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[6] M. Shahri, and M. A. Ghannadi, "Explanatory Analyses of Work Trip Generation Using Mixed Geographically Weighted Regression (mgwr)," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 10, pp. 707-714, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Muhammed Ali Çolak, and Osman Ünsal Bayrak, "Predictive Modeling of Urban Travel Demand Using Neural Networks and Regression Analysis," *Urban Science*, vol. 9, no. 6, pp. 1-25, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[8] Jaideep Mukherjee, and B. Raghuram Kadali, "A Comprehensive Review of Trip Generation Models based on Land Use Characteristics," *Transportation Research Part D: Transport and Environment*, vol. 109, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Alaa Mohammad Yousef Dodeen, "*Developing Trip Generation Models Utilizing Linear Regression Analysis: Jericho City as a Case Study*," M.S. thesis, Master of Roads and Transportation Engineering, Faculty of Graduate Studies, An-Najah National University, 2014. [Google Scholar] [Publisher Link]

[10] C.M Sreeparvathy, R.T Arjun Siva Rathan, and K. Jayakesh, "Development of Trip Generation Model Using Linear Regression for Areas in Hyderabad City," *IOP Conference Series: Earth and Environmental Science*, vol. 1326, no. 1, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Nuriye Kabakuş, and Ahmet Tortum, "Comparative Analysis of Trip Generation Models According to Household Characteristics for Developed, Developing and Non-Developed Provinces in Turkey," *Sādhanā*, vol. 44, no. 5, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] V.A. Profillidis, and G.N. Botzoris, "Statistical Methods for Transport Demand Modeling," *Modeling of Transport Demand*, pp. 163-224, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] Aaditya Bh., and Rahul T.M, "A Comprehensive Analysis of the Trip Frequency Behavior in COVID Scenario," *Transportation Letters*, vol. 13, no. 5-6, pp. 395-403, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] T.M. Rahul et al., "Gender Differences in Work Trip Generation: Evidences from Bangalore," *Transportation Research Procedia*, vol. 48, pp. 2800-2810, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Jerry Julius Edem Dzivon, and Reginald Nii Akotoa Quao, "Modelling the Influence of Socio-Demographic Attributes on Students' Trip Generation at Kwame Nkrumah University of Science and Technology (KNUST) Campus," *Urban, Planning and Transport Research*, vol. 11, no. 1, pp. 1-19, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Sameer Abu–Eisheh, Mohammad S. Ghanim, and Alaa Dodeen, "Trip Generation Model for a Developing City in an Emerging Country," *Transportation Research Interdisciplinary Perspectives*, vol. 24, pp. 1-16, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[17] Thivya P. Amalan, and R. C. Wickramasuriya, "Developing Trip Generation and Attraction Models Using High-Frequency Proxy Data," *Journal of South Asian Logistics and Transport*, vol. 4, no. 1, pp. 45-68, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] Justin S. Chang et al., "Comparative Analysis of Trip Generation Models: Results using Home-Based Work Trips in the Seoul Metropolitan Area," *Transportation Letters*, vol. 6, no. 2, pp. 78-88, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[19] Christopher R. Fleet, and Sydney R. Robertson, "Trip Generation in the Transportation Planning Process," *Highway Research Record*, vol. 240, pp. 11-31, 1968. [Google Scholar] [Publisher Link]

[20] Edward J. Kannel, and Kenneth W. Heathington, "*The Temporal Stability of Trip Generation Relationships : Technical Paper*," Indiana Department of Transportation and Purdue University, pp. 1-35, 1972. [CrossRef] [Google Scholar] [Publisher Link]

[21] Frank J. Cesario, "A Combined Trip Generation and Distribution Model," *Transportation Science*, vol. 9, no. 3, pp. 211-223, 1975. [CrossRef] [Google Scholar] [Publisher Link]

[22] C.J. Ruygrok, and P.G. Essen, "The Development and Application of a Disaggregate Poisson Model for Trip Generation," *Transport Research International Documentation*, pp. 1-27, 1980. [Google Scholar] [Publisher Link]

[23] J. M. Rickard, "Application of a Generalised Linear Modelling Approach to Category Analysis," *Transportation Planning and Technology*, vol. 13, no. 2, pp. 121-129, 1989. [CrossRef] [Google Scholar] [Publisher Link]

[24] Simon Washington, and Jean Wolf, "Hierarchical Tree-Based Versus Ordinary Least Squares Linear Regression Models: Theory and Example Applied to Trip Generation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1581, no. 1, pp. 82-88, 1997. [CrossRef] [Google Scholar] [Publisher Link]

[25] D. A. Badoe, and C. Chen, "Modeling Trip Generation with Data from Single and Two Independent Cross-Sectional Travel Surveys," *Journal of Urban Planning and Development*, vol. 130, no. 4, pp. 167-174, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[26] Daniel A. Badoe, "Forecasting Travel Demand with Alternatively Structured Models of Trip Frequency," *Transportation Planning and Technology*, vol. 30, no. 5, pp. 455-475, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[27] Kwang-Kyun Lim, and Sivaramakrishnan Srinivasan, "Comparative Analysis of Alternate Econometric Structures for Trip Generation Models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2254, no. 1, pp. 68-78, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[28] Sangrok Kim et al., "The Trip Generation Models with Time-Effects," *Journal of Korean Society of Transportation*, vol. 30, no. 1, pp. 103-112, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[29] Leta F. Huntsinger, Nagui M. Rouphail, and Peter Bloomfield, "Trip Generation Models Using Cumulative Logistic Regression," *Journal of Urban Planning and Development*, vol. 139, no. 3, pp. 176-184, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[30] Mohsen Nazem, Martin Trépanier, and Catherine Morency, "Integrated Intervening Opportunities Model for Public Transit Trip Generation–Distribution: A Supply-Dependent Approach," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2350, no. 1, pp. 47-57, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[31] Judith L. Mwakalonge, and Daniel A. Badoe, "Trip Generation Modeling using Data Collected in Single and Repeated Cross-Sectional Surveys," *Journal of Advanced Transportation*, vol. 48, no. 4, pp. 287-376, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[32] Mehran Fasihozaman Langerudi, Taha Hossein Rashidi, and Abolfazl (Kouros) Mohammadian, "Individual Trip Rate Transferability Analysis based on a Decision Tree Approach," *Transportation Planning and Technology*, vol. 39, no. 4, pp. 370-388, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[33] Ryan Westrom et al., "Multimodal Trip Generation Model to Assess Travel Impacts of Urban Developments in the District of Columbia," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2668, no. 1, pp. 29-37, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[34] Xin Wu et al., "Hierarchical Travel Demand Estimation Using Multiple Data Sources: A Forward and Backward Propagation Algorithmic Framework on a Layered Computational Graph," *Transportation Research Part C: Emerging Technologies*, vol. 96, pp. 321-346, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[35] Sameer Abu-Eisheh, and Mohammad Irshaid, "Modelling Trip Generation Using Adaptive Neuro-Fuzzy Inference System in Comparison with Traditional Multiple Linear Regression Approach," *UKSim-AMSS 22nd International Conference on Computer Modelling and Simulation*, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[36] Fan Yang et al., "A Data-Driven Approach to Trip Generation Modeling for Urban Residents and Non-Local Travelers," *Sustainability*, vol. 12, no. 18, pp. 1-15, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[37] S.M. Saleh et al., "Trip Generation And Attraction Model and Forecasting Using Machine Learning Methods," *IOP Conference Series: Materials Science and Engineering*, vol. 1087, no. 1, pp. 1-5, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[38] Zhishuai Li et al., "Urban Trip Generation Forecasting Based on Gradient Boosting Algorithm," *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)*, Beijing, China, pp. 50-53, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[39] Carlos A. Gonzalez-Calderon et al., "Service Trip Generation Modeling in Urban Areas," *Transportation Research Part E: Logistics and Transportation Review*, vol. 160, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[40] Alsayed Alsobky, "Multiple Piecewise Regression for Trip Generation Models," *Alexandria Engineering Journal*, vol. 61, no. 9, pp. 7417-7428, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[41] Safa Ali Lafta, and Mohammed Qadir Ismael, "Trip Generation Modeling for a Selected Sector in Baghdad City using the Artificial Neural Network," *Journal of Intelligent Systems*, vol. 31, no. 1, pp. 356-369, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[42] Songhua Hu et al., "Examining Nonlinearity in Population Inflow Estimation Using Big Data: An Empirical Comparison of Explainable Machine Learning Models," *Transportation Research Part A: Policy and Practice*, vol. 174, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[43] Saumya Anand, Pritikana Das, and G.R. Bivina, "Trip Generation Based on Land Use Characteristics: A Review of the Techniques Used in Recent Years," *Transportation Research*, vol. 434, pp. 367-377, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[44] Yuebing Liang et al., "Deep Trip Generation with Graph Neural Networks for Bike Sharing System Expansion," *Transportation Research Part C: Emerging Technologies*, vol. 154, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[45] Anish Kumar, Amit Kumar, and Sanjeev Sinha, "Advancing Urban Transportation Planning: A Comparative Analysis of Machine Learning and Conventional Regression-Based Trip Generation Models," *International Journal of Advanced Technology and Engineering Exploration*, vol. 11, no. 121, pp. 1768-1783, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[46] Fatemeh Nourmohammadi, Taha H. Rashidi, and Meead Saberi, “Spatial Transferability of Pedestrian Trip Generation Models,” *Transportation Research Part A: Policy and Practice*, vol. 199, pp. 1-21, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[47] Amod N. Damle, and Nilu H. Damle, *Culture of Inequality: The Changing Hindu–Muslim Relations in Maharashtra*, Routledge India, pp. 1-168, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[48] B.N. Sardesai, “The Peoples’ Awakening in Sangli State,” *Proceedings of the Indian History Congress*, vol. 54, pp. 433-440, 1993. [Google Scholar] [Publisher Link]

[49] Adinath D. Gade, and Ratnadeep Govind Jadhav, “A Spatio-Temporal Analysis of Agriculture in Sangli-Miraj-Kupwad City, Maharashtra,” *Agricultural Transformation in India Since Independence*, pp. 59-69, 2016. [Google Scholar]