DEVELOPMENT OF MODE-WISE SPEED PREDICTION MODELS FOR URBAN ROADS WITH SIDE FRICTION CHARACTERISTICS

Shishodiya Ghanshyam Singh , S. Vasantha Kumar *

School of Civil Engineering, Vellore Institute of Technology (VIT), Vellore, Tamil Nadu, India

*E-mail of corresponding author: svasanthakumar1982@gmail.com, svasanthakumar@vit.ac.in

Resume

The present study has developed regression models to study the impact of side friction factors on speeds of vehicles in urban roads. The factors that were considered are on-street parking, pedestrian movement, water stagnation, encroachment, poor road condition and wrong side driving. Three locations from Vellore and Chikmagalur cities in India were selected and vehicle speeds, side friction parameters and flow were collected through videographic survey for 6 hours duration. It was found that the speeds of vehicles were in the range of only 11-23 km/h due to side friction. The study has proposed a "simple arithmetic operand" method to overcome the multicollinearity problem and it predicts well, the speeds with mean absolute percentage error (MAPE) between 8-29%. The arithmetic operand method proposed in the present study can be applied in other fields as well to overcome the multicollinearity problem.

Article info

Received 1 February 2022
Accepted 9 May 2022
Online 13 June 2022

Keywords:

side friction
vehicle speed
prediction
regression
multicollinearity
principal component regression

Available online: https://doi.org/10.26552/com.C.2022.3.E108-E121

ISSN 1335-4205 (print version)
ISSN 2585-7878 (online version)

1 Introduction

An efficient transportation system should ensure easy mobility by free flow of vehicles without any traffic congestion [1-3]. Properly maintained carriageways are essential to meet this goal as there are features on carriageway, which normally hinder the traffic flow. Popularly called as side friction in literature can be in the form of pedestrian walking on the carriageway, on-street parking, road side encroachment, bus stops, wrong movement of vehicles, etc. These side friction elements occupy some portion of the carriageway and thus result in reduction of carriageway width. This phenomenon leads to reduced capacity of the highway and poor level of service (LOS) besides affecting the speeds of vehicles. As seen in Figure 1, due to non-availability of proper sidewalk, the pedestrians are forced to walk on the carriageway and this causes the vehicles to reduce their speeds as the carriageway width has got reduced. These side frictions not only reduce the speeds of vehicles and capacity of the roads but they raise the safety concerns as well, especially for the pedestrians [4].

Studies have been carried out worldwide on the effect of side friction on vehicle speeds and they are reviewed in Table 1. It can be seen that the side friction is a common phenomenon in most of the countries, especially in the developing ones as there are studies reported from India, Indonesia, Africa etc. From the review of literature, one can figure out that the roadside friction activities such as pedestrians, bus stops, entry/exit maneuvers, wrong movement of vehicles, non-motorized traffic and on-street parking were considered mostly and their impacts on the vehicular speeds were analyzed. However, there are other side friction elements, such as water stagnation on roads owing to lack of proper drainage system, carriageway encroachment through road side shops and poor road condition due to improper maintenance. The impact of these side frictions, which likely tends to obstruct the traffic flow, has not been studied before in any of the reported studies. This was one of the main motivations aimed for carrying the present work in which each of these side frictions were taken into account and their impact was studied. One common thing, observed from Table 1, is that in many studies, statistical analysis of explanatory variables was not seriously taken into consideration. For example, R^2 and p-values were not reported in many studies. Even if the R-square was less and/or p-values were more than 0.05, but still the variables were considered and the models were used. The possible reason for not achieving the expected R^2 and p-value is due to the existence of redundancy in the form of multicollinearity

***Figure 1** Side friction on Indian roads*

***Table 1** Studies on side friction*

Sl. No	Authors	Side friction considered	Model fitting based on R^2 value*	Statistical significance of the variables based on p-values**	Whether the problem of Multicolli-nearity addressed	Country where the study was carried out
1	Biswas et al. [5]	On-street parking, Pedestrian movement and Non-motorized vehicles (NMV)	N/A	N/A	No	India
2	Alkaissi and Kamoona [6]	On-street parking, Pedestrians movement	Good	N/A	No	Iraq
3	Mahendra et al. [7]	Pedestrian movement, On-street parking, Entry/ Exit of vehicles, Slow moving vehicles	Moderate	N/A	No	Indonesia
4	Patkar and Dhamaniya [8]	Bus stops, NMV and Pedestrian movement	Good	Significant	No	India
5	Yeboah et al. [9]	Pedestrian movement	N/A	N/A	No	Ghana
6	Bitangaza and Bwire [10]	Pedestrian movement, On-street parking	Good	N/A	No	Rwanda
7	Chauhan et al. [11]	NMV, Roadside parking	N/A	N/A	No	India
8	Najid [12]	NMV, Pedestrian movement, Left-Right Access, Roadside parking	N/A	N/A	No	Indonesia
9	Hidayati et al. [13]	Pedestrian movement, Roadside parking vehicles, Entry/Exit of vehicles	N/A	N/A	No	Indonesia
10	Gulivindala and Mehar [14]	Wrong movement of vehicles, Pedestrian movement, Stopped vehicles and Entry/ Exit manoeuvre	N/A	Moderate	No	India
11	Salini et al. [15]	Bus Stops, Pedestrian movement and Roadside parking	Good	Not Significant	No	India
12	Rao et al. [16]	Pedestrian movement, Roadside parking, NMV, Entry/Exit of vehicles	Poor	Not Significant	No	India
13	Bansal et al. [17]	Bus stops	Poor	Significant	No	India
14	Chiguma [18]	Pedestrian movement, NMV, Roadside parking	Moderate	Not Significant	No	Tanzania
15	Munawar [19]	Pedestrian movement, Roadside Parking, Entry/Exit of Vehicles, Bus Stops	N/A	N/A	No	Indonesia

*R^2 value > 0.7 is Good, 0.5 - 0.7 is Moderate, < 0.5 is Poor [20]

**p- value < 0.05 means the independent variable is significant; N/A = Not Available.

in the dataset, i.e. higher inter-correlations between the two or more independent variables (IVs) in the model. However, none of the studies reviewed before have studied this multicollinearity effect and the ways to overcome it. Hence, in the present study after analyzing the conventional methods like principal component regression (PCR), a new method called "simple arithmetic operand" was proposed to remove this multicollinearity effect and achieve the desired R^2 and p-values. This was one of the main contributions of the present study, i.e. a new method to remove the multicollinearity among the independent variables, which can be applied in other fields, as well, where the multicollinearity is a potential problem while regressing the variables. The details of study area, data collection and data extraction are explained in the following section.

2 Study area

The problem of side friction can be mainly seen in cities and busy towns where the commercial activities are predominant. Hence, in the present study, a city called Vellore located in Tamil Nadu state of India and a busy town in Karnataka state of India called Chikmagalur were considered. Vellore, which has a population of around 0.5 million is one of the fast-growing cities in the state and it is considered as an important educational hub as VIT University and Christian Medical College and Hospital (CMCH) are located in Vellore. In addition, the golden temple and Vellore fort attracts large number of tourists in recent years. The other study area, called Chikmagalur, is a famous hill station in the state of Karnataka and it attracts tourists in huge amounts

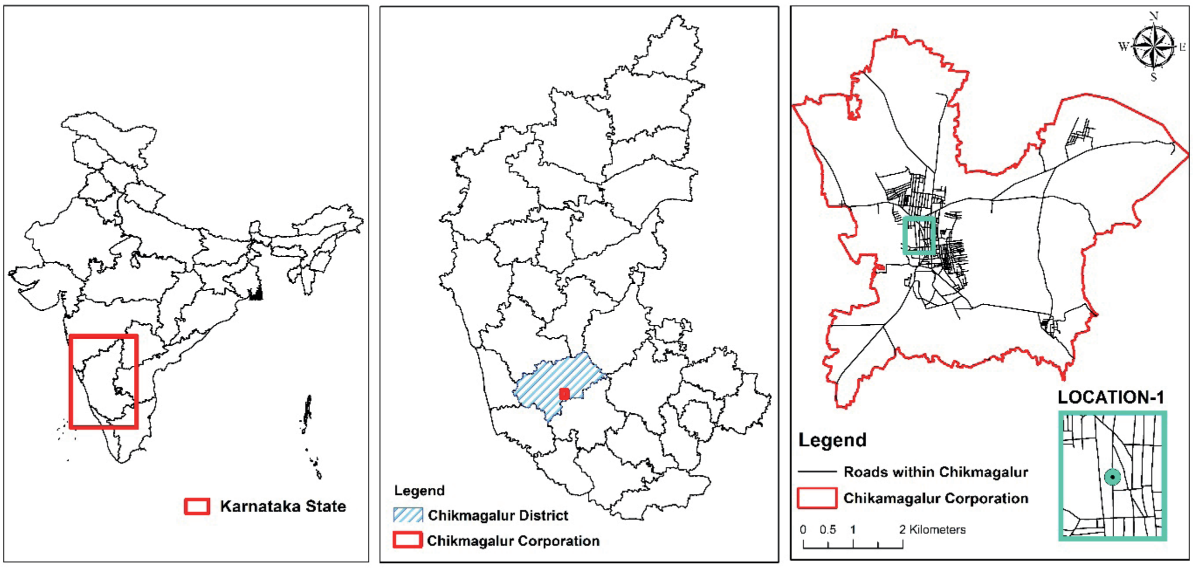

Figure 2 *Map showing the location - 1 in Chikmagalur town in Karnataka, India*

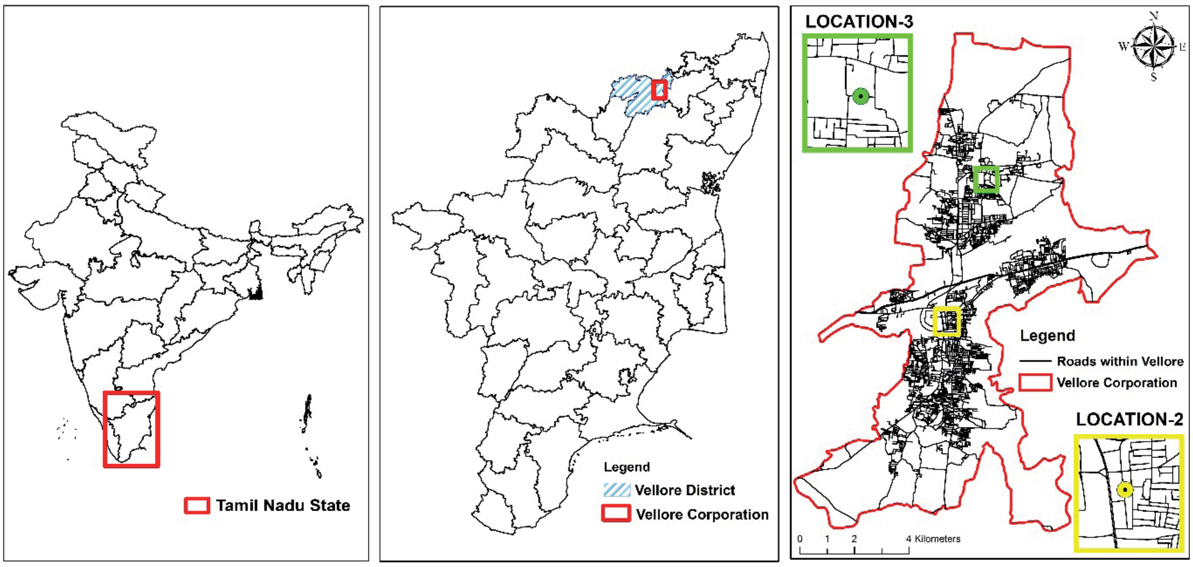

Figure 3 *Map showing the locations - 2 and 3 in Vellore city in Tamil Nadu, India*

***Table 2** Description of the study area*

Name of the study location	Nature of road	Type of side friction
Indira Gandhi (IG) Road, Chikmagalur (Location - 1)	Intermediate undivided arterial road with 10 m wide carriageway (including both directions)	Water stagnation on road, Pedestrian movement, Wrong-side driving.
Long Bazzar Road, Vellore (Location - 2)	4 lane divided road with two lanes per direction (7.5 m wide carriageway per direction)	Encroachments in the form of road-side shops, On-street parking, Pedestrian movement, Wrong-side driving.
VG Rao Nagar Road, Vellore (Location - 3)	Intermediate undivided collector street with 5.5 m wide carriageway	Poor road condition.

***Figure 4** Photos showing the study locations (a) IG road, Chikmagalur; (b) Long bazaar road, Vellore; (c) VG Rao nagar road, Vellore*

almost throughout the year. After careful investigation, three locations were selected as study stretches - one in Chikmagalur and two in Vellore. Figures 2 and 3 shows the maps of three locations selected for the present study and details of the selected locations such as nature of the road and type of the side friction are given in Table 2.

The locations were chosen in such a way that they cover the side frictions, which have not been studied before. For example, in Indira Gandhi (IG) road of Chikmagalur, water stagnation on carriageway is a major side friction, as seen in Figure 4a. It can be noticed that due to water stagnation, almost two thirds of the carriageway in one direction is not utilized, as the vehicles are not using the road space covered with water. This creates a kind of bottleneck situation where the vehicles are forced to reduce its speed when passing through this area. In addition to water stagnation, pedestrian movement on carriageway and wrong-side driving were also noticed in IG road and hence were taken into account. This IG Road is considered as the main commercial hub for the tourists coming to Chikmagalur as it contains the hotels, restaurants, shopping malls and the presence of Karnataka State Road Transport Corporation (KSRTC) bus station makes this road one of the busiest roads in Chikmagalur.

In the case of Vellore, two locations were taken, one amidst of the main inner city comprising CBD (central business district), namely the Long bazaar road, in addition to other one in a residential area called VG Rao Nagar. The study stretch in Long bazaar of Vellore is basically a four-lane divided urban arterial with dual lanes per direction of travel but presence of encroachments in the form of road-side shops and on-street parking occupied almost one whole lane in each direction, as can be seen in Figure 4b. In order to put restrictions on these encroachments, the civic authorities have constructed a median few years ago. However, this median does not serve its original purpose and encourages illegal parking as the two-wheelers have parked their vehicles on both the sides of the median, as seen in Figure 4b. As like in IG road of Chikmagalur, in Long bazaar of Vellore also pedestrians were found to walk on the carriageway and wrong side movement were observed. Hence, these two variables were considered as the side frictions in addition to the road-side encroachment and on-street parking. The two-lane collector street, located in a residential area called VG Rao Nagar, was another study stretch from Vellore. A photograph of the location is shown in Figure 4c. As seen from the photo, there is no proper road constructed with asphalt or concrete and only earthen road is available for the motorists to use. Such a poor road condition is one of the side frictions, which can be seen in interior streets of tier-2 cities in India like Vellore. Though the VG Rao Nagar comes under Vellore municipal corporation (selected for smart city mission by Government of India), the roads are very poor and act as a speed barrier for the vehicles. Such bumpy roads not only cause the vehicles to reduce their speed while passing through the section, but they cause many health-related problems, as well, like the back pain and spinal cord damage especially for the two-wheelers. These earthen roads, where thousands of vehicles plying on it every day, create a dusty environment in the entire location. All the three study stretches from Chikmagalur and Vellore are midblock sections without any intersections nearby, thus, the drop in vehicles' speeds would be mainly due to the prevailing side frictions activities.

3 Data collection and extraction

The data for all the three locations was collected using videographic method with a handycam mounted on the tripod. High rise buildings, which are just adjacent to the study stretch were chosen as the vantage points to get a clear vision of traffic stream. As traffic and side friction activities would be generally more during weekdays, the video data collection also was carried out on typical working days, i.e. 16th August 2019 (Friday), 10th February 2020 (Monday) and 5th February 2020 (Wednesday) for the locations-1, 2 and 3 respectively. As various commercial activities begin from 10 a.m. in IG road of Chikmagalur and Long Bazaar of Vellore, the vehicular and pedestrian population will also start increasing from 10 a.m. onwards. Hence, the data collection was carried out from 10 a.m. to 12 p.m. for a duration of 2 h at these two locations (1 and 2). For the study stretch in VG Rao Nagar of Vellore, the video recording was carried out for a period of 8 a.m. to 10 a.m. The reason for choosing this particular time is VG Rao Nagar is basically a residential area with many schools and colleges nearby. Hence, the educational and office trips were found to be more during this period of 8-10 a.m. and so the data collection was done during this time using a tripod mounted handycam.

The extraction of required data from video depends on variables that are going to be used in the regression model. In this study, the velocity of a specific class of vehicle has been taken as the dependent variable (DV) and the speed corresponding to other classes of the vehicles, side friction elements and flow of vehicles, were considered as independent variables (IV's). Hence, for building up the regression model, the data required are speed, side friction parameters and flow corresponding to different categories of vehicles (four wheeled, three wheeled and two wheeled). The traffic volume count (flow) was done manually by counting class-wise categories of vehicles for every 5 minutes time interval by using the recorded video. Speed of vehicles can be obtained by dividing the distance by the corresponding travel times of vehicles. In order to get the travel times, a length of 63 m, 34 m and 60 m were marked on the road at the three study locations, namely, IG Road, Long Bazaar road and VG Rao Nagar road, respectively. The video taken at the study locations was played and the travel times of vehicles to cover the above mentioned lengths were then noted down using a digital stop watch. With the known distance and time of travel, speed of vehicles was then determined. As the side frictions, namely, the water stagnation, poor road condition and encroachment were taken as categorical variables with presence or absence, there is no data required to be extracted from the video. However, the side frictions like pedestrian movement and wrong movement vehicles were calculated manually for each 5 min time interval.

4 Methodology

The present study has developed the mode-wise regression models with help of data collected from study stretches in Chikmagalur and Vellore. Since each vehicle type exhibit different speed characteristics, separate multiple linear regression equations were built corresponding to each mode. The two wheeled vehicles would generally drive fast, despite having compact space available between other modes of vehicles, due to the smaller vehicle sizes, which may not be possible with three wheeled and four wheeled vehicles. The general form of the regression models that are developed for each mode are given below.

a) The four wheeled vehicles

$$y_{FW} = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 + \alpha_5 x_5 + \alpha_6 x_6 + \alpha_7 x_7 + \alpha_8 x_8 + \alpha_9 x_9, \quad (1)$$

where:

y_{FW} = mean speed of the four wheeled vehicles (km/h) for each 5 min. duration,

x_1 = mean speed of the three wheeled vehicles (km/h) for each 5 min. duration,

x_2 = mean speed of the two wheeled vehicles (km/h) for each 5 min. duration,

x_3 = categorical variable (1 if poor road condition, 0 otherwise),
x_4 = categorical variable (1 if road-side encroachment and on-street parking, 0 otherwise),
x_5 = number of pedestrians moving on the carriageway in each 5 min. interval,
x_6 = number of wrong-side driving vehicles in each 5 min. interval,
x_7 = number of the four wheeled vehicles (flow) observed for each 5 min. duration,
x_8 = number of the three wheeled vehicles (flow) observed for each 5 min. duration,
x_9 = number of the two wheeled vehicles (flow) observed for each 5 min. duration.

b) The three wheeled vehicles

$$y_{ThW} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9, \quad (2)$$

where:
y_{ThW} = mean speed of the three wheeled vehicles (km/h) for each 5 min. duration,
x_1 = mean speed of the four wheeled vehicles in (km/h) for each 5 min. duration,
x_2 = mean speed of the two wheeled vehicles in (km/h) for each 5 min. duration.
x_3 to x_9 as defined before in Equation (1).

(c) The two wheeled vehicles

$$y_{TW} = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 + \gamma_3 x_3 + \gamma_4 x_4 + \gamma_5 x_5 + \gamma_6 x_6 + \gamma_7 x_7 + \gamma_8 x_8 + \gamma_9 x_9, \quad (3)$$

where:
y_{TW} = mean speed of the two wheeled vehicles (km/h) for each 5 min. duration,
x_1 = mean speed of the four wheeled in (km/h) for each 5 min. duration,
x_2 = mean speed of the three wheeled in (km/h) for each 5 min. duration.

In the above equations, α_0, β_0 and γ_0 are the regression constants and α_1 to α_9, β_1 to β_9 and γ_1 to γ_9 are the regression coefficients for the independent variables of the four wheeled, three wheeled and two wheeled vehicles, respectively. The independent variables that were chosen indicate that the speed of a particular category of vehicles in a traffic stream is influenced by speeds of other vehicles, flow or volume and the available side frictions. As per traffic flow theory, for a given flow there could be two extreme cases possible one in the free-flowing state and the other one in the congested state [21]. This says that, just based on flow alone, one cannot ascertain the type of the traffic state. Hence, in the present study, while selecting the independent variables, speeds of other categories of vehicles were also taken into account in addition to flow, because for a given flow, if the traffic is free flowing, the vehicles would go at high speeds, whereas if it is congested, the speeds would be less. As seen from Figure 4, the presence of side frictions also plays a major role in speeds of vehicles. Hence, they were taken into account in the regression models, as seen from Equations (1) to (3). There are total of five side frictions, namely, poor road condition, road-side encroachment and on-street parking, water stagnation, pedestrians movement and wrong-side driving. Out of these five side frictions, the first 3 were taken as categorical variables (presence or absence) and the last 2 were taken as continuous variables. Normally in regression analysis, if there are "n" levels to represent as dummy or categorical variable, then there will be "n-1" dummy variables in the regression model [22]. In the present study, there are 3 categorical responses, which are need to be incorporated and so 2 variables were included, namely, x_3 and x_4. If both x_3 and x_4 were zero, then that indicates the water stagnation condition. A total of 720 observations for each vehicular mode were taken into account for generating the multiple linear regression equations [12 data points/h × 2 h × 3 locations × 10 variables (1 dependent variable and 9 independent variables)]. Thus, an aggregate of 2160 set of measurement points (720×3) was considered for development of all the three multiple linear regression models. While regressing the independent variables, if a specific type of roadside friction activity is absent, then it was treated to be zero. Such as, in VG Rao Nagar data, x_3 (poor road condition) was taken as 1, whereas for the other two locations, it was zero.

Once the regression models were built, the R^2 and p-values were checked to know the strength of the models developed. If R^2 is exceeding the value of 0.70 and probability value remains below 0.05, then the proposed model is reliable and it is statistically significant [20]. However, sometimes, due to the high correlation among the independent variables, popularly called as multicollinearity, the p-values of some variables may exhibit values more than 0.05. The problem with multicollinearity is that the regression coefficient estimates become imprecise and fluctuate drastically and may lead to poor predictions, as well [23-24]. Multicollinearity may result in indeterminate estimates as explained below.

The two variable case is assumed, expressed in the standardized form (subtracting mean and dividing by standard deviation).

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \quad (4)$$

where: i (*observation*) = 1, 2, ... , n (*sample size*).

As $\overline{y} = \beta_0 + \beta_1 \overline{x_1} + \beta_2 \overline{x_2}$, Equation (4) becomes

$$\begin{aligned} y_i - \overline{y} &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} - (\beta_0 + \beta_1 \overline{x_1} + \beta_2 \overline{x_2}) \\ &= \beta_1 (x_{1i} - \overline{x_1}) + \beta_2 (x_{2i} - \overline{x_2}). \end{aligned} \quad (5)$$

If one divides Equation (5) by standard deviation on both sides

$$\frac{y_i - \overline{y}}{S_y} = \left(\beta_1 \frac{S_{x1}}{S_y}\right)\frac{x_{1i} - \overline{x_1}}{S_{x1}} + \left(\beta_2 \frac{S_{x2}}{S_y}\right)\frac{x_{2i} - \overline{x_2}}{S_{x2}}. \quad (6)$$

Equation (6) can be denoted as

$$y_i{}^* = \beta_1{}^* z_1 + \beta_2{}^* z_2\,, \quad (7)$$

where: $y_i{}^* = \frac{y_i - \overline{y}}{S_y}, z_{1i} = \frac{x_{1i} - \overline{x_1}}{S_{x1}}, z_{2i} = \frac{x_{2i} - \overline{x_2}}{S_{x2}}$.

Equation (7) is the standardized form of dependent and independent variables with mean zero and standard deviation one. The matrix form of Equation (7) is $Y^* = Z\beta^*$, w where:

$$Y^* = \begin{bmatrix} y_1{}^* \\ y_2{}^* \\ \cdot \\ \cdot \\ \cdot \\ y_n{}^* \end{bmatrix}, Z = \begin{bmatrix} Z_{11} & Z_{21} \\ Z_{12} & Z_{22} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ Z_{1n} & Z_{2n} \end{bmatrix}, \beta^* = \begin{bmatrix} \beta_1{}^* \\ \beta_2{}^* \end{bmatrix}..$$

The least square estimator of β^* is

$$\beta^* = (Z'Z)^{-1}Z'Y^* = \frac{1}{(n-1)(1-r_{z1z2}{}^2)} \times \begin{bmatrix} 1 & -r_{z1z2} \\ -r_{z1z2} & 1 \end{bmatrix}\begin{bmatrix} (n-1)r_{z1y^*} \\ (n-1)r_{z2y^*} \end{bmatrix} = \frac{1}{(1-r_{z1z2}{}^2)} \times \begin{bmatrix} r_{z1y^*} - r_{z1z2} \cdot r_{z2y^*} \\ r_{z2y^*} - r_{z1z2} \cdot r_{z1y^*} \end{bmatrix}. \quad (8)$$

Thus, the standardized regression coefficients are

$$\beta_1{}^* = \frac{r_{z1y^*} - r_{z1z2} \cdot r_{z2y^*}}{1 - r_{z1z2}{}^2}, \quad (9)$$

$$\beta_2{}^* = \frac{r_{z2y^*} - r_{z1z2} \cdot r_{z1y^*}}{1 - r_{z1z2}{}^2}, \quad (10)$$

where, r_{z1y^*} is the correlation coefficient between z_1 and y^*, r_{z2y^*} is the correlation coefficient between z_2 and y^* and r_{z1z2} is the correlation coefficient between z_1 and z_2. $\beta_1{}^*$ and $\beta_2{}^*$ can also be obtained directly using $\beta_1 \frac{S_{x1}}{S_y}$ and $\beta_2 \frac{S_{x2}}{S_y}$ respectively as seen in Equations (6) and (7). From Equations (9) and (10), can be seen that, if the independent variables are highly correlated, $|r_{z1z2}| \to 1$, what would result in poor estimates of $\beta_1{}^*$ and $\beta_2{}^*$ as denominators in Equations (9) and (10) would approach zero. In addition, the variances $Var\beta^* = (Z'Z)^{-1}\sigma^2$ will become very large and approach ∞ when $|r_{z1z2}| \to 1$. Such large variances imply the poor and indeterminate estimates of the regression coefficients. Hence, if there exists highly correlated variables with one another in a regression model, multicollinearity pose a serious issue that has to be estimated for accuracy of the models.

Detection of multicollinearity can be achieved by calculating "tolerance" for each independent variable. Tolerance is computed by subtracting R^2 from 1. HereR^2 is obtained by regressing an independent variable over the remaining IV's. Whenever the least correlation exists between considered IV's, it indicates R^2 to be small and, thus, the tolerance become large. On the other hand, if strong correlation is found among the independent variables, R^2 would be high and the tolerance would be low. Thus, the terms tolerance and multicollinearity are inversely proportional to each other. In general, any tolerance value less than 0.2 indicates the presence of high degree of multicollinearity in the dataset used for model development [25-26]. Hence, the tolerance value of 0.2 was considered as a benchmark for indicating the existence or non-existence of multicollinearity in the dataset for the present study also. If the multicollinearity is detected, methods like principal component regression (PCR) can be done to remove the multicollinearity. In the present study, a new method, called "simple arithmetic operand" method, was proposed by defining a new set of independent variables obtained by adding, subtracting, multiplying or dividing the existing independent variables. Both the PCR and the proposed method were evaluated for its statistical significance by checking the R^2, p-values and tolerance. The model was then validated by taking only 75% of the data and predicting for the remaining 25% data. Predicted speeds were then checked with the observed speeds for similarity and the accuracy of prediction was reported through one of the popular error measures called "Mean Absolute Percentage Error (MAPE)". The reason for choosing MAPE is that it has many advantages when compared to other error measures, such as mean absolute error (MAE) or root mean squared error (RMSE). For example, MAPE is scale independent, whereas the other two error measures are scale dependent, i.e. they depend on the range of the target variable and hence cannot be compared across the datasets. As the present study compares the performance of the mode-wise regression models, it would be advantageous if one uses MAPE rather than MAE or RMSE. In addition, if the value of MAPE is below 10%, then prediction can be considered excellent, 10-20% is considered good and if MAPE ranges between 20-50% then its reasonable, whereas MAPE exceeding more than 50% is unacceptable [27-28]. Thus, based on MAPE, one can ascertain how good the prediction results are, which may not be possible with MAE or RMSE. Results are discussed in the following section.

5 Results and discussion

In this section, the findings were presented in two parts. The first part explains the basic descriptive statistics for the collected data. The second part deals with development of multiple linear regression model, treatment for multicollinearity, which is then

Table 3 *Descriptive statistics*

	Indira Gandhi (IG) road, Chikmagalur			
	Mean	Std. Dev.	Min.	Max.
Mean speed of the four wheeled vehicles	15.11	2.31	9.43	19.34
Mean speed of the three wheeled vehicles	17.47	2.74	10.61	22.29
Mean speed of the two wheeled vehicles	21.37	1.01	19.7	23.26
Number of pedestrians (x_5)	49.38	11.5	32	77
Wrong-side driving (x_6)	16.33	3.76	10	24
Number of the four wheeled vehicles (x_7)	52.33	10.01	33	78
Number of the three wheeled vehicles (x_8)	33.92	4.56	25	42
Number of the two wheeled vehicles (x_9)	84.46	9.94	67	103
	Long bazzar road, Vellore			
	Mean	Std. Dev.	Min.	Max.
Mean speed of the four wheeled vehicles	2.67	4	0	10.68
Mean speed of the three wheeled vehicles	10.6	1.85	7.99	14.85
Mean speed of the two wheeled vehicles	16.5	3.05	11.31	21.89
Number of pedestrians (x_5)	72.42	12.84	46	97
Wrong-side driving (x_6)	27.92	7.57	10	44
Number of the four wheeled vehicles (x_7)	0.96	1	0	3
Number of the three wheeled vehicles (x_8)	6.88	2.64	3	13
Number of the two wheeled vehicles (x_9)	44.83	10.13	27	63
	VG Rao nagar road, Vellore			
	Mean	Std. Dev.	Min.	Max.
Mean speed of the four wheeled vehicles	9.3	1.17	7.35	11.97
Mean speed of the three wheeled vehicles	5.76	5.08	0	12.67
Mean speed of the two wheeled vehicles	13.66	1.07	11.95	16.03
Number of pedestrians (x_5)	0	0	0	0
Wrong-side driving (x_6)	0	0	0	0
Number of the four wheeled vehicles (x_7)	5.88	2.82	1	13
Number of the three wheeled vehicles (x_8)	1.42	1.64	0	5
Number of the two wheeled vehicles (x_9)	32.83	7.8	20	48

followed by validation of the developed model.

5.1 Descriptive statistics

The statistics, like mean, standard deviation (SD), minimum and maximum values, were calculated location-wise for the speed, flow and side friction data collected and the results are shown in Table 3. The categorical variables x_3 and x_4 were not considered for calculation of descriptive statistics, as they take the values of 0 or 1. Many interesting results can be obtained by looking into Table 3. Overall, the two-wheeled vehicles speeds were high when compared to that of other vehicle categories in all the 3 locations considered. As said before, the two-wheeler being the smallest vehicle in terms of size, possesses the ability to freely maneuver between the large seized vehicles and, hence, the average speed is high when compared to others. Due to this phenomenon, the two-wheelers did not exhibit completely stopped condition ("zero" speed) in any of locations, whereas other vehicles have stopped completely and reported zero speed under the "minimum speed" category as seen from Table 3. One can notice in Table 3 that the maximum speed ranges between 19-23 km/h, 11-22 km/h and 12-16 km/h, for IG road, Long bazaar road and VG Rao nagar road, respectively. Though the central government has raised the speed limits of cars and two-wheelers in Indian cities to 70 km/h and 50 km/h respectively [29], however the maximum speeds of cars and two-wheelers were found to be only 19 km/h and 23 km/h, respectively, as seen from Table 3. This is mainly due to the presence of different side frictions on the urban roads and unless they are removed it is difficult to achieve increment in the speed of vehicles in India. The results suggest that the minimum, maximum and average speed of vehicles were different for different vehicle categories and this clearly indicates the need for the mode-wise speed prediction models as attempted in the present study.

Table 4 Results of the mode-wise regression models

	Four wheeled (R^2 = 0.815)			Three wheeled (R^2 = 0.720)			Two wheeled (R^2 = 0.811)		
Variable	Reg. Coeff. value	p value	Tolerance	Reg. Coeff. value	p value	Tolerance	Reg. Coeff. value	p value	Tolerance
Intercept	35.494	1.84E-05	-	-0.672	0.952	-	28.106	6.05E-08	-
x_1	0.008	0.936	0.280	0.012	0.936	0.184	-0.055	0.500	0.185
x_2	-0.131	0.500	0.189	0.578	0.016	0.206	0.154	0.016	0.307
x_3	-21.041	0.000	0.014	-1.321	0.865	0.011	-13.997	0.000	0.014
x_4	-21.829	6E-06	0.022	1.840	0.780	0.016	-7.673	0.021	0.017
x_5	-0.007	0.831	0.072	0.0009	0.984	0.072	-0.067	0.003	0.082
x_6	-0.114	0.104	0.131	-0.057	0.520	0.127	0.034	0.460	0.127
x_7	-0.045	0.407	0.058	-0.045	0.515	0.057	-0.018	0.600	0.057
x_8	-0.165	0.129	0.039	0.292	0.032	0.041	-0.099	0.165	0.039
x_9	-0.088	0.016	0.135	-0.012	0.790	0.123	-0.017	0.480	0.123

"-" implies not applicable

From the descriptive statistics was found that, though the Long bazaar road of Vellore has bigger carriageway width, compared to the IG road, the average speeds of vehicles were smaller than those of IG road. The possible reason for that is presence of the side friction activities in the form of roadside shops and on-street parking in the Long bazaar road, as seen in Figure 4b. When the effective carriageway width available at Long bazaar was measured using the tape, it was shocking to notice that out of 7.5 m carriageway, only 2.74 m was available for vehicles to go. The remaining portion was completely encroached by the side friction activities and this may be the reason for lower speeds in the Long bazaar road. The average speed of the four wheelers was found to be only 2.67 km/h in Long bazaar, which is far less when compared to other locations' four-wheelers speeds. The possible reason for this is that the Long bazaar is a market area where more pedestrians walk on the carriageway due to absence of proper footpath. This has been shown in Table 3 where the pedestrian count was higher (72 pedestrians/5 min) in the Long bazaar than in any other location. In such a scenario, considering the pedestrians' safety, the drivers of the four-wheelers are forced to reduce their speed drastically, as the pedestrians walk on the carriageway. The results indicate the need for proper sidewalks also in locations where there are pedestrian activities, in addition to removal of side frictions like the on-street parking and roadside shops.

5.2 Development of the speed prediction models

Table 4 shows the multiple linear regression model results obtained for all the modes (four wheeled, three wheeled and two wheeled vehicles). It includes the coefficient of determination (R^2), estimated regression coefficients, p-value for each of the independent variables, as well as the tolerance values. It was found that R^2 was good as it is more than 0.7 in all the three modes. In addition, the signs of the coefficients were logical as most of the side friction and flow variables exhibited negative sign. For example, the negative coefficients of categorical variables x_3 and x_4 indicate that if the road is poor in condition or encroached by the road-side shops with on-street parking or water stagnation is there, then speeds will get reduced. Similarly, if the flow is increasing then speeds of vehicles would get reduced. Though the developed mode-wise regression models have good R^2 values and logical signs, however the explanatory variables are not statistically significant, as most of the coefficients have p-values more than 0.05, as seen from Table 4. In addition, the tolerance was found to be less than 0.2 in 24 out of 27 tolerance values calculated and this indicates the presence of multicollinearity in the given dataset. It is logical to expect multicollinearity, i.e. correlation among the independent variables because if there are road-side shops, it may attract people to purchase the goods, which in turn leads to gathering of people, on-street parking and pedestrian movement and finally leading to reduction of vehicles speed in that road stretch.

In order to break the multicollinearity between the independent variables in the model, the conventional approach is to use methods like principal component regression (PCR) or ridge regression. In this study, the PCR was used to check whether it can remove the multicollinearity or not. The PCR involves basically the conversion of given data of independent and dependent variables to standardized form and then calculating the covariance matrix. The covariance matrix was then used to calculate the eigen values and vectors, which were then used to calculate the principal components. The results of the PCR, done using NCSS 2020 statistical software [30], is shown in Table 5. As before, the R^2 values were good as they are more than 0.7 for all the three models. Since some degree of bias is added to the regression estimates in the PCR, it is not meaningful

Table 5 *Results of mode-wise regression models developed using PCR*

	Four wheeled (R^2 = 0.762)			Three wheeled (R^2 = 0.719)			Two wheeled (R^2 = 0.781)		
Variable	Reg. Coeff. value	p value	Tolerance	Reg. Coeff. Value	p value	Toleance	Reg. Coeff. value	p value	Tolerance
Intercept	6.139	-	-	1.876	-	-	14.509	-	-
x_1	-0.007	0.949	0.280	-0.006	0.962	0.244	0.075	0.326	0.249
x_2	0.206	0.305	0.228	0.555	0.011	0.247	0.173	0.011	0.310
x_3	1.323	0.466	0.174	-3.145	0.136	0.162	-3.251	0.010	0.143
x_4	-5.480	0.026	0.097	0.371	0.889	0.099	1.798	0.196	0.115
x_5	0.018	0.661	0.074	-0.0007	0.988	0.073	-0.069	0.005	0.082
x_6	-0.064	0.413	0.135	-0.063	0.469	0.135	0.077	0.108	0.139
x_7	0.055	0.317	0.071	-0.052	0.398	0.071	0.032	0.348	0.072
x_8	0.152	0.087	0.077	0.269	0.006	0.081	0.052	0.349	0.073
x_9	-0.041	0.288	0.149	-0.017	0.685	0.148	0.014	0.535	0.150

Table 6 *Results of multiple linear regression model developed using simple arithmetic operand method*

Variable	p value	Tolerance
	Four wheeled (R^2 = 0.702)	
Intercept	7.03E-21	N/A
Variable-1	0.000	0.349
Variable-2	1.74E-05	0.712
Variable-3	0.014	0.354
	Three wheeled (R^2 = 0.700)	
Intercept	4.39E-05	N/A
Variable-1	1.13E-06	0.623
Variable-2	0.013	0.505
Variable-3	0.016	0.400
	Two wheeled (R^2 = 0.707)	
Intercept	1.96E-18	N/A
Variable-1	0.009	0.258
Variable-2	0.006	0.923
Variable-3	0.014	0.248

to infer the statistical significance based on the t-stat (coefficient divided by standard error) and p-value for PCR models [31]. However, the tolerance must be more than 0.2 for all the variables to ensure that the multicollinearity is removed. However, results in Table 5 show that tolerance was less than 0.2 for majority of the variables and this clearly indicates that the PCR failed to remove the multicollinearity in the dataset. The present study tried a new method called "simple arithmetic operand" to remove the multicollinearity, which is explained below.

The proposed approach, named as "simple arithmetic operand" is quite simple and straightforward, as the name suggests. Because in this method, the independent variables were added, differenced or ratioed to create new variables, until the desired results were achieved, i.e. R^2 exceeding 0.70, p values of new IV's (independent variables) less than 0.05 accompanied by tolerance of more than 0.2. While developing the regression model using the proposed approach, 75% of the data were used so that the remaining 25% can be used for validation. The developed models are shown below in Equations (11)-(13).

a) Four wheeled vehicles

$$y_{FW} = 14.210 + 4.530\underbrace{[(x_1/x_2) - (x_3 + x_4)]}_{\text{Variable} - 1} - 0.061\underbrace{[(x_5 + x_6)]}_{\text{Variable} - 2} + 0.098\underbrace{[(x_7) + (x_8 - x_9)]}_{\text{Variable} - 3}. \quad (11)$$

b) Three wheeled vehicles

$$y_{ThW} = -8.722 + 0.452[(x_1 + x_2) + (x_3 + x_4)] + 0.041[(x_5 + x_6)] - 0.099[(x_7 - x_9) - (x_8)]. \quad (12)$$

(c) Two wheeled vehicles

$$y_{TW} = 10.425 + 0.157[(x_1 + x_2) - (x_3 + x_4)] + 0.020[(x_5 + x_6)] + 0.024[(x_7) + (x_8 + x_9)]. \quad (13)$$

It can be seen from Equations (11)-(13) that the regression model has been reduced from 9 variables to 3 variables, by adding, subtracting or ratioing the given independent variables. Sometimes, the combinations were applied, as well. For example, x_3 and x_4 were added and subtracted from the ratio of x_1 and x_2. As seen from Table 6, R^2 values were 0.7 and above, p-values were less than 0.05 and tolerance values were above 0.2. Tolerance of above 0.2 indicates that the proposed approach has removed the multicollinearity, which was not possible with the conventional methods like the PCR. The proposed approach is simple, does not require any specialized softwares like NCSS and produces the desired results finally. In addition, no independent variables were dropped as seen from Equations (11)-(13). Thus, with the proposed approach no variables are missing. The only limitation with this approach is that it involves the "trial and error" practice, that is done manually until the expected results are obtained. If a separate tool is developed in MSEXCEL to automate this trial-and-error procedure, then there will be no limitation for this method.

With the aim of validating the potential of the developed model in forecasting the mode wise vehicular speeds, the remaining 25% data were used for comparing the actual/observed speeds with the predicted ones. The results shown in Figure 5 clearly exhibit the similar trend of observed speed to that of the predicted speed. In addition, R^2 was greater than 0.7 for all the three modes, which says that the results were acceptable. The MAPE calculated between actual and predicted were 17.26%, 29.60% and 8.23% for the four wheeled, three wheeled and two wheeled vehicles respectively. Generally, the value of MAPE below 10% is considered to be excellent, 10-20% termed as good and a range of 20-50% is considered to be reasonable, but any value exceeding more than 50% is unacceptable [27- 28]. With the obtained results, one can draw a conclusion that the prediction accuracy is excellent in the case of the two wheeled vehicles, good for the four wheeled and reasonable for the three wheeled vehicles.

In Figure 5 was noticed that the predicted and observed values were in good agreement for the two wheeled and four wheeled vehicles, whereas for the three wheeled vehicles, the predicted speeds were lower than the actual ones. The possible reason for this could be that the three-wheelers include the normal autos and the shared type autos both of which have different characteristics in terms of its size, as well as operation.

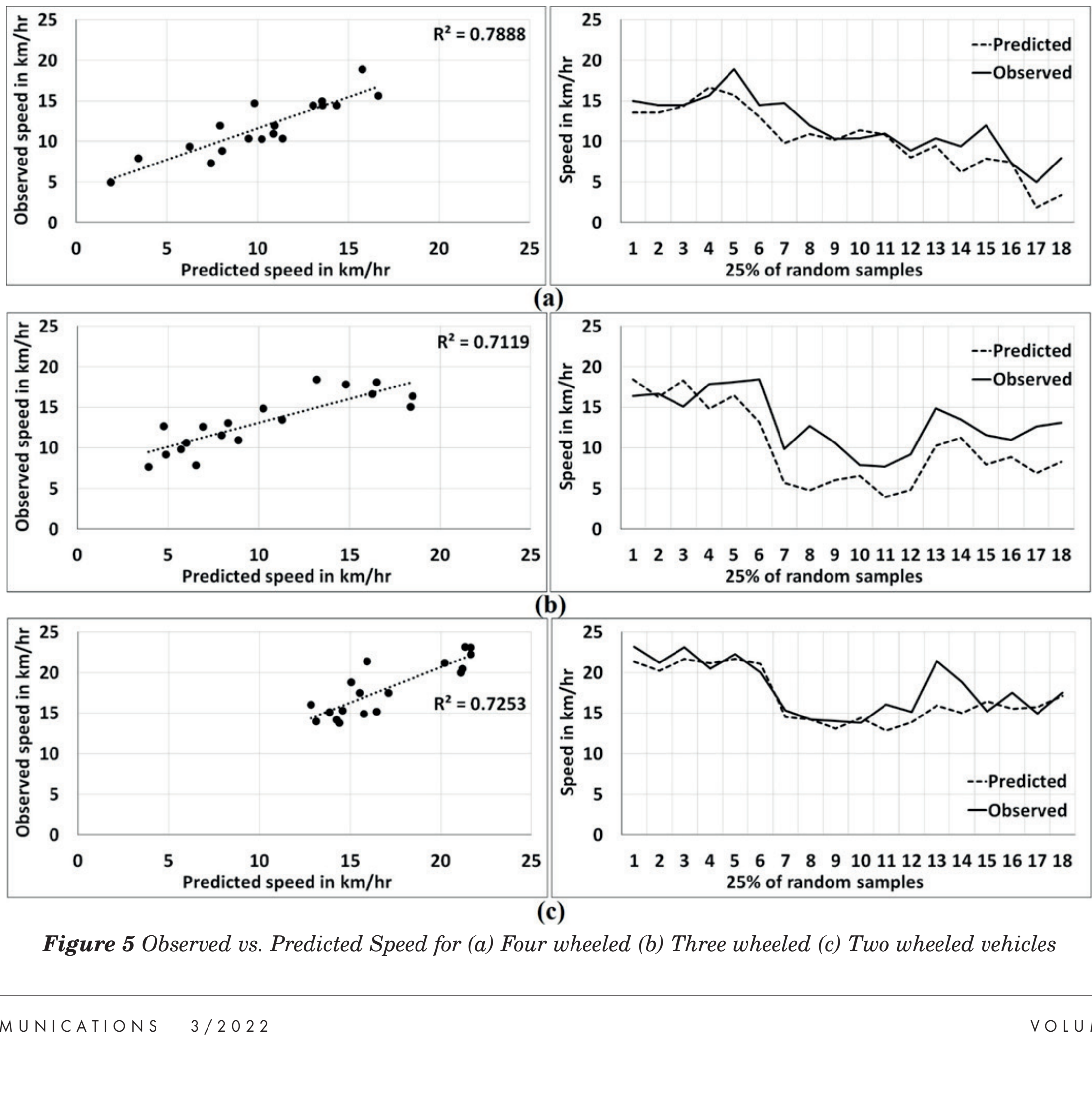

Figure 5 *Observed vs. Predicted Speed for (a) Four wheeled (b) Three wheeled (c) Two wheeled vehicles*

In addition, the normal autos were only present in Chikmagalur, whereas the shared type autos were counted more in Vellore. In model development, both types of autos have considered as one class and the data from all the locations were combined to develop the regression model. Perhaps it may likely to be the probable cause for three wheeled vehicles alone to exhibit slightly higher range of MAPE when compared to other modes. Overall, the MAPE and R^2 are within the acceptable limits for the all the three modes and, hence, the developed models could be considered for predicting the mode-wise speeds using speed corresponding to other vehicular classes, side friction activities data and traffic volume count.

6 Concluding remarks

One of the important factors that causes reduction in speed of vehicles on urban roads of India is the side friction. The carriageway, which is meant for vehicles to go has not been utilized properly for the purpose for which it has been constructed, because of these side frictions. Studies on investigating the impact of side frictions on vehicular speeds have been mostly focused on common type of roadside frictional activities like curb side bus stops, on-road parking, movement of pedestrians, etc. However, there are other side frictions like water stagnation, road-side shops and poor road condition, which are not studied before. Present study attempted this by collecting data from cities like Vellore and Chikmagalur in India and developed mode-wise speed prediction models using the regression analysis. One of the main highlights of the present study is that it has proposed a method to remove the multicollinearity, which is a common problem in the regression analysis. Even the popular methods, like the PCR, failed to remove multicollinearity in the dataset. However, the simple arithmetic operand method proposed in this study performs well, as it finally results in a model with only few variables, good R^2, low p-values and acceptable tolerance values. The mode-wise regression models, developed using the simple arithmetic operand method, were found to predict well the speeds of various categories of vehicles if the side friction and other causative variables are given as inputs. The arithmetic operand method that has been proposed in the present study can be applied in other fields, as well, to overcome the multicollinearity problem. The main limitation of the present study is that the traffic and side friction data were collected during the peak hours on selected working days. If one collects the data during the off-peak hours or weekends, the results may be slightly different as the impact of the side friction would be minimum during those times. Again, it depends on which side friction one considers, for example, poor road condition will not change on a weekday or weekend, whereas the on-street parking would be high during the weekdays when compared to weekends. Another limitation is that the present study considered cars, light commercial vehicles (LCVs) and buses under "four-wheelers" category, while extracting the flow data. If one considers them as separate classes, it may be possible to see how the speeds of those four-wheeled vehicle categories are influenced by different side frictions, which is the future scope of the present work.

References

[1] SOLTESZ, T., KOZEL, M., CSISZAR, C., CENTGRAF, T., BENYO, B. Information system for road infrastructure booking. *Periodica Polytechnic Transportation Engineering* [online]. 2011, **39**(2), p. 55-62. ISSN 0303-7800, eISSN 1587-3811. Available from: https://doi.org/10.3311/pp.tr.2011-2.02

[2] BALAKRISHNAN, S., SIVANANDAN, R. Influence of lane and vehicle subclass on free-flow speeds for urban roads in heterogeneous traffic. *Transportation Research Procedia* [online]. 2015, **10**, p. 166-175. ISSN 2352-1465. Available from: https://doi.org/10.1016/j.trpro.2015.09.066

[3] SILVA, D., FOLDES, D., CSISZAR, C. Autonomous vehicle use and urban space transformation: a scenario building and analysing method. *Sustainability* [online]. 2021, **13**(6), 3008. ISSN 2071-1050. Available from: https://doi.org/10.3390/su13063008

[4] ANIS, S., CSISZAR, C. Management of potential conflicts between pedestrians and autonomous vehicles. In: IEEE Smart City Symposium Prague SCSP 2019: proceedings [online]. 2019. ISBN 978-1-7281-0498-0, eISBN 978-1-7281-0497-3, p.1-6. Available from: https://doi.org/10.1109/SCSP.2019.8805678

[5] BISWAS, S., CHANDRA, S., GHOSH, I. Side friction parameters and their influences on capacity of Indian undivided urban streets. *International Journal of Transportation Science and Technology* [online]. 2021, **10**(1), p. 1-19. ISSN: 2046-0430. Available from: https://doi.org/10.1016/j.ijtst.2020.03.007

[6] ALKAISSI, Z., KAMOONA, W. Influence of roadside impedance on speed and capacity of urban streets. *IOP Conference Series: Materials Science and Engineering* [online]. 2021, **1090**, p. 1-13. ISSN 1757-8981, eISSN 1757-899X. Available from: https://doi:10.1088/1757-899X/1090/1/012121

[7] MAHENDRA, M., WICAKSONO, A., DJAKFAR, L. The effect of side friction on delays in one-way urban road sections. *Journal of Southwest Jiaotong University* [online]. 2021, **56**(5), p. 265- 274. ISSN 0258-2724. Available from: https://doi: 10.35741/issn.0258-2724.56.5.24

[8] PATKAR, M., DHAMANIYA, A. Influence of nonmotorized vehicles on speed characteristics and capacity of mixed motorized traffic of urban arterial midblock sections. *Journal of Transportation Engineering, Part A: Systems* [online]. 2020, **146**(4), p. 1-11. ISSN 2473-2907, eISSN 2473-2893. Available from: https://doi.org/10.1061/JTEPBS.0000325
[9] YEBOAH, A., AMOAH, A., KOVE, M. Analysis of congestion on some road link sections using roadside friction in congestion index in Kumasi. *International Journal for Traffic and Transport Engineering* [online]. 2020, **10**(1), p. 31-40. ISSN 2217-544X, eISSN 2217-5652. Available from: http://dx.doi.org/10.7708/ijtte.2020.10(1).03
[10] BITANGAZA, M., BWIRE, H. Effects of roadside friction elements on travel performance and level of service in Kigali. *International Journal for Traffic and Transport Engineering* [online]. 2020, **10**(2), p. 236-251. ISSN 2217-544X, eISSN 2217-5652. Available from: http://dx.doi.org/10.7708/ijtte.2020.10(2).09
[11] CHAUHAN, R., DHAMANIYA, A., ARKATKAR, S., SAHU, P. K., VIKRAM, D. Effect of side friction parameter on urban road traffic: under mixed traffic scenario. *Journal of the Eastern Asia Society for Transportation Studies* [online]. 2019, **13**, p. 314-330. eISSN 1881-1124. Available from: https://doi.org/10.11175/easts.13.314
[12] NAJID. Evaluation of side friction in IHCM for highway two lanes two ways. In: 2nd International Symposium on Transportation Studies in Developing Countries ISTSDC 2019: proceedings [online]. Vol. 193. Advances in Engineering Research. 2019. ISBN 978-94-6252-913-7, p. 116-119. Available from: https://doi.org/10.2991/aer.k.200220.024
[13] HIDAYATI, N., SUNARJONO, S., AWAD, S., MAGFIRONA, A. Different impact of side friction condition on traffic flow along Yosodipuro street Surakarta. *AIP Conference Proceedings* [online]. 2019, **2114**(1), 040011. ISSN 0094-243X, eISSN 1551-7616. Available from: https://doi.org/10.1063/1.5112440
[14] GULIVINDALA, P., MEHAR, A. Analysis of side friction on urban arterials. *Transport and Telecommunication* [online]. 2018, **19**(1), p. 21-30. eISSN 1407-6179. Available from: https://doi.org/10.2478/ttj-2018-0003
[15] SALINI, S., GEORGE, S., ASHALATHA, R. Effect of side frictions on traffic characteristics of urban arterials. *Transportation Research Procedia* [online]. 2016, **17**, p. 636-643. ISSN 2352-1465. Available from: https://doi.org/10.1016/j.trpro.2016.11.118
[16] RAO, A. M., RAO, K. R. Free speed modelling for urban arterials - a case study on Delhi. *Periodica Polytechnica Transportation Engineering* [online]. 2015, **43**(3), p. 111-119. ISSN 0303-7800, eISSN 1587-3811. Available from: https://doi.org/10.3311/PPtr.7599
[17] BANSAL, P., AGRAWAL, R., TIWARI, G. Impacts of bus-stops on the speed of motorized vehicles under heterogeneous traffic conditions: a case-study of Delhi, India. *International Journal of Transportation Science and Technology* [online]. 2014, **3**(2), p. 167-178. eISSN 2046-0430. Available from: https://doi.org/10.1260/2046-0430.3.2.167
[18] CHIGUMA, M. Analysis of side friction impacts on urban road link-case study, Dar-Es-Salaam. Doctoral thesis in traffic and transport planning. infrastructure and planning. Sweden: Royal Institute of Technology Stockholm, 2007. ISBN 978-91-85539-17-8.
[19] MUNAWAR, A. Speed and capacity for urban roads, Indonesian experience. *Procedia - Social and Behavioral Sciences* [online]. 2011, **16**, p. 382-387. eISSN 1877-0428. Available from: https://doi.org/10.1016/j.sbspro.2011.04.459
[20] MOORE, D., NOTZ, W., FLIGNER, M. *The basic practice of statistics*. 7. ed. New York, USA: W. H. Freeman, 2015. ISBN 978-1464142536.
[21] ROESS, R. P., PRASSAS, E. S., MCSHANE, W. R. *Traffic engineering*. 5. ed. London, UK: Pearson Education; 2019. ISBN 9780134599717.
[22] DEVORE, J. L. *Probability and statistics for engineering and the sciences*. 9. ed. Boston, USA: Brooks/Cole Cengage Learning, 2010. ISBN 978-0-538-73352-6.
[23] MONTGOMERY, D. C., RUNGER, G. C. *Applied statistics and probability for engineers*. 6. ed. New York, USA: John Wiley and Sons, Inc., 2016. ISBN 9788126562947.
[24] LEVINE, D. M., STEPHAN, D. F., SZABAT, K. A. *Statistics for managers using microsoft excel*. 8. ed. London, UK: Pearson Education Limited, 2016. ISBN 0134173058.
[25] KIM, J. H. Multicollinearity and misleading statistical results. *Korean Journal of Anaesthesiology* [online]. 2019, **72**(6), p. 558-569. ISSN 2005-6419, eISSN 2005-7563. Available from: https://doi.org/10.4097/kja.19087
[26] WEISBURD, D., BRITT, C. *Statistics in criminal justice*. New York, USA: Springer, 2014. ISBN 9781461491699.
[27] KUMAR, S. V., VANAJAKSHI, L. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review* [online]. 2015, **7**(21), p. 1-9. eISSN 1866-8887. Available from: https://doi.org/10.1007/s12544-015-0170-8
[28] BHARADWAJ, N., KUMAR, P., ARKATKAR, S., MAURYA, A., JOSHI, G. Traffic data analysis using image processing technique on Delhi-Gurgaon expressway. *Current Science* [online]. 2016, **110**(5), p. 808-822. ISSN 0011-3891. Available from: https://doi.org/10.18520/cs/v110/i5/808-822

[29] The Times of India. It's legal now: cars allowed to go up to 70 km/h in cities [online]. 2018. Available from: https://timesofindia.indiatimes.com/india/its-legal-now-cars-allowed-to-go-up-to-70km/h-in-cities/articleshow/63308093.cms
[30] Statistical Software - NCSS [online]. 2020. Available from: ncss.com/software/ncss
[31] Statistical software. Principal components regression. Chapter 340 - NCSS [online]. 2020. Available from: https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/ Principal_Components_Regression.pdf