Thailand Statistician
January 2025; 23(1): 162-180
http://statassoc.or.th
Contributed paper

Clustering Based Approach to Quantify the Unsafe Driving at Uncontrolled Median Openings due to Forced Deceleration: A Case Study in India

Malaya Mohanty*, Rachita Panda, Ritik Raj Arya and Sarthak Lenka
School of Civil Engineering, KIIT Deemed to be University Bhubaneswar, India.
*Corresponding author; e-mail: malaya.mohantyfce@kiit.ac.in

Received: 22 May 2022
Revised: 2 August 2022
Accepted: 30 September 2022

Abstract

Uncontrolled median openings with no lane discipline lead to complex driving phenomenon since the approaching through vehicles (major traffic stream) generally have to slow down their vehicles due to the presence of U-turns in the opposing traffic stream (minor traffic stream). It is observed that the approaching through vehicles despite being the major stream vehicles are forced to slow their vehicles, which sometimes also leads to evasive deceleration beyond limits of safety. In the present study, field data has been collected and the speed reduction from start of median opening to center of median opening is evaluated. Further, the effect of vehicle category and lane in which they are travelling, on speed reduction is also examined. Linear and quadratic probabilistic models have been developed to estimate the reduction in speed based on speed at the center of median opening, vehicle category, and zone/lane in which the vehicle is travelling. Safe reduction in speed is determined by using the IRC 66-1976 guidelines regarding SSD calculation. Thereafter, an undesirable driving index/safety index is proposed by conducting k-mean clustering utilizing the percentage reduction in speed. If the average reduction in speed increases beyond the proposed index, traffic facilities require deeper analysis for smooth flow of vehicles. The present study shall help in identifying the proportion of safe and unsafe reduction in speeds at median openings and proper steps can be undertaken to improve the safety and comfort of road users.

Keywords: Safety, speed reduction, probabilistic models, clustering analysis, driver behavior.

1. Introduction

A nation can't develop without a strong infrastructure, and roads play a major role in building a nation. Akin to building of roads, ensuring smooth flow of traffic is also very important. Further, in developing countries like India where heterogeneous traffic prevails, management of traffic flow is a big challenge. It is essential to assess various kinds of traffic facilities along with various categories of vehicles plying on it to quantify various traffic parameters. Uncontrolled median openings are one such kind of traffic facility which has not been much in the realm of research scope in developed nations (Mohanty and Dey 2021). This might be because of lesser presence of such facilities in developed countries along with the homogeneous traffic flow with proper lane discipline and design of reserve lanes for U-turning maneuver. However, in case of developing countries, the presence of uncontrolled median opening on urban roads affects the traffic flow in adverse manner (Mohapatra et al. 2016, Mohanty and Dey 2021, Sil et al. 2018). The presence of shared priority and sometimes even reversal of priority in case of median openings leads the approaching through vehicles (major stream vehicles) to face undesirable delay which in turn leads to traffic congestion during peak periods (Mohanty and Dey 2019-a). With increasing congestion and delays, it is difficult to maintain a smooth flow of traffic at median openings.

Studies pertaining to median openings have been majorly biased towards the study of U-turns. While few researchers (Al-Masaeid 1999, Mohapatra et al. 2016) have tried to estimate the capacity of U-turning traffic, many researchers (Mazaheri et al. 2020, Gupta et al. 2018) have tried to calculate the critical gaps and study the merging behaviour of U-turns. Even being the minor stream traffic, a lot of research have been conducted to access the movement of U-turns. Yao et al. (2019) provided a structured guideline as where a median opening should be provided albeit considering the movement of only U-turns and signal timings of the minor street vehicles. Similarly, Liu et al. (2008) determined the capacity of median openings based on the critical gap acceptance by U-turning vehicles. However, it has been observed by few researchers (Mohapatra et al. 2016, Mohanty and Dey 2019a) that the approaching through vehicles are forced to slow down and face undesirable delay even though they are the major stream traffic. Mohanty and Dey (2019a) have calculated the delay faced by the approaching through vehicles as they approach the median openings and within the median opening area. According to the authors the delay faced by the vehicles is due to presence of U-turns. However, if observed thoroughly, it can be noted that even during the presence of very less U-turns, sometimes, the approaching through vehicles also face some delay. The reason behind this delay might be attributed to the geometry of median openings coupled with psychology of drivers. The dangerous reduction in speed and its effects have been studied by various researchers (Maze et al. 2000, Antov et al. 2009, Aarts and Schagen 2006, Afukaar and Damsere-Derry 2010, Mohanty et al. 2021). However, most of the studies deal with either the speed reduction at work zones or at speed humps and uncontrolled intersections. The studies related to speed reduction at median openings haven't been dealt more. The reduction in speed at median openings beyond a certain threshold, might lead to evasive braking and unsafe movement at median openings. It may impact the safety and life of vehicle. Therefore, this undesirable slowing down of vehicles must be evaluated and proper measures should be ensured if majority of vehicles are forced to reduce their speed beyond a certain threshold. Further, the major reasons behind the slowing down of vehicles except the presence of U-turns must also be investigated.

In the present study field data is collected and the speed reduction of vehicles from start of median opening to the center of median opening is determined. The effect of vehicle category and lane in which its travelling is also investigated with respect to reduction in speed. Thereafter, based on IRC 66-1976, stopping sight distance (SSD) is utilized to determine the safe percentage of speed reduction at median opening areas. Cluster analysis is conducted, and a safety index/undesirability index is proposed. Generalized linear and quadratic probabilistic models have been developed to estimate the percentage reduction in speed at median openings based on contributing parameters.

2. Site Selection and Data Collection

The present study requires the speed data from median openings. The vehicular speeds need to be extracted for all vehicles at 2 points of the road i.e., at the start of median opening and at the center of median opening area. To obtain these speeds, instantaneous speeds are measured at these 2 locations by considering 10 meters of road stretch each. The reduction in speed within the two points shall be analyzed to determine whether the reduction is significant. Therefore, the first step in data collection is to utilize two cameras to record the video, (i) considering 10 m at the start of median opening, and (ii) considering 10 m within the median opening area. The test sections were selected in such a way that the traffic flow at these sites is not influenced by horizontal/vertical curvature, presence of intersection in downstream/upstream direction, bus stop, parked vehicles, pedestrian movements, or any kind of side friction. All the road sections were 6-lane divided i.e., 3-lanes on each direction of travel with raised curb on either side of the road. The widths of median openings are observed to be 20 meters. Therefore, the distance between start of median opening to the center of median opening is 10 meters. Data were collected by video recording technique on typical weekdays. The camera set up is shown in Figure 1. The speed measured in between AB to CD (10 meters) is assumed to be the average speed at the start of median opening (MO). Similarly, the speed measured in between CD to EF (10 meters) is assumed to be the average speed at the center of MO.

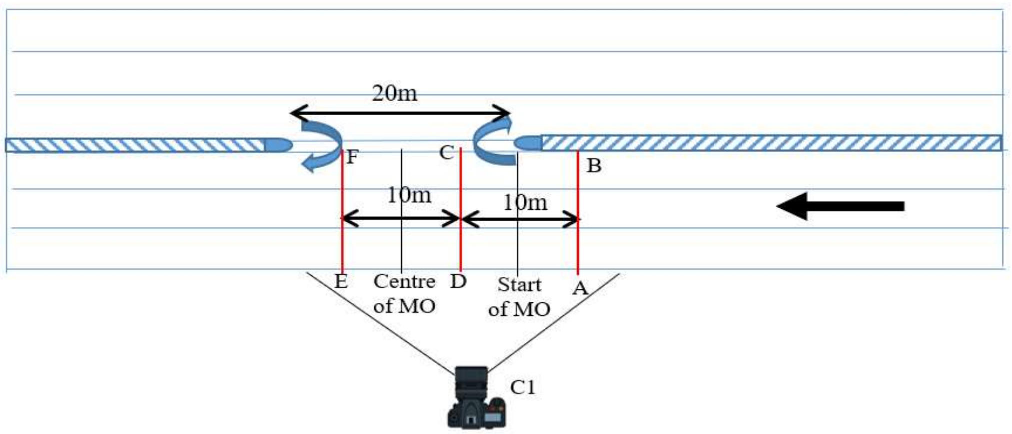

Figure 1 Camera set-up for data collection

The recorded films were played on desktop monitors and laptops with the help of software (Avidemux version 2.6) which plays the videos at frame per second at which it was recorded. For the present study, the camera used had a precision level of 25 to 30 frames per second (fps). From the video, the following information was noted:

- The time intervals at which a vehicle crossed AB, CD, and EF (noted for all vehicles individually).
- Category of every vehicle whose times are noted down.
- Lane in which vehicle travels (lane 1/median lane, lane 2/central lane, lane 3/kerb lane, lane 1&2 sharing, lane 2&3 sharing)

The following formulae are used to determine speed at the start and center of median opening

$$V_{reduced} = 0.89 - (1.47 \times SCMO) + (5.40 \times 2W/3W) + (7.83 \times Cars) + (6.23 \times Z1) + (2.80 \times Z2) \quad (1)$$

$$V_{CentreMO} = \left(\frac{10}{t_{EF} - t_{CD}} \right), \quad (2)$$

where $V_{StratMO}$ = Average speed at start of MO, $V_{CentreMO}$ = Average speed at center of MO and T_{ij} = time taken to cross ij (where ij refers to AB, CD, EF).

3. Results and Discussions

Classified traffic data has been extracted from the recorded video along with operating speeds at 2 successive road stretches to study the undesirable speed reduction at uncontrolled median openings. The speed data of 5600 vehicles across all traffic volumes have been utilized for the present study. The average speed reduction for each vehicle is noted down. It is also observed that the reduction in speed varies with category of vehicle and the lane in which the vehicles are travelling. Further, sometimes reduction in speed seemed to be very evasive and sudden whose threshold value needs to be identified too. To understand the phenomenon of every factor on undesirable speed reduction at the median openings, the results and analysis section is divided into 5 parts as follows:

1) Assessment of reduction in speed at median openings, 2) Effect of vehicle category on speed reduction, 3) Effect of zone on speed reduction, 4) Analysis of safe and unsafe deceleration to develop a safety/undesirability index and 5) Development of probabilistic models to estimate reduction in speed.

3.1. Assessment of reduction in speed at median openings

Firstly, the descriptive statistics for all the three variables i.e., the speed at the start of median opening, speed within median opening, and percentage reduction in speeds from start to center of the median opening have been calculated using IBM SPSS to better interpret the data. Table 1 shows the mean, standard deviation, and skewness of the data.

Table 1 Descriptive statistics for the input variables

Statistical parameters	Speed at the start of MO	Speed within MO	Percentage reduction of speed
Mean	40.81	30.58	25.33
Std. Dev	6.39	8.09	15.01
Skewness	0.301	–0.183	0.510

Table 1 shows a positive skewness of data as compared to normal distribution for speeds of vehicles at start of MO. However, a negative skewness is observed in case of speeds within MO. This variation in structure of data is observed due to sudden and undesirable reduction in speed of vehicles while entering median openings. This variation in speed distribution at successive stretches leads to high skewness in the distribution of percentage reduction in speed. Moreover, while the standard deviation (SD) to mean ratio is around 0.16 to 0.26 (i.e., SD is 16-26% of mean) for speed values at start and center of MO, the SD to mean increases to almost 0.6 for percentage reduction in speed. The increase in SD reflects the greater spread of speed reduction along with more data of speed reduction on the higher side (positively skewed). If assumed to be normally distributed, 67% of the speed reduction values lie between 10% speed reduction to as high as 40% of speed reduction. The field data shows reduction in speed in the range of even 70%. The values of speed reduction and skewness clearly depict that the vehicles are adversely affected by the presence of median opening and U-turning vehicles which lead them to decrease their speed non-uniformly. Majority of the vehicles are found to have reduced their speeds at a percentage of 10 to 30%. Although the speed reduction is clearly visible, the significant statistical difference between the speeds at the start of median opening and at the center of median opening area, is tested by conducting a t-test as shown in Table 2. The sig. value of < 0.05 confirms that both the speeds are significantly different from each other.

Table 2 t-test between speeds up to at start of MO and within the MO

t-test	Mean	Std. Dev	Std. Error Mean	t	Sig.
Speed at the start of median opening and speed within the median opening	10.23	5.96	0.163	62.81	0.00

The above analysis clearly reveals that approaching through vehicles, although being the major stream vehicles face undesirable obstacles in form of U-turning traffic which sometimes is so severe that they have to reduce their speeds even by 60-70% in a stretch of 10 meters only, while entering the median opening area. According to Perco et al. (2012), speed reduction of less than 2.5 kmph can be considered practically same. Therefore, the cases where reduction in speed is less than 2.5 kmph (45 data points out of 1337) haven't been considered for the study. Figure 2 depicts the individual value plot of all the values of percentage reduction in speed.

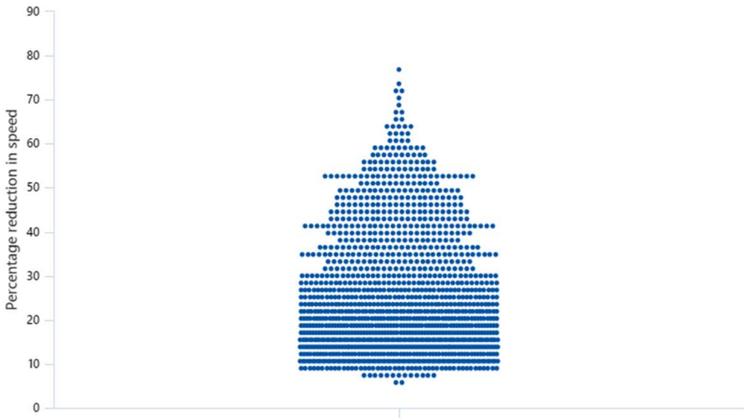

Figure 2 Individual value plot for percentage in speed reduction

Figure 2 clearly indicates that although maximum vehicles have reduced their speeds in the range of 10-30%, still many vehicles also reduce their speeds up to 40-45%. There are also few vehicles who reduce their speed in the order of 60-80%. In order to examine this reduction in speed at a deeper level, it needs to be known whether speed reduction is different for different category of vehicle and whether travelling on any specific lane also affects the reduction in speed. Further, Figure 3 presents the histogram of frequencies of all the speed data which represents a near-normal distribution with marginal positive skewness.

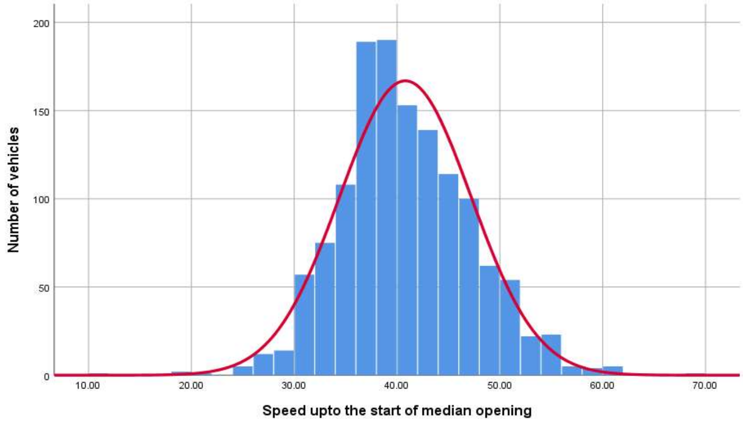

Figure 3 Histogram depicting the frequencies of speed data at median opening

3.2. Effect of vehicle category on speed reduction

In the present study, 4 major categories of vehicles were identified which are HV (Heavy vehicles), Car, 3W & 2W. Car includes cars, jeeps, and SUVs. Thereafter, category wise speed reduction was observed & extracted from video for analysis. First, a general descriptive statistical analysis of the extracted data was done using IBM SPSS and is provided in Table 3. The table shows that the HVs (Buses and trucks) exhibit the lowest reduction in speed, whereas cars exhibit the highest at around 28% mean reduction in speed. The speed reduction in case of 2W and 3W exhibit similar values. The data points can be observed graphically in Figure 4 with the help of box plots. Statistical significance among the speeds of various category of vehicles can be determined by conducting t-test between the speeds of different category of vehicles. T-test compares the mean between 2 samples and declares whether the 2 data sets are statistically similar or dissimilar at any specific level of significance. Therefore, to test the statistical significance of difference/indifference between the reduction in speed values for various categories, 2-sample t-tests have been conducted at 5% significance level whose results have been showcased in Table 4.

Table 3 Descriptive statistics of speed reductions by various category of vehicles

Variable	Mean	Std. Dev	Minimum	Maximum
Speed Reduction (2W)	25.5	12.7	7.0	63.1
Speed Reduction (3W)	25.3	14.0	7.2	70.0
Speed Reduction (Car)	28.7	16.4	7.2	76.1
Speed Reduction (HV)	15.5	03.3	11.7	17.6

We can conclude from the results of the t-test that the speed reduction between every category of vehicles, except between 2W & 3W is statistically different. The reason for this might be because of the manoeuvrability freedom of 2W (bikes, scooties, etc) & 3W (autos, minivans, etc) which is much more compared to Cars & HV. Therefore, 2W and 3W tend to exhibit more lane changing behavior than other categories leading to lesser reduction in speed. They change their lanes quickly to avoid reduction in speed near the median opening area. Similarly, the speed reduction between cars & HV is also statistically different since heavy vehicles generally travel at lower speeds as compared to cars due to their heavy weights, and mostly they stick to lane 2 or lane 3 (middle lane or kerb lane) to avoid unnecessary congestion near lane 1 at median opening areas. Cars being light weight and more dynamic as compared to HVs, require less braking distance to come to stop and hence exhibit higher reduction in speed. Secondly, from safety aspect, heavy vehicles are known to be more unsafe when involved in crashes as compared to cars (Brodie et al. 2009, Assemi and Hickman 2018, Anderson and Hernandez 2017). Therefore, HVs usually travel at lower speeds as compared to cars. The study also gives a viewpoint that while developing a predictive model to estimate percentage reduction in speed, 2W and 3W can be considered as a single category.

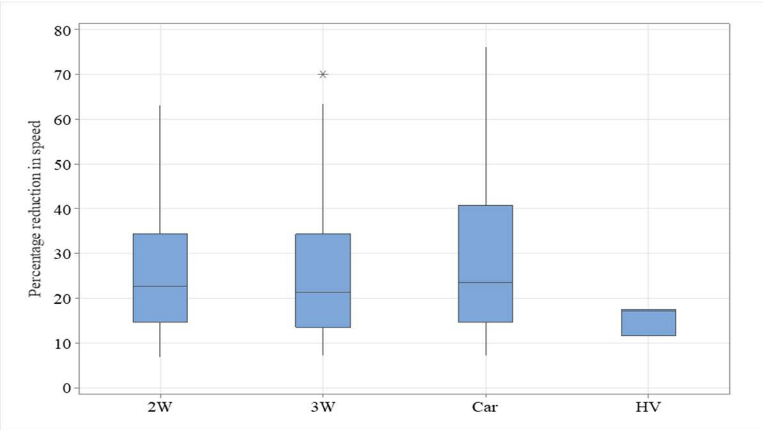

Figure 4 Box plots of percentage in speed reduction based on vehicle category

Table 4 t-test for comparing reduction in speed across various category of vehicles

Category of vehicle for conducting t-test	t-value	p-value	Remarks
2W and 3W	0.21	0.835 > 0.05	No statistical difference
2W and Car	–3.29	0.001 < 0.05	Statistically different
2W and HV	5.05	0.037 < 0.05	Statistically different
3W and Car	3.07	0.002 < 0.05	Statistically different
3W and HV	4.08	0.041 < 0.05	Statistically different
Car and HV	6.38	0.024 < 0.05	Statistically different

3.3. Effect of zone on speed reduction

In a developing country like India, heterogeneous traffic conditions prevail with no proper lane discipline (Mathews et al. 2015, Mohanty and Dey 2021). According to Mohapatra et al. (2016), on a 6-lane divided road, where each direction of movement is 3-lane, usually 5 queues of vehicles are observed. In other words, many vehicles are seen to be sharing 2 adjacent lanes simultaneously while driving. Further, while approaching towards median opening area, the vehicles are observed to be changing their lanes from median lane (lane 1) towards central lane (lane 2) or even kerb lane (lane 3). The visualization of this lane changing behavior is presented in Figure 5. To address this issue of heterogeneity and absence of lane discipline, the zone concept was introduced by Mohanty and Dey (2019b). Zone 1 is defined as the vehicles who enter the study area at median lane (lane 1) but exit the study area at lane 1 or sharing lane 1&2 or central lane (lane 2). Similarly, zone 2 and zone 3 are also defined based on entry and exit of vehicles on specific lanes as exhibited in Table 5. The descriptive statistics showcasing the average speed reduction for vehicles travelling in each zone is presented in Table 6.

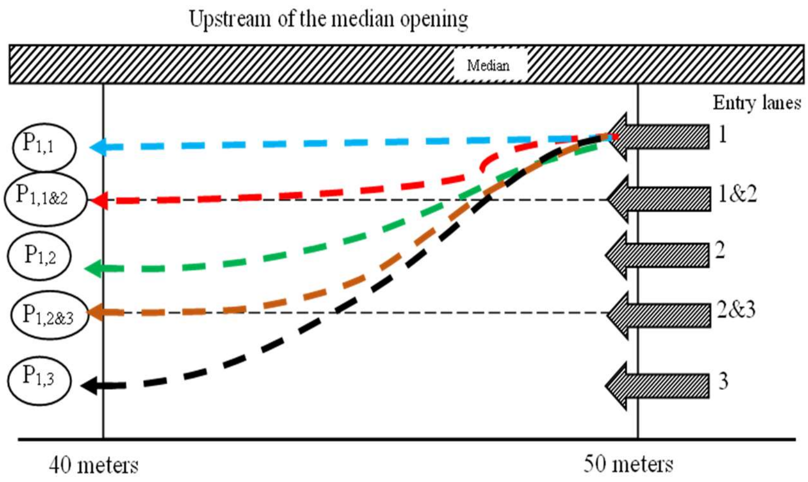

Figure 5 Visualization of lane changing behavior at median openings

Table 5 Classification of zones

Zone	Entry lane	Exit lane
1	1, 1&2	1, 1&2, 2
2	2, 2&3	2, 2&3, 3
3	3	3

Table 6 Descriptive statistics of speed reductions at different zones

Variable	Mean speed reduction (%)	Std. Dev	Minimum	Maximum
Zone 1	28.6	15.7	7.2	72.0
Zone 2	25.4	13.5	7.0	76.1
Zone 3	22.4	11.1	7.2	45.6

Table 6 clearly indicates a decrease in mean speed reduction while driving away from the median towards kerb. In other words, vehicles travelling in zone 3 are observed to be reducing their speeds in lesser quantity as compared to vehicles who are travelling in zone 2, and zone 1 respectively. Mainly 2 reasons contribute to this peculiarity in results. Firstly, the U-turning vehicles usually encroach the first lane and to some extent the second lane in the major stream traffic, while making a U-turning maneuver. According to an independent study by Mohapatra et al. (2016), 2W, 3W, cars and SUVs usually make a U-turn maneuver by encroaching maximum upto 6-7 meters (upto second/central lane) on the major traffic stream. Due to this phenomenon, vehicles who travel in these zones, face higher reduction in speed. Secondly, vehicles who generally have lower desirable speeds prefer to move on kerb lane. Thus, vehicles whose speeds are already less won't experience higher reduction in speed. It can also be observed that even though the average reduction in speeds is in the range of 22-28%, however, the maximum reduction in speed even reaches 72-76% on zone 1 and 2. The statistical significance of this decrement is tested via a t-test which is presented in Table 7. As can be seen from Table 7, all the sig. values or p-values lie below 0.05 indicating the mean speed reduction to be different for different zones statistically. Box plots are also drawn (Figure 6) which exhibit this decrement of speeds across the 3 zones in a better way.

Table 7 t-test for comparing reduction in speed across various zones

Zone for conducting t-test	t-value	p-value	Remarks
Zone 1 & 2	3.61	0.000	Statistically different
Zone 2 & 3	2.11	0.037	Statistically different
Zone 1 & 3	4.06	0.000	Statistically different

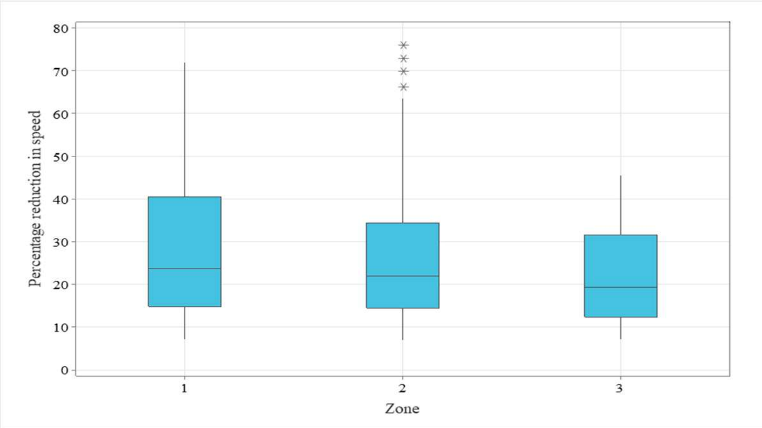

Figure 6 Box plots of percentage in speed reduction based on zone

3.4. Analysis of safe and unsafe deceleration to develop a safety/undesirability index

To acquire safe and efficient operation of vehicles, the design of roadway must permit the drivers sufficient time and distance to control their vehicles in different situation. As per Indian Code IRC: 66 -1976, the minimum sight distance needed by a driver to stop his vehicle so as to avoid unforeseen accidents is known as stopping sight distance (SSD). SSD has two prime components from which one is the distance travelled to perceive a situation and apply brakes (also known as lag distance) and the second one is the distance travelled after application of brakes till the vehicle stops. (braking distance). The following formula is used to determine SSD for a road (IRC 66-1976)

$$SSD = \left(0.278 \times V \times t\right) + \frac{V^2}{254 f}, \tag{3}$$

where SSD is the stopping sight distance in m, V is the initial speed of the vehicle in kmph, t is the reaction time of the driver in sec, and f is the coefficient of longitudinal friction (0.35-0.4).

Based on the data regarding initial speed and SSD in the IRC code, along with the study by early researchers (Mohanty et al. 2021-b), the critical deceleration rate for vehicles is obtained to be 3.81 m/s^2. To calculate the safe critical deceleration rate, an assumption is made, regarding the lag distance. It is assumed that since a vehicle is driving towards a median opening, the driver already has a perception to stop/slow down the vehicle. Therefore, the reaction time of driver leading to lag distance is not considered. Only the braking distance is considered as present in IRC 66. To analyze the critical safe reduction rate/deceleration rate of speed the following equation of motion was used:

$$v^2 - u^2 = 2 \times a \times s, \tag{4}$$

where v is the final speed in m/s, u is the initial speed in m/s, s is the stopping distance in m, and a is the deceleration rate in m/s^2.

From Table 1, the mean speed at the start of median opening is calculated to be 40.81 kmph. Applying the values of 10 meters (s) of road stretch along with v as 0 m/s and u as 40.81 kmph (which is equivalent to 11.34 m/s) in (2), the critical deceleration rate is computed as 3.81 m/s2 which was earlier reported by Mohanty et al. (2021-b). Utilizing the values of critical deceleration rate, initial mean speed, and distance, the percentage reduction in speed is calculated. It is observed that the critical safe reduction in speed from start to center of MO should be around 36.2%. Thus, reduction in speed more than 36.2% is dangerous and showcases evasive behavior. Similarly, speed reduction below the critical value can be considered as safe. Table 8 shows the preliminary segmentation of data based on the usage of IRC guidelines.

Table 8 Preliminary severity indices based on usage of IRC guidelines

Speed Reduction (%)	Severity Index (SI)	Traffic Safety Characteristic at Median Opening
< 36.2	0	Safe
36.2	Critical Safe	Critically Safe
>36.2	1	Unsafe

Table 8 suggests that if the speed reduction is below 36.2% between start to center of median opening (based on an average speed of 40.81 kmph), the vehicles are safe. Similarly, speed reduction of greater than 36.2% indicates Insafe driving condition. However, if keenly observed, 36.2% is a critical value; therefore, speed reduction of 35% and 10% represent 2 completely different scenarios. A speed reduction of 35% is more unsafe as compared to speed reduction of 10% since the latter involves general braking operation as compared to the former which involves forced braking operation, which is not beyond limits of IRC codes, but is evasive. Ofcourse, speed reduction values of higher than 36.2% should be considered with same impression since once above the critical value, any percentage reduction is beyond the limits of evasive braking and has every chance of leading to rear-end collisions with the same stream traffic or side collision with U-turns. Percentage of speed reduction values less than 36.2% need more analysis to define the safety classifications. In the present study, clustering technique has been utilised to classify the speed reduction percentage values which are below 36.2%.

Cluster analysis is a common mathematical tool to classify the data into various ranges/groups. The objective of this analysis is to perform a partition where objects within a cluster form a group of similar structure and it is widely utilized in various fields of science and technology (Jain and Dubes 1988, Sangngam et al. 2022). The data points in a cluster group are closer to the center of that same group, as compared to the center of other cluster groups. An efficient clustering method will produce clusters with property that their intra-cluster distance is small, and their inter-cluster distance is large (Jain and Dubes 1988). The commonly used clustering algorithms those can be used for defining LOS criteria are K-means, K-medoid, and hierarchical agglomerative. In this study, K-means has been used for the grouping of data points and determine the LOS criteria. SPSS software has been used for the cluster analysis and its validation. The commonly used clustering techniques in the field of traffic and transportation engineering are k-means, and hierarchical clustering (Mohapatra et al. 2015). K-mean clustering is one of the unsupervised methods for solving the classification problems where the groups are classified based on their Euclidean distances (Mohapatra et al. 2015). K-mean method uses the variation within cluster as a quantity to form homogenous clusters. Precisely, the technique aims at segmenting the data in such a way that the variation within a cluster is minimized. Here, the method of clustering starts by randomly assigning objects to several clusters. The objects are then successively reassigned to other clusters to minimize the within-cluster variation, which is basically the (squared) distance from each observation to the center of the associated cluster. If the reallocation of an object to another cluster decreases the within-cluster variation, this object is reassigned to that cluster (Mooi and Sarstedt 2014). According to Mooi and Sarstedt (2014), when the data set exceeds 500, K-means clustering process provides better partitioning results. Many researchers (Pecheux et al. 2000, Mohapatra et al. 2015, Patnaik et al. 2016, Biswas et al. 2017, Mohanty and Dey 2019b, Mohanty et al. 2021) have successfully utilized this technique for studying various aspects of traffic flow modelling like safety analysis, LOS determination, etc. According to Mooi and Sarstedt (2014), the number of clusters in a K-mean clustering must be pre-specified by researcher or the number of optimum clusters can be obtained from Hierarchical clustering and can be specified here. But generally, K-means is superior to hierarchical methods as it is less affected by outliers and the presence of irrelevant clustering variables. Furthermore, K-means can be applied to very large datasets, as the procedure is less computationally demanding than hierarchical methods (Mooi and Sarstedt, 2014). They concluded that silhouette could be used to define the number of clusters (K) for clustering analysis. A clustering is usually continued until a predetermined number of iterations are reached, or convergence is reached (Mooi and Sarstedt, 2014). Convergence is an important key to the K-means clustering technique. Convergence of the clusters simply means that there is no more change in the cluster affiliations. This convergence is achieved by the help of a number of iterations. At first SPSS finds out the value of center of clusters based on the number of predefined clusters. Usually, Lloyd's algorithm is used in K-means clustering to reach at convergence while finding out the center of clusters. A popular heuristic for K-means clustering is Lloyd's algorithm (Jain, 2010). According to Yadav and Sharma (2013), Lloyd's algorithm can be explained very simply in two phases. In the first phase, K number of centroids are selected by randomly where K refers to the number of defined clusters. In the second phase, each object or data point in the set is associated with the nearest centroid. Euclidean distance is used to measure the distance between each data object and cluster centroid. If the distance of one object is nearer to another cluster center than the random defined center, the center changes until all the data objects of a particular cluster are near to the center of that cluster only. The algorithm used in IBM SPSS for convergence can also be written simply as follows:

Input

K: number of desired clusters

$D:\{d1, d2, ..., dn)$ a data set containing n objects

Output

A set of K clusters as specified in input

Method

1) K data items are arbitrarily chosen from D dataset as initial cluster centers.

2) Each data item d_i is assigned to the cluster to which object is most similar based on the mean value of the object in cluster.

3) New mean value of the data items for each cluster is calculated and the mean value is updated.

4) The process is repeated until no change in mean value is observed in the next iteration.

In the present study, at first, a 2-step clustering is conducted to check the optimal number of clusters based on average silhouette value. Three number of clusters were obtained from preliminary screening to be the most optimal number of clusters for the data. Figure 7 shows the graphical representation of the average silhouette value for 3 number of clusters.

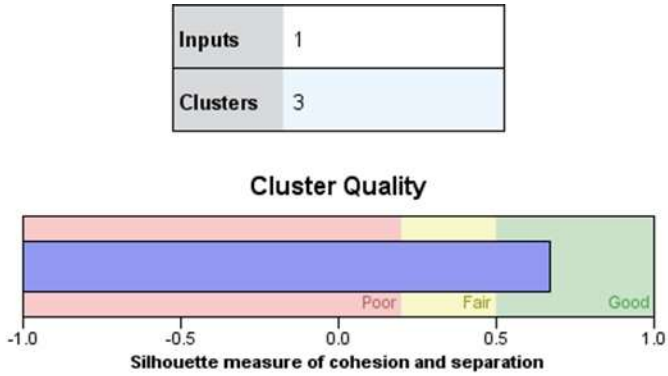

Figure 7 Average silhouette value for 3 clusters

Next, K-mean clustering is conducted with output set at 3 number of clusters. All the speed reduction percentage values below 36.2% were considered for the clustering. Since the number of data points are very high (965) which is greater than 500, K-means clustering technique can be considered as one of the best techniques (Boora et al. 2017). It was observed that for the K-mean clustering technique, the cluster centers converged after 15 iterations as shown in Table 9. The final cluster centers/means are provided in Table 10.

The viability of the cluster analysis is checked by conducting a variable clustering of the K-means clustering technique. The R square value and total proportion of variation explained for the cluster structure is obtained to be 0.994. Figure 8 presents the cluster ranges which is tabulated in Table 11. Table 11 gives the final ranges of speed reduction percentages for evaluating the safety aspect and undesirable driving index within median openings. The remarks column demarcates the traffic situation at every safety index/undesirability level (A to D). While index A (when percentage reduction speed < 15.53%) represents the safest movement in terms of applying brakes, reducing speed, and possibility of rear end collision along with most desirable driving by drivers at median openings, index D represents traffic flow failure along with acute forced driving within the MO. Median openings should be optimally designed so that even during high traffic flow, the speeds of vehicles should not reduce by more than 24-25% from start to center of MO and maintain the safety index level B.

Table 9 Iteration history of K-means clustering technique

Iteration	Change in Cluster Centers		
	1	2	3
1	0.230	6.549	7.821
2	2.203	1.106	0.415
3	0.773	0.784	0.215
4	0.590	0.743	0.310
5	0.422	0.689	0.358
6	0.454	0.781	0.326
7	0.181	0.455	0.313
8	0.167	0.354	0.196
9	0.069	0.176	0.117

Table 9 (Continued)

Iteration	Change in Cluster Centers		
	1	2	3
10	0.060	0.164	0.118
11	0.041	0.050	0.000
12	0.010	0.042	0.039
13	0.000	0.015	0.019
14	0.000	0.015	0.020
15	0.000	0.015	0.020
16	0.000	0.000	0.000

Table 10 Final cluster centers/ means along with number of cases

Cluster	Number of Cases	Percentage Reduction in Speed
1	183.50	11.93
2	532.00	19.67
3	831.50	29.57

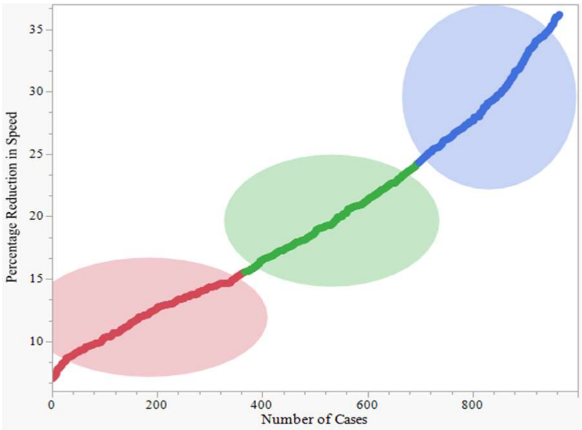

Figure 8 Scatterplot matrix for cluster groups

3.5. Development of probabilistic model to estimate reduction in speed

Field studies have revealed that the reduction of speed within MO is influenced by speeds at start and center of MO, albeit it depends more on the speeds at the center of MO (as depicted in Table 12). Moreover, it is not always the case that 2 vehicles driving at a definite speed will reduce their speeds by similar percentage. Considering the field conditions, it is easier to collect and extract speed data of one road stretch rather than 2. A Pearson correlation was conducted among percentage reduction in speed with speed at start of MO and speed at the center of MO which is shown in Table 12.

Table 11 Safety indices in terms of speed reduction percentages within median openings

Safety Index/Undesirability index	Range of percentage reduction in speed	Remarks
A	<15.53	Safe manoeuvre within MO. Speed reduction is not undesirable. Speed reduction mostly due to road geometry and psychology of drivers.
B	15.54–24.30	Stable movement within MO. Marginally unsafe situation. Speed reduction starts to get affected by presence of U-turns and heavy approaching through traffic stream.
C	24.30–36.20	Critical movement within MO Critically unsafe situation for drivers increases. Speed reduction is undesirable and is caused due to presence of U-turns. Forced braking operation starts
D	>36.20	Completely unsafe situation. Movement within MO is highly unsafe and congested. Sudden evasive braking may lead to rear end collisions. Undesirable delay and bottleneck situation.

Table 12 Pearson correlation between speed at start, center of MO, and percentage reduction in speed

	Speed at start of MO	Speed at center of MO	Per. Reduction in speed
Speed at start of MO	1.000	0.710	–0.198
Speed at center of MO	0.710	1.000	–0.825
Per. Reduction in speed	–0.198	–0.825	1.000

It can be clearly observed that speed at the center of MO has a large impact on percentage reduction in speed (higher absolute correlation value) as compared to speed at the start of MO. Therefore, to predict the percentage reduction in speed for a traffic stream, speed at the center of the median opening area has been considered. The amount of speed reduced by a driver suggests about his/her uncomfortableness while driving through the median opening and indicates about the forced braking operation conducted by the driver. The average reduction in speed for a traffic stream can be used to understand the level of unsafe manoeuvres at MO. Various mathematical relations (linear, logarithmic, quadratic, exponential) are developed to estimate the percentage reduction in speed for the traffic stream, based on speed at center of MO as independent variable. The R square values for all the models were checked along with p-value/significance value. The details of the statistics pertaining to various curve estimations are provided in Table 13. It is observed that the p-value in case of all curve fittings have come less than 0.05. Therefore, the model with highest R square value has been used for determining the percentage reduction in speed. In the present study, quadratic model has been found to estimate the percentage reduction in speed, with higher accuracy at the center of median opening area with an R square of 0.765 as shown in Table 13.

Table 13 Results of different curve estimates to determine percentage reduction in speed

Equation	Model Summary		Parameter Estimates		
	R Square	Sig.	Constant	b1	b2
Linear	0.680	0.000	70.52	−1.45	
Logarithmic	0.735	0.000	167.42	−41.81	
Quadratic	**0.765**	**0.000**	**119.36**	**−5.00**	**0.06**
Exponential	0.637	0.000	117.12	−0.05	

The developed mathematical equation to determine the percentage reduction in speed for the whole traffic stream considering the speed at center of median opening area is provided in (5).

$$PRS = 119.36 - (5 \times SWMO) + (0.06 \times SWMO^2), \quad (5)$$

where *PRS* is the percentage reduction in speed and *SWMO* is the speed within the median opening area in kmph.

The equation works best for speeds ranging from 9 to 50 kmph within the median opening area. To validate the equation, the difference between field data and model result is compared for the data that has not been used for model development. The mean absolute percentage error (MAPE) has been calculated for the data. The formula used to calculate MAPE is presented in (6).

$$M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|, \quad (6)$$

where A_t is the actual value, F_t is the model value and n is the number of data used for validation.

The highest mean absolute percentage error (MAPE) for the developed equation (Equation 5) were obtained to be around 8% which is acceptable since MAPE value less than or equal to 10% is considered strong (Liu et al., 2008). Therefore, by utilising speeds at the center of MO, the rate of reduction in speed within MO can be determined using developed equation (5) with good levels of accuracy.

Field investigations have also shown that the presence of multiple vehicle category (2W, 3W, Cars and HVs) and the zone in which the vehicles are travelling also have significant influence over the speed reduction within MO. Further, Equation (5) will not be sufficient if speed reduction of a particular category travelling in any specific zone is required. Thus, another multilinear probabilistic regression model has been developed to predict the percentage reduction in speed based on speed at center of MO, along with category of vehicle and zone in which it is travelling. The equation has a better MAPE value as compared to previously developed model and can be used to find the average reduction in speed of the traffic stream if vehicle composition is known.

The category of vehicle and zone, both are indicator/categorical variables. These variables designate the absence or presence of an attribute. Such factors are generally quantified by creating artificial dummy variables that take a value of 0 or 1, 0 representing absence of the variable's feature and 1 specifying the presence of that attribute. The dummy variables must be one less than the categories of the variable or zone where vehicle is travelling. In the present study, it was observed that the speed reduction of 2W and 3W are statistically similar, and therefore can be considered as a single category. 2 dummy variables (2W/3W and cars) have been used. For average speed reduction by car, car takes the value of 1 and rest 0. If both the values take 0, the speed reduction is obtained for HVs. Similarly, for zone 3 calculation, both zones 1 and 2 shall assume 0. The equation is as follows in (7) with an adjusted R square value of 0.78. The p-values came < 0.05 for all factors except for 2W/3W since they don't have very high effects in determining the average reduction in speed. The conducted validation showed less than 6% error. It can be observed from (7) that cars travelling in zone 1 have highest reduction in speed:

$$V_{reduced} = 0.89 - (1.47 \times SCMO) + (5.40 \times 2W/3W) + (7.83 \times Cars) + (6.23 \times Z1) + (2.80 \times Z2), \tag{7}$$

where $V_{reduced}$ is the percentage reduction in speed, $SCMO$ is the speed at center of MO, $2W/3W$ is the two wheelers/three wheelers, $Cars$ is the cars, jeeps and other 4-wheelers, $Z1$ is the zone 1, and $Z2$ is the zone 2.

With knowledge of traffic composition and zone composition, the average reduction in speed can be calculated with better accuracy by utilizing in (7). After calculating the average reduction in speed for each category and zone, the composition of vehicles and travelled zones composition can be used as weights to calculate total reduction in speed of traffic stream. Based on the average reduction in speed at a certain time of day, the safety undesirability index can be known. If the index is below Level B, proper steps need to be undertaken to ensure smooth traffic flow. One of the possible measures is to make the uncontrolled median openings signalized for performing U-turning manoeuvres at certain times of day like peak hours. Along with U-turning manoeuvre, a reserve lane should be provided which should be used only by U-turning vehicles. This shall ensure that approaching through traffic don't get affected by the U-turns.

4. Conclusions

In the present study, a detailed analysis has been done regarding sudden and undesirable reduction in speed of approaching through vehicles at uncontrolled median openings. Due to the presence of U-turning vehicles, the vehicles in the approaching traffic stream compel to decelerate at a rate higher than the desirable rate of braking. Due to this forced and evasive braking behavior, the chance of rear end collision increases. Moreover, these kinds of braking operations reduce the life of a vehicle and deteriorates the riding service and comfort to the road user. In the present study, field data in the form of video recording has been collected at 2 stretches of MO (10m at the start of MO and 10m within the MO) on typical weekdays on 6 lane divided MO. The descriptive statistical analysis has been done by using IBM SPSS software to find out mean, standard deviation, and skewness of speed data. The result shows positive and negative skewness of vehicular speed at the start and center of MO respectively indicating high skewness in the distribution of percentage reduction in speed. Moreover, while SD is 16-26% of mean speed values at start and center of MO, the SD to mean increases to almost 0.6 for percentage reduction in speed within the 2 stretches of MO. Speed reduction percentage has been observed to be as high as 60-70% when the vehicle enters MO. Majority of the vehicles are found to have reduced their speeds at a percentage of 10 to 30%. 2-sample t-tests between the reduction in speed values for various categories of vehicles (2W, 3W, Cars and HVS) have been conducted. The result showed that 2W and 3W show less and similar kind of speed reduction as compared to HV due to more flexibility of changing lanes while moving at median openings. In order to analyze the effect of 3 zones of 6-lane divided carriageway on which the vehicles are travelling on percentage reduction in speed, the average reduction in speed values have been collected across 3 zones. The average reduction in speeds is in the range of 22-28%, however, the maximum reduction in speed even reaches

72-76% on zone 1 and 2. Again, utilizing 3.81 m/s^2 as critical deceleration rate (IRC 66-1976), the critical safe reduction in speed from start to center of MO is calculated to be around 36.2% based on an average speed of 40.81 kmph as per IRC 66-1976. Thus, speed reduction below the critical value can be considered as safe.

Additional investigation has been done to classify the speed reduction percentage values which are below 36.2% through K-mean clustering technique. It was observed that the cluster centers converged after 15 iterations and the R square value is obtained to be 0.994. The safety index/undesirability level (A to D) depicts that index A (when percentage reduction speed < 15.53%) the safest movement in terms of applying brakes, reducing speed, and possibility of rear end collision and Index D (when percentage reduction speed > 36.20 %) is the traffic flow failure within the MO. Hence median openings should be optimally designed so that even during high traffic flow, the speeds of vehicles should not reduce by more than 24-25% from start to center of MO and maintain the safety index level B.

Moreover, it can be clearly observed that speed at the center of MO has a larger effect on percentage reduction in speed (higher absolute correlation value) as compared to speed at the start of MO. In the present study, quadratic model shows higher accuracy at the center of median opening area with an R square of 0.765 to estimate the percentage reduction in speed and is best fitted for speed range 9-50 kmph with highest mean absolute percentage error (MAPE) around 8% (<10%) hence considered strong. To establish the effect of zone and vehicle category, another multilinear probabilistic regression model is developed to predict the percentage reduction in speed considering speed at center of MO, category of vehicle and zone in which it is travelling. The model has higher accuracy value with very less MAPE of around 6%.

The detailed investigation of the present study will be helpful to counteract the forced speed reduction phenomenon at uncontrolled median openings and encourage safe and congestion free movement of approaching through vehicles. The developed equations can be used at any median opening to check for its traffic scenario and develop strategies to mitigate the undesirable and unsafe driving behavior. Further, similar studies can also be carried out for different types of roads (National Highways, Rural roads, etc.) with various lane configurations to bolster the equations with higher accuracy. For forced speed reductions above a certain percentage of vehicles, appropriate mitigation measures should be adopted. The developed equations can be applied on field to understand which category or categories of vehicles are majorly responsible for the unsafe driving behavior at a certain median opening. The study can help practitioner engineers to maintain a smooth flow of vehicles at uncontrolled median openings.

References

Aarts L, Van Schagen I. Driving speed and the risk of road crashes: a review. Accid Anal Prev. 2006; 38(2): 215-224.

Afukaar FK, Damsere-Derry J. Evaluation of speed humps on pedestrian injuries in Ghana. Inj Prev. 2010; 16(Suppl 1): A205-A206.

Al-Masaeid HR. Capacity of U-turn at median openings. ITE J. 1999;69(6): 28-30.

Anderson J, Hernandez S. Heavy-vehicle crash rate analysis: comparison of heterogeneity methods using Idaho crash data. Transp Res Rec. 2017; 2637(1): 56-66.

Antov D, Abel K, Sürje P, Rõuk H, Rõivas T. Speed reduction effects of urban roundabouts. Balt J Road Bridge Eng. 2009;4(1): 22-26.

Assemi B, Hickman M. Relationship between heavy vehicle periodic inspections, crash contributing factors and crash severity. Transp Res Part A Policy Pract. 2018; 113(C): 441-459.

Biswas S, Chandra S, Ghosh I. Estimation of vehicular speed and passenger car equivalent under mixed traffic condition using artificial neural network. Arab J Sci Eng. 2017; 42(9): 4099-4110.

Boora A, Ghosh I, Chandra S. Clustering technique: an analytical tool in traffic engineering to evaluate the performance of two-lane highways. Eur Transp. 2017; 66.

Brodie L, Lyndal B, Elias IJ. Heavy vehicle driver fatalities: learnings from fatal road crash investigations in Victoria. Accid Anal Prev. 2009; 41(3): 557-564.

Gupta A, Mondal S, Sharma VK. Modelling U-turning behaviour of vehicles at mid-block median openings in multilane urban roads. Curr Sci. 2018; 114(7): 1461-1473.

Jain AK, Dubes RC. Algorithms for clustering data. Upper Saddle River, New Jersey: Prentice-Hall; 1988.

Jain AK. Data clustering: 50 years beyond K-means. Pattern Recognit Lett. 2010; 31(8): 651-666.

Liu P, Lu JJ, Hu F, Sokolow G. Capacity of U-turn movement at median openings on multilane highways. J Transp Eng. 2008; 134(4): 147-154.

Mathews TV, Munigety CR, Bajpai A. Strip-based approach for the simulation of mixed traffic conditions. J Comput Civil Eng. 2015; 29(5), http://doi.org/10.1061/(ASCE)CP.1943-5487.0000378.

Mazaheri A, Akbarzadeh A, Rahimi AM, Akbarzadeh M. Estimation of critical gap of U-turns at uncontrolled median openings considering Iran’s driver behavior. Transp Lett. 2020; 14(6): 1-13.

Maze T, Schrock SD, Kamyab A. Capacity of freeway work zone lane closures. Proceeding of the Mid-Continent Transportation Symposium; 2000 May 15-16; USA. Iowa: Iowa State University; 2000. pp. 178-182.

Mohanty M, Dey PP. Modelling the major stream delay due to U-turns. Transp Lett. 2019a; 11(9): 498-505.

Mohanty M, Dey PP. Quantification of LOS at uncontrolled median openings using area occupancy through cluster analysis. Arab J Sci Eng. 2019b; 44(5): 4667-4679.

Mohanty M, Dey PP. Operational effects of U-turns at median opening. Transp Lett. 2021; 14(3), http://doi.org/10.1080/19427867.2021.1908491.

Mohanty M, Panda B, Dey PP. Quantification of surrogate safety measure to predict severity of road crashes at median openings. IATSS Res. 2021; 45(1): 153-159.

Mohanty M, Raj Y, Rout S, Tiwari U, Roy S, Samal SR. Operational effects of speed breakers: a case study in India. Eur Transp. 2021; 81(1), https://doi.org/10.48295/ET.2021.81.1.

Mohapatra SS, Dey PP, Chandra S. Modeling the critical position of U- turning vehicles at uncontrolled median openings. KSCE J Civil Eng. 2016; 20(1): 411-420.

Mohapatra SS, Sil GO, Dey PP. Quantification of LOS at median openings through cluster analysis. Indian Highways. 2015; 43(3): 25-31.

Mooi E, Sarstedt M. A concise guide to market research: the process, data, and methods using IBM SPSS Statistics. Berlin: Springer; 2011.

Patnaik AK, Bhuyan PK, Rao KK. Divisive analysis (DIANA) of hierarchical clustering and GPS data for level of service criteria of urban streets. Alex Eng J. 2016;55(1): 407-418.

Pecheux KK, Pietrucha MT, Jovanis PP. User perception of level of service at signalized intersections: Methodological issues. In: Proceedings of the Fourth International Symposium on Highway Capacity, Transportation Research Circular E-C018. Washington, DC: TRB, National Research Council. 2000 Jun; pp. 322-335.

Perco P, Marchionna A, Falconetti N. Prediction of the operating speed profile approaching and departing intersections. J Transp Eng. 2012; 138(12): 1476-1483.

Sil G, Mohapatra SS, Dey PP, Chandra S. Merging process of U-turns at uncontrolled median openings under mixed traffic conditions. Transport. 2018; 33(2): 370-379.

Sangngam P, Laoarun W, Rodchuen M. Adjusted clustered rank tests for clustered data in unbalanced design. Thail Stat. 2022; 20(2): 435-451.

Yao R, Sun L, Sun F, Zhao S. Optimizing the location of U-turn median openings at signalized intersections with multitype vehicles. J Transp Eng Part A Syst. 2019; 145(9), https://doi.org/10.1061/JTEPBS.0000267.

Yadav J, Sharma M. A review of K-mean algorithm. Int J Eng Trends Tech. 2013; 4(7): 2972-2975.