



Pedestrian perception-based level-of-service model at signalized intersection crosswalks

S. Marisamynathan¹  · P. Vedagiri²

Received: 15 June 2018/Revised: 1 August 2019/Accepted: 19 August 2019/Published online: 13 September 2019
© The Author(s) 2019

Abstract Pedestrian level of service (PLOS) is an important measure of performance in the analysis of existing pedestrian crosswalk conditions. Many researchers have developed PLOS models based on pedestrian delay, turning vehicle effect, etc., using the conventional regression method. However, these factors may not effectively reflect the pedestrians' perception of safety while crossing the crosswalk. The conventional regression method has failed to estimate accurate PLOS because of the primary assumption of an arbitrary probability distribution and vagueness in the input data. Moreover, PLOS categories in existing studies are based on rigid threshold values and the boundaries that are not well defined. Therefore, it is an important attempt to develop a PLOS model with respect to pedestrian safety, convenience, and efficiency at signalized intersections. For this purpose, a video-graphic and user perception surveys were conducted at selected nine signalized intersections in Mumbai, India. The data such as pedestrian, traffic, and geometric characteristics were extracted, and significant variables were identified using Pearson correlation analysis. A consistent and statistically calibrated PLOS model was developed using fuzzy linear regression analysis. PLOS was categorized into six levels

(A–F) based on the predicted user perception score, and threshold values for each level were estimated using the fuzzy *c*-means clustering technique. The developed PLOS model and threshold values were validated with the field-observed data. Statistical performance tests were conducted and the results provided more accurate and reliable solutions. In conclusion, this study provides a feasible alternative to measure pedestrian perception-based level of service at signalized intersections. The developed PLOS model and threshold values would be useful for planning and designing pedestrian facilities and also in evaluating and improving the existing conditions of pedestrian facilities at signalized intersections.

Keywords Pedestrian · Signalized intersection · Level of service · Fuzzy regression · Fuzzy *c*-means

1 Introduction

A sustainable transportation system should provide mobility and flexibility in a safe and environment-friendly mode. Walking is one of the best examples for sustainable mode of transportation which is well suited for relatively short distance in urban cities. Due to the demands of vehicular traffic, most studies have investigated the operational condition of vehicular traffic and only a few have addressed pedestrian issues. Presently, pedestrians are part of most roadway environments; therefore, transport planners and engineers must paid more attention toward increasing the walkability and improving the pedestrian safety.

✉ S. Marisamynathan
marisamy1989@gmail.com

P. Vedagiri
vedagiri@civil.iitb.ac.in

¹ Transportation Engineering and Management, Department of Civil Engineering, National Institute of Technology Tiruchirappalli, Tiruchirappalli, Tamil Nadu 400076, India

² Transportation Systems Engineering, Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India

India is characterized by mixed traffic condition, which is defined as the condition where pedestrians and vehicles share the same space without proper lane discipline. Since this condition could lead to conflicts and accidents between road users, it is necessary to provide suitable facilities to improve the serviceability and safety of pedestrians. Mumbai is the most populous city in India with a population estimate of 21 million as of 2014. In 2012 and 2014, deaths due to conflicts between pedestrian and vehicle accounted for 8.3% and 12.9%, respectively, of the total number of road accident deaths in Mumbai, India [1, 2]. It is inferred that an increase in the percentage of pedestrian death is due to the decrease in pedestrian serviceability and safety. Thus, understanding the need of pedestrians is important in transportation design and management of pedestrian facilities.

Transport planners and designers need to examine performance measures that could make pedestrians safe, convenient, and comfortable while crossing signalized intersections. User perception-based classification of pedestrian level of service (PLOS) is important for traffic engineers to distinguish the operational performance of pedestrian facilities. In this study, we made an attempt to develop a PLOS model using pedestrians' perceptions of safety, convenience, and efficient level at signalized intersections. Based on PLOS classification, pedestrian facilities can be improved and better planned for future usage.

2 Review of earlier studies

Estimation of PLOS is much more complicated than that of motorized vehicle level of service (LOS), as it is more flexible [3]. Since PLOS is the most common measure of effectiveness in evaluating the walking conditions of existing pedestrian facilities, various studies have focused on PLOS model development at intersections, sidewalks [4–6], midblocks [7], stairways [8], and roadway segments [9, 10].

In 1965, the concept of LOS was originally introduced in the Highway Capacity Manual (HCM) for determining vehicle LOS based on the average stopped delay at signalized intersections. Later, HCM 2000 provided guidance to design and develop pedestrian facilities based on the quantitative measure of pedestrian delay and space requirements in six LOS categories (*A–F*) [11]. Several research studies have identified influencing variables other than pedestrian delay and space requirements, involving pedestrian characteristics, traffic characteristics, geometric characteristics, or combination of any of these characteristics. Pedestrian volume, bidirectional effect, platoon effects, area occupancy, pedestrian flow, walking speed,

and flow ratio were used to develop PLOS models at signalized intersections [12–14]. Most of the existing PLOS models focus on improving pedestrian safety and comfort by considering potential conflicts between pedestrians and vehicles, vehicle volume, left and right turning vehicles, and speed of vehicles [15–19]. Various studies have proposed PLOS models with respect to waiting time or delay, green signal, congestion level, crosswalk width, level of space at corner, crossing facilities, traffic signs, intersection geometry features, and land use [13, 16, 20–24].

Recently, researchers have started to address PLOS using qualitative and quantitative data with the qualitative data representing the perception of pedestrian safety and comfort at signalized intersections [15, 21, 22]. Considering the various research outcomes, HCM 2010 has also revised the method for PLOS estimation with respect to qualitative and quantitative data. Demand flow, pedestrian delay, vehicle volume, speed, number of lanes, and number of right turn channelizing islands were incorporated in the mathematical PLOS score model, and LOSs based on traveler perception are listed [25].

Limited studies have developed PLOS models for Indian conditions [17, 26–29]. The pedestrian facilities such as sidewalks and crosswalks were evaluated with respect to peak hour pedestrian flow, and LOS rating was defined with respect to flow by using IRC 103-2012 [30]. One PLOS model at sidewalk was developed to assess various infrastructure facilities from land use pattern considering pedestrians' perceptions, and the qualitative data-based PLOS model was built by adopting point system techniques [27]. Another PLOS model for sidewalk was developed by combining qualitative and quantitative data and adopting stepwise regression analysis method and Khisty's relationship chart [29]. The recently developed Indian Highway Capacity Manual (Indo-HCM) and Indian Road Congress (IRC 103-2012) defined the PLOS rating (*A–F*) with respect to pedestrian delay (in seconds) for crosswalk at signalized intersections [31, 32]. The Indo-HCM defined that PLOS is a measure for assessing the operating conditions of facilities in a quantitative manner. Therefore, the PLOS rating at crosswalk developed in existing Indian studies and manuals are based on quantitative or qualitative measure, but combination of qualitative and quantitative measure is still missing in Indian conditions, which is a research gap to be addressed in the present study.

Many researchers have developed PLOS models using conventional regression methods such as linear, generalized linear, stepwise multivariables, or binary/cumulative/ordered logit models [13, 17, 18, 21, 24]. The most widely used conventional regression method is generally based on two assumptions: (1) the observations follow some probability distributions, especially the normal distribution, and (2) the mean values are varying with respect to independent

variables; and the regression equation is very simple to develop. In practice, it has been found that the data are fuzzy in nature and cannot be defined by probability distributions. The data described in a subjective type such as “very dangerous” and “approximately equal to 5” are typical examples. Fuzzy linear regression (FLR) method is a more suitable technique to develop regression models when the observations are subjective, complex, and of qualitative data type. The FLR based on a possibility distribution reflects the membership values of the dependent variable rather than a probability distribution [33].

In HCM 2010, the listed range of scores associated with each LOS criterion are common for both pedestrian and bicycle travel modes at signalized intersections, and also the threshold values are not stated [25]. Various studies follow a medium value or step function method to define threshold values for PLOS criteria [17, 18, 22], and the rigid boundaries between LOS criteria have failed to represent the pedestrian operational conditions. Cluster analysis is recognized as an important technique for classifying data, by finding clusters and dissimilarities between different data clusters. Most of the existing studies use *K*-means clustering in transportation problems, such as defining LOS for urban arterial roads, LOS for freeways, and vehicle LOS at intersections [34]. Few researchers have utilized fuzzy *c*-means clustering technique for defining vehicle LOS classification at intersections and urban roads [35–38]. None of the studies, at least to the author’s knowledge, have utilized fuzzy *c*-means clustering techniques for defining threshold values for PLOS at signalized intersections in India.

Earlier works in PLOS for signalized intersections have considered almost all of the factors influencing pedestrian sense of safety, comfort, and convenience while crossing intersections. However, pedestrian delay with respect to pedestrian crossing behavior, traffic volume under heterogeneous conditions, median width, and conflict between pedestrian and vehicle, which also represent pedestrian safety, convenience and efficiency level while crossing signalized intersections, are missing in the existing PLOS studies. A few studies have developed PLOS for Indian conditions using the conventional linear regression (CLR) method but the estimates fail to provide accurate PLOS [17, 27, 29]. Existing international standard threshold values for PLOS do not suit for Indian conditions due to mixed traffic and variations in users behaviors. This study attempts to contribute an alternative methodology for PLOS model development and defines threshold values that may overcome limitations of the conventional regression analysis and the medium value method. The FLR method is used to develop the PLOS model, and the fuzzy *c*-means clustering technique is applied to define the threshold value of each PLOS classification.

3 Methodology

The PLOS model proposed in this study follows a novel procedure of combining quantitative as well as qualitative data, and the major steps involved in PLOS model development are outlined in Fig. 1. The data collection and extraction from selected nine signalized intersections are discussed in detail. Details about identification of significant variables, PLOS model development, definition of threshold values, and validation of the model are discussed next.

3.1 Model formulation

The present study considers two techniques, fuzzy linear regression (FLR) and fuzzy *c*-means (FCM) clustering techniques for modelling and classifying PLOS.

3.1.1 Fuzzy linear regression model

In past, a variety of deterministic and stochastic models have been developed and different kinds of formulas, equations, and assumptions were used to solve all kinds of complex transportation engineering problems. Most of the transportation problems are solved using subjective data such as user response, which is very difficult to quantify. Transportation parameters are characterized by uncertainty, subjectivity, imprecision, and ambiguity. Drivers and pedestrians use this subjective knowledge information on a daily basis while making decisions. In the conventional regression analysis, the observations are assumed to follow some particular probability distributions, most of the time, the normal distributions, and the means of these particular

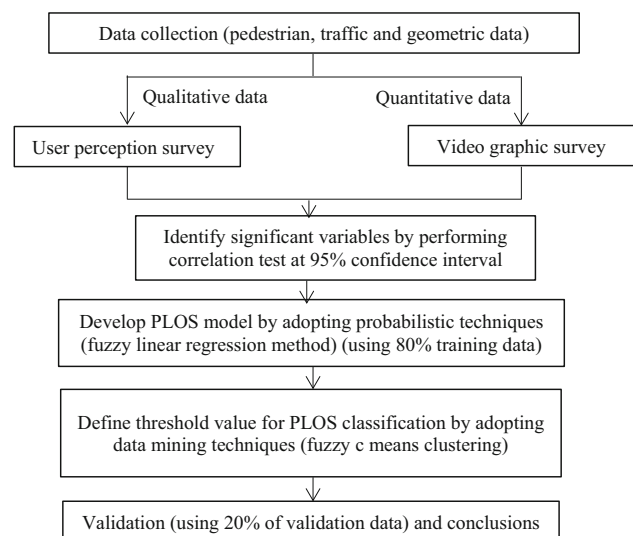


Fig. 1 Major steps involved in PLOS model development

probability distributions vary in some systematic fashion with the values of the independent variables. In practice, sometimes it is difficult to find a probability distribution of the observations, especially when the observations are fuzzy in nature. The observations described by subjective judgment or linguistic terms such as excellent (equal to 1), good (equal to 2), average (equal to 3), poor (equal to 4), and terrible (equal to 5) are typical examples. The parameter estimation under this fuzzy environment is a challenge to the CLR analysis. Therefore, the existing CLR-based model was not fitted to estimate solution for qualitative or subjective data, especially in LOS model development [7, 24]. However, these problems can be rectified with the FLR method [39, 40] to develop LOS model at signalized intersections. FLR analysis is the best method to develop model with fuzzy data and also it allows possibility distributions [41–43].

FLR based on possibility distribution reflects the membership values of the dependent variables rather than a probability distribution [41, 42, 44]. The generalized form of FLR model is as follows:

$$Y_i = A_0 + A_1X_{i1} + A_1X_{i1} + \dots + A_jX_{ij} + \varepsilon_i, \quad (1)$$

$$i, j = 1, 2, \dots, n,$$

where for each j , $X_{1j}, X_{2j}, \dots, X_{nj}$ are the values of n fuzzy or crisp independent variables X_1, X_2, \dots, X_n , and Y_i is the corresponding value of fuzzy dependent variable Y ; A_0, A_1, \dots, A_n are the coefficients of the model and ε_i is the fuzzy error term. The observations of X_i and Y_i are fuzzy; therefore, it falls into the category of fuzzy regression analysis. In this section, the four-stage FLR model development is described step by step with a single independent and dependent fuzzy variable, and the same methodology is applied to solve the multivariable problem.

Stage 1 Initialize the dependent and independent variables. Identify the fuzzy and crispy variables from the listed variables.

Stage 2 Apply the membership function to the fuzzy variables. In this section, triangular membership function is considered with three point values such as lower, middle or average, and upper value. The FLR model for one independent and one dependent fuzzy variable is

$$\check{y}_{i(L,M,U)} = \check{\beta}_{0(L,M,U)} + \check{\beta}_{1(L,M,U)}\check{x}_{i(L,M,U)} + \varepsilon_{i(L,M,U)}, \quad (2)$$

$$i = 1, 2, 3, \dots, n,$$

where L, M , and U indicate the lower, middle or average, and upper values of corresponding variables at selected interval or cycle; $\check{y}_i, \check{\beta}_0, \check{\beta}_1$ and \check{x}_i are all triangular fuzzy numbers, and therefore, $\check{\beta}_{0(L,M,U)} + \check{\beta}_{1(L,M,U)}\check{x}_{i(L,M,U)}$ is the

approximately triangular fuzzy number $(\check{\beta}_{0(L)} + \check{\beta}_{1(L)}\check{x}_{i(L)}, \check{\beta}_{0(M)} + \check{\beta}_{1(M)}\check{x}_{i(M)}, \check{\beta}_{0(U)} + \check{\beta}_{1(U)}\check{x}_{i(U)})$.

Stage 3 To find an estimate for $\check{\beta}_0$ and $\check{\beta}_1$, and minimize the following sum of squared distances [42]:

$$S(A_0, A_1) = \sum_{i=1}^n d^2(\check{y}_i, \check{\beta}_0 + \check{\beta}_1\check{x}_i) \quad (3)$$

$$= \sum_{i=1}^n \left[\left(y_{i(L)} - \beta_{0(L)} - \beta_{1(L)}x_{i(L)} \right)^2 + \left(y_{i(M)} - \beta_{0(M)} - \beta_{1(M)}x_{i(M)} \right)^2 + \left(y_{i(U)} - \beta_{0(U)} - \beta_{1(U)}x_{i(U)} \right)^2 \right], \quad (4)$$

where $S()$ is total sum of squares; d is explained sum of squares; x_i and y_i are the triangular fuzzy member of independent and dependent variables, respectively; β_0 and β_1 are coefficients of model.

Differentiating with respect to $\beta_{0(L)}, \beta_{0(M)}, \beta_{0(U)}, \beta_{1(L)}, \beta_{1(M)}$, and $\beta_{1(U)}$, we obtain the following estimates of $\check{\beta}_0$ and $\check{\beta}_1$:

$$\check{\beta}_1 = (\check{\beta}_{1(L)}, \check{\beta}_{1(M)}, \check{\beta}_{1(U)}) \quad \text{and}$$

$$\check{\beta}_0 = (\check{\beta}_{0(L)}, \check{\beta}_{0(M)}, \check{\beta}_{0(U)}),$$

$$\check{\beta}_{1(L)} = \frac{\sum_{i=1}^n \check{x}_{i(L)}\check{y}_{i(L)} - n\bar{x}_L\bar{y}_L}{\sum_{i=1}^n \check{x}_{iL}^2 - n\bar{x}_L^2}, \quad \text{and} \quad (5)$$

$$\check{\beta}_{0(L)} = \bar{y}_L - \check{\beta}_{1(L)}\bar{x}_L,$$

$$\check{\beta}_{1(M)} = \frac{\sum_{i=1}^n \check{x}_{i(M)}\check{y}_{i(M)} - n\bar{x}_M\bar{y}_M}{\sum_{i=1}^n \check{x}_{iM}^2 - n\bar{x}_M^2}, \quad \text{and} \quad (6)$$

$$\check{\beta}_{0(M)} = \bar{y}_M - \check{\beta}_{1(M)}\bar{x}_M,$$

$$\check{\beta}_{1(U)} = \frac{\sum_{i=1}^n \check{x}_{i(U)}\check{y}_{i(U)} - n\bar{x}_U\bar{y}_U}{\sum_{i=1}^n \check{x}_{iU}^2 - n\bar{x}_U^2}, \quad \text{and} \quad (7)$$

$$\check{\beta}_{0(U)} = \bar{y}_U - \check{\beta}_{1(U)}\bar{x}_U,$$

where $\bar{x}_L, \bar{x}_M, \bar{x}_U, \bar{y}_L, \bar{y}_M$, and \bar{y}_U are mean of the respective independent and dependent variables, respectively.

Stage 4 Many defuzzification methods have been proposed in the existing studies and the centroid method is the most prevalent and physically appealing one, as it weighs all the values with different possibilities to form a single value [45, 46]. Let \check{y}_i be the defuzzified values of $\check{y}_{i(L,M,U)}$ and the solution based on centroid method is

$$\check{y}_i = \frac{\int_{-\infty}^{+\infty} y\mu_Y(y)dy}{\int_{-\infty}^{+\infty} \mu_Y(y)dy} = \frac{1}{3} (\check{y}_{i(L)} + \check{y}_{i(M)} + \check{y}_{i(U)}), \quad (8)$$

where $\mu_Y(y)$ is the deviation of each variable from the mean.

This method is used to develop PLOS models, and the same procedure can be applied for solving the multiple variable fuzzy regression models.

3.1.2 Fuzzy *c*-means clustering

Existing studies define six LOSs for each type of facility, from *A* to *F*, and the medium value method or step function approach is used to estimate threshold values for various LOS. The use of distinct boundaries limits the subjective and continuous characteristics of the quality of service and the rigid threshold values fail to provide accurate LOS. From the existing literature, it was found that data mining is the most suitable method for the classification of large data, especially cluster analysis. Cluster analysis refers to the process of partitioning a finite data set of objects (*X*) into a certain number of natural and homogeneous clusters (*c* subsets). There are various methods of data clustering and every method has its own advantages and disadvantages. In this study, an FCM clustering technique was used to determine the threshold values for different classification of LOS. In the FCM technique, each qualitative data point is not rigidly assigned to only one cluster, but rather to a few clusters with certain membership values [35]. The framework of FCM method is given below in a step-by-step procedure.

Step 1 For a data set consisting *N* data points, select the desired number of clusters, *k*, where $2 < k < N$.

Step 2 Generate a starting center point for each of the *k* clusters.

Step 3 Calculate the distance from each of the *N* points to each of the *k* clusters, i.e., d_{ik} . The distance is calculated using Euclidean distance, the most widely used method, as follows:

$$d_{ik} = \sqrt{\left(\sum_{i=1}^N |x_i - y_k|^2 \right)}, \quad i = 1, 2, 3, \dots, n. \quad (9)$$

Step 4 Assign a proportional or fuzzy membership of the *N* points to each of the *k* clusters.

Step 5 Find the new center point (*c*) for each of the *k* clusters by finding the weighted average:

$$c_k = \frac{\sum_{i=1}^n [(u_{ki})^q (x_i)]}{\sum_{i=1}^n [(u_{ki})^q]}, \quad (10)$$

where *x* is the *n*th data point, *u* is the cluster membership of the data point for the *n*th cluster *k*, and *q* is the fuzzy exponent, calculated from Step 4.

Step 6 Repeat Steps 3, 4, and 5 until there are no changes in the cluster membership or proportional.

3.2 Factors influencing PLOS at signalized intersections

Based on the outcome of literature reviews and field observations (preliminary survey), several possible factors influencing PLOS at signalized intersections were identified for Indian conditions. The selected factors are listed in three main categories as shown in Table 1.

Interaction between pedestrians and vehicles was also considered as an important factor of influence in PLOS. India is known for its heterogeneous traffic conditions with high pedestrian and traffic flow; thus, interaction between pedestrians and vehicles is highly possible, due to the 60% pedestrian noncompliance behavior [47]. In this study, interaction between a pedestrian and a vehicle is defined in two ways: (1) due to pedestrian noncompliance behavior with signal during pedestrian nongreen phase, and (2) due to free left turn or noncompliance behavior of drivers during pedestrian green phase. During noncompliance behavior, the rolling behavior of pedestrians was not considered in this study. According to the above explanation, there are only two possible outcomes which are 0 to 1: 0 for pedestrian not interacting with vehicle in crosswalk and 1 for pedestrian interacting with vehicle in crosswalk. The model for probabilities of interaction between pedestrian and vehicle in crosswalk is adopted from the authors' published works [47]. The study identified that pedestrian crossing direction, suitable gap, approaching vehicle direction, and approaching vehicle lane have a significant effect on pedestrian–vehicular interaction in crosswalks according to Pearson's correlation test conducted at the 95% confidence interval in SPSS 16.0 software. A binary logit model was developed with NLOGIT 4 software.

4 Data collection

The data were collected by conducting user perception and video-graphic surveys during peak hour at signalized intersections. The user perception survey was conducted with the help of investigators placed on selected signalized intersections under typical traffic conditions to question about pedestrians' perception on level of safety, convenience, and efficiency while crossing the intersections. Simultaneously, video-graphic survey was performed to collect information on several primary factors influencing PLOS at signalized intersections.

4.1 Site selection

To cover the typical traffic and roadway conditions in India, nine signalized intersections were selected from a list of all signalized intersections in Mumbai suburban area

Table 1 Selected variables and descriptions

| Category | Variable | Measures | Description |
|------------------------------|---|----------|--|
| Pedestrian characteristics | Pedestrian gender | Ordinal | 0 for female pedestrian and 1 for male pedestrian |
| | Age groups | Ordinal | 0 for child pedestrian (less than 18 years), 1 for adult pedestrian (18–50 years) and 2 for elderly pedestrian (more than 50 years) |
| | Baggage | Ordinal | 0 for without baggage and 1 for with baggage |
| | Pedestrian delay (s/ped/cycle) | Scalar | Pedestrians waiting to cross the intersection from nongreen phase to green phase |
| | Pedestrian crossing speed (m/s) | Scalar | Crosswalk length is divided by actual travel time of particular pedestrian |
| Traffic characteristics | Compliance behavior | Scalar | 0 for pedestrian crossing the crosswalk during green phase and 1 for pedestrian crossing the crosswalk during nongreen phase |
| | Traffic volume along with pedestrian crossing (veh/crosswalk/cycle) | Scalar | Sum of three kinds of vehicle volumes were considered: (1) The volume of through movement vehicles and left turning vehicles from the street parallel to the crosswalk during pedestrian green phases (2) Volume of moving vehicles through crosswalk while pedestrian waiting in waiting area or median during pedestrian nongreen phases and (3) The number of vehicles occupied in crosswalk at upstream side for each pedestrian |
| | Probability of interaction between pedestrian and vehicle | Scalar | The probabilities of pedestrian interaction with vehicle in crosswalk. It ranges from 0 to 1. 0 for pedestrian not interacting with vehicle in crosswalk and 1 for pedestrian interacting with vehicle in crosswalk |
| | Approaching vehicle direction | Ordinal | Pedestrian finding suitable gap from approaching vehicle and start crossing during pedestrian nongreen phases. 0 for through movement vehicles, 1 for right turning vehicle and 2 for left turning vehicles |
| | Approaching vehicle lane (AVL) | Ordinal | The lane at which the first vehicle reaches the crosswalk during pedestrian noncompliance with traffic signal. From curb to median considered as 1, 2, 3 |
| | Suitable gap (s) | Scalar | Time difference between pedestrian departing time and nearby approaching vehicle to reach the pedestrian crosswalk |
| | Pedestrian crossing direction | Ordinal | 0 for pedestrian upstream to downstream movement (U-D) and 1 for pedestrian downstream to upstream movement (D-U) |
| Intersection characteristics | Crosswalk marking | Ordinal | 0 for absence of pedestrian crosswalk marking and 1 for presence of pedestrian crosswalk marking |
| | Green time ratio | Scalar | Ratio between allotted green time for pedestrian and cycle time |
| | Median width (m) | Scalar | The width of provided median |
| | Connection between sidewalk and crosswalk | Ordinal | 0 for absent connection and 1 for available connection between sidewalk and crosswalk |
| | Free left turning for vehicles | Ordinal | 0 for prohibited free left turn and 1 for allowed free left turn |
| | Exclusive left turning lane | Ordinal | 0 for absence of exclusive left turning lane and 1 for availability of exclusive left turning lane |
| | Crosswalk length (m) | Scalar | Length of crosswalk |

Ordinal—discrete variable; scalar—continuous variable

with high pedestrian demand. All approaches at the selected signalized intersections have two-way traffic and bidirectional pedestrian flow without any table top crossing coupled with blinking green signal for left turning traffic movement. The sites were of a typical four-arm type with fixed signal time and phase time. Basic information on the study sites are presented in Table 2.

4.2 Video-graphic survey

The characteristics of the pedestrian, traffic, and geometric conditions at signalized intersections were collected with a video-graphic survey. Cameras were set up in the direction of pedestrian upstream to downstream movement and downstream to upstream movement at selected crosswalk in each intersection. The video-graphic survey covered the selected crosswalk and recorded pedestrian movements throughout the study area. The required data were extracted manually in laboratory; it takes longer duration but has the

Table 2 Selected study locations and basic information

| Sl. no. | Location | Crosswalk identity | Pedestrian flow (ped/h) | Crosswalk length (m) | Pedestrian signal time (s) | | | | Free left turn |
|---------|------------------------------|--------------------|-------------------------|----------------------|----------------------------|----------------|-----|------------------|----------------|
| | | | | | Green | Flashing green | Red | Number of cycles | |
| 1 | Link Road Junction | <i>A</i> | 175 | 27 | 23 | 4 | 152 | 25 | Prohibited |
| 2 | Malad Junction | <i>B</i> | 395 | 22.4 | 26 | 3 | 149 | 25 | Prohibited |
| 3 | Mahim Junction | <i>C</i> | 402 | 20 | 35 | 2 | 106 | 31 | Permitted |
| 4 | Mahatma Gandhi Road Junction | <i>D</i> | 89 | 27.6 | 36 | 3 | 114 | 29 | Permitted |
| 5 | Holkar Junction | <i>E</i> | 337 | 31.5 | 22 | 3 | 121 | 30 | Prohibited |
| 6 | Samaj Junction | <i>F</i> | 148 | 19 | 12 | 3 | 118 | 33 | Permitted |
| 7 | Chembur Naka Junction | <i>G</i> | 303 | 27 | 20 | 3 | 132 | 25 | Prohibited |
| 8 | Andheri-Link Road Junction | <i>H</i> | 271 | 27 | 19 | 3 | 138 | 28 | Permitted |
| 9 | Santacruz Junction | <i>I</i> | 120 | 18 | 16 | 2 | 116 | 33 | Permitted |

advantage of providing more accurate data and long-lasting record of events.

4.3 User perception survey

Questions were asked to pedestrians immediately after crossing the crosswalk by a number of trained investigators who stood on either side of the crosswalk. Meaning and importance of each question have been explained to each pedestrian personally and were also stated in user perception survey form before acquiring their rating. Ratings range from 1 to 5 for pedestrian safety, convenience, and efficiency level. The sample photos of the investigators conducting the survey are shown in Fig. 2. Figure 3 shows a sample user perception survey form and the explanation of LOS that was given to pedestrians.

4.4 Data

The required data were extracted from collected video tapes using AVS video editor software at an accuracy of 33 ms (30 frames per second). The listed variables were extracted from video for each pedestrian who have responded to the user perception survey, and the descriptive statistics are shown in Table 3.

The user perception survey form comprised three questions about how pedestrians perceive when they cross crosswalk at signalized intersections. A total of 654 pedestrians participated in the user perception survey: 522 males and 132 females. About 7% of pedestrians were less than age 18, 79% were 18–50, and 14% were aged 50 or older. The gender and age distribution of each survey site are shown in Table 4. A selected portion of a larger population is chosen as samples to gain information about the whole population. Sample size test were conducted at 99%

confidence interval. Total population is 2,240 pedestrians from selected sites, and it is reasonable to assume a margin error of 5% with a response rate of 50%, which yields an approximate sample size of 513 pedestrians. The value of statistical sample size result (513 ped) is less than real-time respondents' size (654 ped); it shows the significance level of collected samples for developing PLOS models.

Three sets of scores were collected from respondents and analyzed in four groups, namely the safety, convenience, efficiency, and combined score (sum of safety, convenience and efficiency scores). The observed PLOS score was the average combined score, and was computed by applying the centroid method between combined scores at each cycle. In this study, the triangular membership function was used for model formulation. Therefore, the outcome of the centroid method was equal to the average combined scores. The detailed statistics of each score group for all locations are summarized in Table 5.

5 PLOS model

PLOS is an important measure of performance in analyses of signalized intersections based on user perceptions. In this study, an alternative methodology is proposed for evaluation of PLOS at signalized intersections. A PLOS model is developed and validated using data from nine locations, which consists of 654 pedestrian samples. The modeling process includes the following steps: (a) identification of significant variables via Pearson's correlation, (b) estimation of the coefficients for significant variables in the FLR model, and (c) definition of the threshold value for each LOS category based on FCM clustering.



Fig. 2 Selected crosswalk (E) and investigators with respondents pedestrians

User perception survey form for Pedestrian LOS

1) Gender: Male Female

2) Age: Child(<18 years) Adult(18-50 years) Elderly (>50years)

3) How do you feel the efficient level when you cross the crosswalk?

1 Excellent
2 Good
3 Average
4 Poor
5 Very poor

It was explained to the respondents: the purpose of this question is to ask the respondent's sense of efficient level (ex: signal time, delay)

4) How do you rate with respect to safety when you cross this crosswalk?

1 Highly safe (excellent)
2 Safe (normal)
3 Average
4 Risk (dangerous)
5 Highly risk (very dangerous)

It was explained to the respondents: the purpose of this question is to ask the respondent's sense of safety (ex: interaction, exclusive turning lane, traffic volume)

5) Mark your convenient level while crossing this crosswalk

1 More convenient
2 Convenient
3 Moderate
4 Inconvenient
5 Highly inconvenient

It was explained to the respondents: the purpose of this question is to ask the respondent's sense of convenient in the areas of geometric features (ex: crosswalk, median, connection, waiting area), and traffic conditions (vehicle flow, Interaction)

6) Pedestrian crossing during - Green Red Flashing red

Fig. 3 Pedestrians perceptions survey form

Table 3 Statistics of selected variables and Pearson correlation results

| Variable | Minimum | Maximum | Average | Pearson correlation | |
|---|---------|---------|---------|----------------------|-------|
| | | | | Value | Sign |
| Traffic volume along with pedestrian crossing (veh/crosswalk/cycle) | 5 | 268 | 50.46 | 0.341 ^a | 0.000 |
| Probability of interaction between pedestrian and vehicle | 0 | 0.8238 | 0.2978 | 0.173 ^a | 0.028 |
| Crosswalk length (m) | 19 | 31.5 | 25.19 | - 0.014 | 0.815 |
| Median width (m) | 0.8 | 2.5 | 1.7 | - 0.212 ^a | 0.000 |
| Crosswalk marking | 0 | 1 | - | 0.079 | 0.184 |
| Connection between sidewalk and crosswalk | 0 | 1 | - | - 0.092 | 0.119 |
| Crossing speed (m/s) | 0.67 | 1.66 | 1.11 | - 0.024 | 0.691 |
| Green time ratio | 0.11 | 0.26 | 0.17 | 0.027 | 0.612 |
| Pedestrian delay (s/ped/cycle) | 0 | 120 | 7.03 | 0.232 ^a | 0.004 |
| Combined score ^b | 4 | 15 | 8.71 | 1 | - |

^aSignificant at 95% confidence level

^bCombined score is the sum of safety, convenience and efficiency scores

Table 4 Gender and age distribution of each surveys site

| Crosswalk identity | Number of respondents | Gender | | Age | | |
|--------------------|-----------------------|--------|--------|-----------------|-------|--------------|
| | | Male | Female | Younger than 18 | 18–50 | 50 and older |
| <i>A</i> | 51 | 39 | 12 | 2 | 43 | 6 |
| <i>B</i> | 40 | 26 | 14 | 3 | 32 | 5 |
| <i>C</i> | 83 | 61 | 22 | 6 | 66 | 11 |
| <i>D</i> | 67 | 58 | 9 | 3 | 54 | 10 |
| <i>E</i> | 126 | 112 | 14 | 3 | 105 | 18 |
| <i>F</i> | 87 | 82 | 5 | 9 | 68 | 10 |
| <i>G</i> | 58 | 47 | 11 | 3 | 46 | 9 |
| <i>H</i> | 76 | 52 | 24 | 6 | 63 | 7 |
| <i>I</i> | 66 | 45 | 21 | 7 | 41 | 18 |

Table 5 Descriptive statistics of pedestrian perception score

| Perceived score | Crosswalk identity | | | | | | | | |
|----------------------|--------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| | <i>A</i> | <i>B</i> | <i>C</i> | <i>D</i> | <i>E</i> | <i>F</i> | <i>G</i> | <i>H</i> | <i>I</i> |
| Safety score | | | | | | | | | |
| Min | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 1 |
| Avg | 3.14 | 3.60 | 3.07 | 3.16 | 2.83 | 2.99 | 3.38 | 2.84 | 3.26 |
| Max | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| SD | 1.059 | 1.150 | 1.135 | 1.123 | 1.026 | 1.186 | 1.023 | 0.939 | 0.788 |
| Convenience score | | | | | | | | | |
| Min | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 2 |
| Avg | 2.73 | 2.93 | 2.67 | 3.10 | 2.52 | 2.60 | 3.02 | 2.50 | 3.19 |
| Max | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 4 | 4 |
| SD | 0.896 | 0.888 | 0.843 | 0.923 | 0.817 | 0.855 | 0.946 | 0.721 | 0.652 |
| Efficiency score | | | | | | | | | |
| Min | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| Avg | 3.24 | 3.48 | 2.92 | 2.90 | 2.77 | 2.68 | 3.17 | 2.92 | 2.91 |
| Max | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 4 |
| SD | 0.790 | 0.816 | 0.844 | 0.855 | 0.792 | 0.856 | 0.798 | 0.796 | 0.564 |
| Combined score | | | | | | | | | |
| Min | 5 | 5 | 4 | 6 | 4 | 4 | 5 | 4 | 7 |
| Avg | 9.10 | 10.00 | 8.66 | 9.16 | 8.13 | 8.26 | 9.57 | 8.26 | 9.36 |
| Max | 14 | 14 | 13 | 15 | 15 | 13 | 13 | 13 | 13 |
| SD | 2.119 | 2.088 | 2.091 | 2.307 | 2.063 | 2.165 | 2.177 | 1.942 | 1.442 |
| Samples (<i>N</i>) | 51 | 40 | 83 | 67 | 126 | 87 | 58 | 76 | 66 |

SD standard deviation

5.1 Identification of significant variables

Observed PLOS scores (sum of safety, convenience, and efficiency scores) of each pedestrian who have responded to user perception survey were calculated and compared with all the possible influencing variables which were listed in Table 3. The significant variables were identified by Pearson correlation test for PLOS model development. A

95% confidence level was selected with a corresponding critical *p* value of 0.05. If the calculated *p* values were less than 0.05, then the corresponding variables were included in the model and the other variables were neglected. The Pearson correlation test was performed using the statistical package for the social science (SPSS 16.0) and the results are also shown in Table 3.

From Table 3, the factors, traffic volume along with pedestrian crossing, pedestrian delay, probability of interaction between pedestrian and vehicle, and median width were found to have significant effect on the PLOS at signalized intersections. The following section briefly outlines some of the aspects of the variables and how the dependent variable responds to them.

5.1.1 Vehicular volume

The vehicular volume had more significant effect with a positive sign at 99% confidence interval. The effect of vehicular volume is obtained by adding volume of parallel through movement vehicles (V_{parallel}), volume of moving vehicles while pedestrian waiting in the waiting area or median (V_{moving}), and volume of occupied vehicles in crosswalk (V_{occupied}) while pedestrian crossing crosswalk. As vehicle frequency increases, the pedestrian sense of safety and comfort decreases and the PLOS score increases.

From selected study locations, pedestrian green phases were found to share with vehicular movement which are moving parallel to pedestrian crosswalk and it highly affects pedestrian sense of safety. Pedestrian's waiting in the waiting area or in the median feel insecure and inconvenient while facing high vehicle movement. Vehicles were also found to occupy pedestrian crosswalk marking and crossing area, influencing pedestrian safety and convenience levels.

5.1.2 Pedestrian delay

Ensuring efficiency level, proper signal phase, and signal timing is essential in pedestrian sense of efficiency at the signalized intersection. Efficiency level is affected by pedestrian waiting time delay or pedestrian nongreen phases. Pedestrian waiting time delay had significant effect on PLOS score with a positive sign. It is inferred that an increase in the value of pedestrian delay results in increase of PLOS score. The pedestrians arriving at green phase did not receive any waiting time delay, and the minimum value for waiting time delay is zero.

5.1.3 Probability of pedestrian interaction with vehicle in crosswalk

The probability of pedestrian interacting with vehicle in crosswalk had significant effect on PLOS score with positive sign, and it indicates that an increase in the value of probabilities of interaction results in an increase in the PLOS score. The value of probability of interaction is calculated by the following equation [47]:

$$P_{pv} = -0.161 - 0.469D_{av} + 0.489L - 0.35G_s + 0.908D_{pc}, \quad (11)$$

where P_{pv} denotes the probability of pedestrian interaction with vehicle in crosswalk, D_{av} the approaching vehicle direction, L the approaching vehicle lane, G_s the suitable gap, and D_{pc} the pedestrian crossing direction. The mentioned parameters are described in Table 1.

In Eq. (11), the sign of approaching vehicle direction is negative; i.e., pedestrians are feeling vulnerable when they encounter a left turning vehicle. The signal phases allow vehicles for free left turning and this affects the pedestrian safety during green phases. Usually, pedestrians identify the gap size from the first lane of the crosswalk to the approaching vehicle and will not comply with traffic signal based on that. However, for lanes farther from the pedestrian, they are unable to identify the gap size and often interact with vehicles which may result in accidents or delay to pedestrians, and hence, the sign of approaching vehicle lane is positive. The sign of suitable gap is negative; i.e., as the pedestrian acceptance gap size increases, the frequency of interaction between vehicles and pedestrian decreases. The acceptance of a smaller gap size causes the pedestrian to feel risky. The sign of pedestrian crossing direction is positive, i.e., the pedestrian safety level during crossing the D-U direction is lower than the U-D direction crossing. The probability of pedestrian interacting with vehicle in crosswalk is zero when the pedestrian and driver comply with traffic signals.

5.1.4 Pedestrian facilities

Pedestrian sense of safety and convenience in the signalized intersection environment are highly affected by the presence of refuge island and median width conditions. According to Pearson's correlation result, the value of median width had significant effect on PLOS at signalized intersection: The sign of model variable is negative; i.e., as the median width increases, the PLOS score decreases. Shorter median width or refuge island affects the pedestrian convenience level. When the median width is decreased, pedestrians are forced to wait in crosswalk or road and it ultimately affect pedestrians' sense of convenience and safety.

5.2 Development of PLOS model

The four significant variables identified from the correlation test were considered to be the most probable primary factors that affect pedestrians' sense of safety, convenience, and efficiency at signalized intersections, and these variables were used to develop the PLOS model. The

Table 6 Model parameters with descriptions

| Model parameters | Expressions | Descriptions |
|--|---|---|
| X_1 Traffic volume | $V_{\text{parallel}} + V_{\text{moving}} + V_{\text{occupied}}$ | V_{parallel} = Volume of parallel through movement V_{moving} = Volume of moving vehicle V_{occupied} = Volume of occupied vehicle in crosswalk |
| X_2 Pedestrian delay | $\frac{\alpha_1(C-(G+\alpha_2R))^2}{2C}$ | α_1 = Correction factor for non-uniform arrival rate and this value is calculated from the equation $\alpha_1 = 0.002V_{\text{Red-Ped}} + 0.734$, where $V_{\text{Red-Ped}}$ = Number of pedestrians arriving during pedestrian red time per hour C = Cycle time in seconds G = Pedestrian green time in seconds α_2 = Percentage of pedestrians who start crossing during nongreen phases R = Pedestrian red time in seconds |
| X_3 Probability of pedestrian interaction with vehicle in crosswalk | P_{pv} | See Eq. (11) |
| X_4 Median width | W_{median} | W_{median} = Median width |

X_0 = constant, E_0 = error term, and S_{PLOS} = PLOS score

Table 7 Results from FLR analysis

| Parameters | Coefficient values | | | | t value |
|----------------------|--------------------|----------|----------|-------------------|-----------|
| | Lower | Medium | Upper | Defuzzified value | |
| X_1 | 0.0198 | 0.0209 | 0.0153 | 0.0187 | 38.6202 |
| X_2 | 0.0166 | 0.0294 | 0.0130 | 0.0196 | 15.1964 |
| X_3 | 3.6442 | 3.8370 | 2.5208 | 3.3340 | 4.1167 |
| X_4 | - 1.1661 | - 1.3094 | - 0.0351 | - 0.8369 | - 39.6955 |
| Constant (X_0) | 7.7315 | 8.2703 | 7.3669 | 7.7895 | - |
| Error term (E_0) | 0.01 | 0.45 | 0.07 | 0.18 | - |

preliminary structure of the proposed PLOS model is represented in the following mathematical expression:

$$S_{\text{PLOS}} = A_1X_1 + A_2X_2 + A_3X_3 + A_4X_4 + X_0 + E_0, \quad (12)$$

and the terms in the expression are described in Table 6. The traffic volume per crosswalk per cycle was counted from video, and the median width was measured directly from the field using measuring wheels. Values of the pedestrian delay and the probability of pedestrian interaction with vehicles in crosswalk are calculated using the equations developed by the authors [47–49], as shown in Table 6.

The four-step FLR technique was applied to the mathematical model using the MATLAB R2014 to find the estimate of the independent variables. Table 7 shows the estimated values of independent variables from FLR

analysis. The fuzzy regression model for the PLOS constructed from the four-step method is

$$S_{\text{PLOS}} = 0.0187X_1 + 0.0196X_2 + 3.3340X_3 - 0.8369X_4 + 7.7895 + 0.18. \quad (13)$$

The PLOS score (S_{PLOS}) was predicted for all input observations using the developed PLOS model from the FLR technique. The predicted and observed PLOS scores were compared using Origin Pro 9.1 and the root mean square error value (1.0905) shows that the developed PLOS model results are close to the observed values. The R -square value (0.6117) indicates that 61.17% of the variation in the predicted PLOS score has been explained by explanatory variables and the predicted model fairly fits with the observed values. The t values for independent variables were estimated and the results presented in

Table 7. Comparison of estimated t values with t table values indicates that all variables are significant at 95% confidence interval. Therefore, the model results are useful in estimating the pedestrian perceived level of service score with respect to safety, convenience and efficiency at signalized intersections.

5.3 Proposed PLOS threshold values for LOS classification

According to the HCM 2010, pedestrian conditions in a signalized intersection are categorized by six LOSs: $A-F$, where A represents the best operational conditions and F the worst operational conditions. A minimum of 4 and a maximum of 15 for the predicted combined PLOS score were used in this analysis. The clustering technique was performed in MATLAB R2014 using FCM algorithm. In FCM clustering, the objective function was minimized to obtain the best threshold values. The optimization is stopped when the improvement in objective function between consecutive iterations is less than 0.001. Figure 4 displays the clustering results of the input data and Table 8 shows the proposed threshold value for each PLOS category.

In Table 8, the conditions for pedestrian crossing a crosswalk have been classified into A (best possible) to F (worst possible) and rated with respect to safety, convenience, and comfort level. Based on pedestrian perception survey, the percentages of PLOS $A-F$ were 2.75, 16.75, 28.75, 32.875, 7.625, and 11.25, respectively.

The best threshold values for a fixed number of clusters and the parameterized cluster shapes were regularly identified using the clustering technique. While performing clustering technique, fitness function was plotted using the identified threshold values in the end of each iteration process. The best fitness function was identified based on

Table 8 PLOS rating

| PLOS rating | PLOS score |
|-------------|---------------|
| A | < 6.25 |
| B | $6.25-7.25$ |
| C | $7.25-8.55$ |
| D | $8.55-10.55$ |
| E | $10.55-11.55$ |
| F | ≥ 11.55 |

statistical performance evaluation parameters such as MAPE and RMSE. Then, a box plot was used to test the threshold values and to verify which of the PLOS scores falls under the same category. The box plot was made using the predicted and observed PLOS scores. In order to verify the correctness of the obtained threshold values, the threshold values with respect to the predicted PLOS score from FCM and the threshold ranges with respect to the observed PLOS score from the box plot were compared.

For an observed PLOS range from 6 to 7, the predicted PLOS score is 6.25 using the FCM technique and the observed threshold value is 6.5 for the field data obtained using the box plot technique. Thus, the predicted value lies in threshold 1. The observed threshold values closely match with the FCM clustering predicted threshold values. Likewise, other threshold values for the predicted PLOS scores were compared with threshold ranges of the observed PLOS scores. All of the predicted values from the FLR technique were found to lie between the observed ranges found using the box plot technique. Further, the error values between observed and predicted threshold values were calculated for FCM (0.65) clustering techniques and are found to be small. The results infer that a reasonable good classification has been found for PLOS categories at signalized intersections using the FLR technique.

5.4 Validation

The validation of the developed model and the defined threshold values were carried out with data collected at a new crosswalk in the Santacruz Junction, Mumbai, India. Video-graphic survey and user perception survey were also conducted at this location, and a total of 66 pedestrians were interviewed with the same user perception survey form immediately after crossing the crosswalk. From the response of the 66 respondents, the observed pedestrians perceived LOS score was 9.76 and the observed LOS rating with respect to the observed LOS score was D . From the field survey, the observed median width and traffic volume per crosswalk per cycle at the selected crosswalk were 0.7 m and 41 veh/crosswalk/cycle, respectively. The

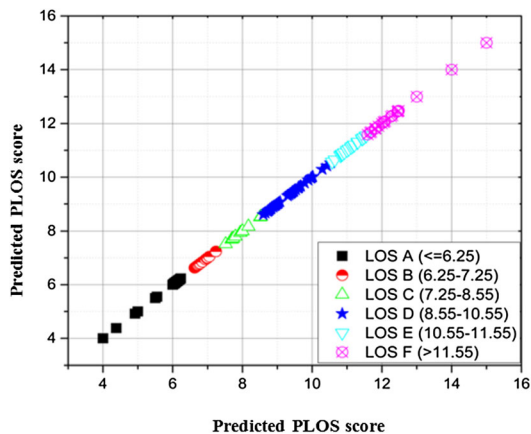


Fig. 4 Results of PLOS ranges from FCM analysis

estimated pedestrian delay and probability of pedestrian interaction with vehicle in crosswalk were 40.73 s/ped and 0.261, calculated using expressions given in Table 6. The predicted PLOS score and rating were 9.82 and PLOS *D*, respectively, obtained from Eq. (9) and Table 7. The percentage difference between the observed PLOS score and the predicted PLOS score was 0.62% and the error shows that the predicted value was close to the field-observed value. PLOS categories for predicted and observed scores were also identical at PLOS *D*. The minimum difference between the observed and predicted PLOS scores proves that the proposed PLOS models have high accuracy in estimation of the perceived PLOS score with respect to safety, convenience, and efficiency level at signalized intersections.

Furthermore, the individual pedestrian perceived LOS categories were compared with their predicted counterparts. Statistical performance tests were performed, and the results were compared using statistical software Origin Pro 9.1. The mean absolute percentage error (MAPE), root mean square error (RMSE), and rate of success prediction were compared between the observed values and predicted values and the results were 12.57%, 1.463%, and 56%, respectively. The performance results of the proposed threshold for PLOS category imply more precise and reliable solutions.

5.5 Sensitivity analysis

A sensitivity analysis is a technique used to determine how different values of independent variables impact a particular dependent variable under a given set of assumptions. It is used to test the presence of uncertainty and model simplification. In the sensitivity analysis, a special type of bar chart, such as tornado diagram, is used to analyze the impact of variables and the chart is categorized vertically and ordered. In this study, a sensitivity analysis was

performed between PLOS model dependent variable and independent variables. From the tornado diagram, the importance of each variable was examined and the impacts of independent variables were ranked with respect to the dependent variable. The data were taken from Santacruz Junction, and values were considered as baseline values for sensitivity analysis. Independent variables such as traffic volume along with pedestrian crossing, pedestrian delay, probability of pedestrian interaction with vehicle, and median width were changed by $\pm 5\%$, 10% , 15% , and 20% , and the effect of these changes on the dependent variable (PLOS score) are given in Table 9. Tornado diagram is plotted for understanding the positive and negative changes of 20% from base values in independent variables and the result is shown in Fig. 5.

Figure 5 shows that if traffic volume (X_1) is reduced by 20%, then there is a huge reduction in PLOS score; and if increased by 20%, a more increase in PLOS score is witnessed. Likewise, other variables such as pedestrian delay (X_2) and the probability of pedestrian interaction with vehicle in crosswalk (X_3) have the same effect on PLOS score. The median width (X_4) has reversed effect on PLOS score but the effect is comparatively lesser. From the results variables are prioritized and the ranks between the variables are rated as follows:

Severity rank 1: Variable X_3 —probability of pedestrian–vehicle interaction.

Severity rank 2: Variable X_2 —pedestrian delay.

Severity rank 3: Variable X_1 —traffic volume.

Severity rank 4: Variable X_4 —median width.

Finally, PLOS can be improved by changing the most significant variables at Santacruz Junction. From the ranking results, it is observed that variable X_3 has more impact than other variables on PLOS score. Therefore, the first remedial measure to improve the PLOS is to reduce

Table 9 Sensitivity analysis of PLOS variables

| Changes in variables (%) | X_1 | PLOS score | X_2 | PLOS score | X_3 | PLOS score | X_4 | PLOS score |
|--------------------------|-------|------------|-------|------------|-------|------------|-------|------------|
| 80 | 32.80 | 9.67 | 32.59 | 9.67 | 0.210 | 9.65 | 0.56 | 9.94 |
| 85 | 34.85 | 9.71 | 34.63 | 9.71 | 0.226 | 9.69 | 0.59 | 9.91 |
| 90 | 36.90 | 9.75 | 36.67 | 9.75 | 0.237 | 9.74 | 0.63 | 9.88 |
| 95 | 38.95 | 9.79 | 38.70 | 9.79 | 0.249 | 9.78 | 0.67 | 9.86 |
| 100 ^a | 41.00 | 9.83 | 40.74 | 9.83 | 0.263 | 9.83 | 0.70 | 9.83 |
| 105 | 43.05 | 9.86 | 42.78 | 9.87 | 0.276 | 9.87 | 0.74 | 9.80 |
| 110 | 45.10 | 9.90 | 44.81 | 9.91 | 0.289 | 9.91 | 0.77 | 9.77 |
| 115 | 47.15 | 9.94 | 46.85 | 9.95 | 0.303 | 9.96 | 0.81 | 9.74 |
| 120 | 49.20 | 9.98 | 48.89 | 9.99 | 0.316 | 10.00 | 0.84 | 9.71 |

^aRepresents existing baseline values for Santacruz Junction (existing conditions); X_1 , X_2 , X_3 and X_4 represent traffic volume, pedestrian delay, the probability of pedestrian–vehicle interaction and median width, respectively

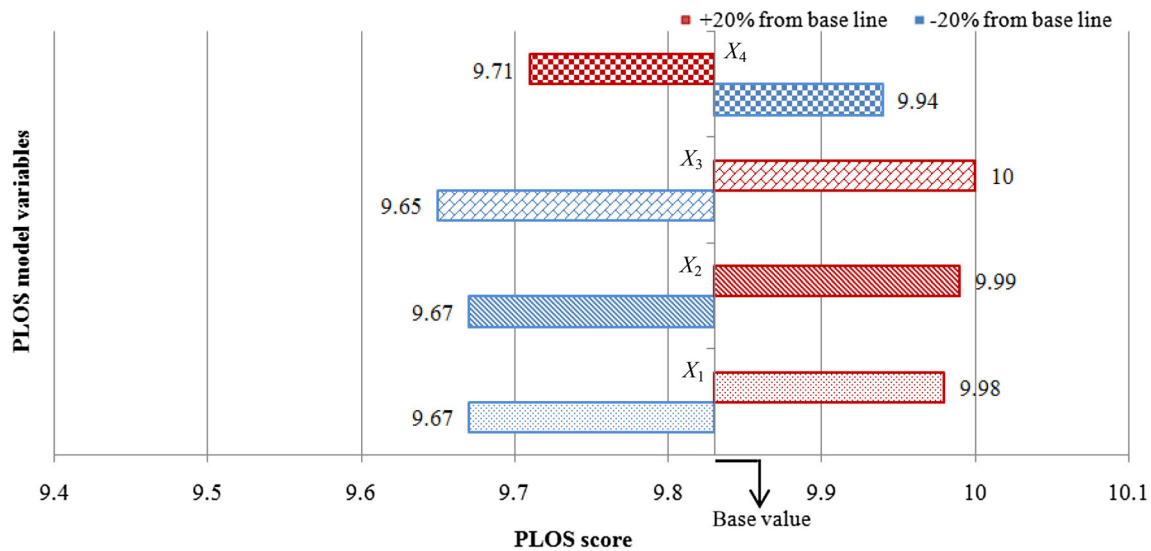


Fig. 5 Tornado diagram for sensitivity analysis of PLOS model variables

the impact of variable X_3 . Likewise, all variables can be improved as per the ranking order.

6 Conclusion

This study has identified several factors which significantly affect pedestrians' perception of safety, convenience, and efficiency while crossing signalized intersections. To provide a safe environment for pedestrians, proper methodologies that support transport planners and decision makers to find and evaluate the elements to improve walking are needed. In India, most signalized intersections are operated by sharing the pedestrian signal phase, which allows pedestrian crossing and parallel through and turning vehicle movement simultaneously. Earlier studies showed that more than 60% of pedestrians do not comply with signal phase due to impatience and experience interaction with vehicles in crosswalk. The existing condition of pedestrian facilities has been measured using LOS categories. This study has identified that pedestrian delay, traffic volume, probability of pedestrian interaction with vehicle, and median width significantly affect PLOS while pedestrians are crossing signalized intersections under mixed traffic conditions. The developed PLOS model accurately predicts pedestrians' perceptions of crossing at signalized intersections. The field data consist of qualitative data (pedestrian perception score with respect to safety, convenience, and efficiency) and quantitative data (pedestrian, traffic, and geometric characteristics) for the selected signalized intersections. To overcome the limitations of the CLR techniques, the FLR technique is used to develop a PLOS model that fits in mixed traffic conditions and is found able

to predict accurate PLOS score. In this study, we defined threshold values for six levels of PLOS ratings using fuzzy data mining techniques.

The study shows that the pedestrian delay has a greater effect on PLOS than other variables and an increase in pedestrian delay led to a decrease in pedestrian perceived efficiency and convenience. It is recommended to reduce the pedestrian delay by shortening cycle length and to increase pedestrian flow by providing more pedestrian green phase for improving PLOS at signalized intersections. Prohibiting free left turning and pedestrian non-compliance behavior, and providing exclusive free left turn lane can reduce interaction between pedestrians and vehicles in crosswalk. Meanwhile providing signal control scheme to separate pedestrians from vehicles, proper crosswalk markings, and placing adequate space between crosswalk and vehicle stop line can also be considered to increase pedestrian perceived safety and convenience. Besides, providing adequate median width for accommodating pedestrians while waiting to cross the crosswalk can improve pedestrian convenience level and increase PLOS at signalized intersections.

Most of the existing Indian studies and Indo-HCM have developed qualitative or quantitative measurement-based PLOS models and a combined qualitative and quantitative PLOS model for Indian traffic conditions is still missing in existing works. The proposed PLOS model has been developed by considering the pedestrian behavioral characteristics, influencing variables in mixed traffic conditions, and combination of qualitative and quantitative data. Therefore, the accuracy of the proposed PLOS model is better than the existing PLOS models. In this study, the FLR technique has been used for the first time to develop

the PLOS model. This method is effective in dealing with qualitative input data and has provided more accurate results when compared with CLR techniques. Apart from this, we have also used a novel approach of combining qualitative and quantitative data in the model development, which can provide new directions for planners and designers to improve existing conditions of intersections. Furthermore, the fuzzy clustering technique has been applied for the first time to identify the threshold values of PLOS classification at signalized intersections under Indian conditions. The developed model and proposed threshold values are useful to transport designers for understanding about the convenience of a specific intersection and its ability to accommodate pedestrian movement. It also helps in assessing and prioritizing the requirements of pedestrians at existing intersections.

Nevertheless, the developed model still has a few limitations to be addressed in the future. (1) In this study, triangular membership was utilized to develop the PLOS model. Other membership shapes have not been attempted and can be tried to identify the variation in FLR model and model consistency. (2) The developed model is specific in nature and works well for high flow conditions but requires additional calibrations for locations with lesser flow conditions. (3) Euclidian distance method was used for identifying the distance between the variables while performing FCM clustering. Application of other methods such as the nearest neighbor or single linkage can be considered for future scope of this study. (4) For field application purpose, we have simplified the developed model by establishing various relationships to estimate the independent variables of the PLOS model, for example, establishing the relation between suitable gap and the probability of pedestrian–vehicle interaction. In addition, the comparisons of pedestrian perceptions related to complaints and non-compliants can be added into the future research.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- (2014) Accidental deaths and suicides in India. Natioan Crime Records Bureau, Ministry of Home Affairs, India. Report
- (2011) Road accidents in India. Government of India, Ministry of Road Transport and Highways (MoRTH), India. Report
- Landis B, Vattikuti V, Brannick M (1997) Real-time human perceptions: toward a bicycle level of service. *Transp Res Rec J Transp Res Board* 1578:119–126. <https://doi.org/10.3141/1578-15>
- Polus BA, Ushpiz A, Division UT (1983) Pedestrian flow and level of service. *J Transp Eng* 109:46–56
- Mori M, Tsukaguchi H (1987) A new method to evaluation of level of serice in pedestrain facilities. *Transp Res Part A* 21:223–234
- Dandan TAN, Wei W, Jian LU, Yang B (2007) Research on methods of assessing pedestrian level of service for sidewalk. *J Transp Syst Eng Inf Technol* 7:5–10
- Kadali BR, Vedagiri P (2015) Evaluation of pedestrian crosswalk level of service (LOS) in perspective of type of land-use. *Transp Res Part A Policy Pract* 73:113–124. <https://doi.org/10.1016/j.tra.2015.01.009>
- Ya W, Kefei Y, Chaowei Y (2007) Level of service standards for pedestrian facilities in shanghai metro stations. In: International conference on transportation engineering, pp 2072–2078
- Jaskiewicz F (1999) Pedestrian level of service based on trip quality. *TRB circular E-C019 Urban street symposium*, pp 1–14
- Asadi-Shekari Z, Moeinaddini M, Zaly Shah M (2013) Non-motorised level of service: addressing challenges in pedestrian and bicycle level of service. *Transp Rev* 33:166–194. <https://doi.org/10.1080/01441647.2013.775613>
- Transportation Research Board (2000) Highway Capacity Manual. National Research Council, Washington, DC
- Lee JYS, Goh PK, Lam WHK (2005) New level-of-service standard for signalized crosswalks with bi-directional pedestrian flows. *J Transp Eng* 131:957–960
- Hubbard S (2009) Right turns on green and pedestrian level of service: statistical assessment. *J Transp Eng* 135:153–159
- Kuanmin C, Xiao-qiang L, Hai J, Yang-dong Z (2010) Towards the pedestrian delay estimation at intersections under vehicular platoon caused conflicts. *Sci Res Essays* 5:941–947
- Zhang L, Prevedourous PD (2003) Signalized intersection LOS that accounts for safety risk. In: *TRB 2003 annual meeting CD-ROM*
- Pan F, Lu J, Xiang Q, Zhang G (2007) Safety level of service at signalized intersections. *Int Conf Transp Eng 2007*:1499–1504
- Nagraj R, Vedagiri P (2013) Modeling pedestrian delay and level of service at signalized intersections crosswalks under mixed traffic condition. *J Transp Res Board* 2394:70–76
- Ni Y, Ling Z, Li K (2013) A new evaluation method combining efficiency and safety: multimodal comprehensive level of service of signalized intersections. In: International conference on transportation information and safety, pp 1449–1457
- Song H, Yang X (2011) Analysis of level of safety service at T-signalized intersection of Beijing. In: International conference on transportation information and safety, pp 1019–1026
- Lam WH, Lee JY (2001) A study of level of service for signalized crosswalk in hongkong urban areas. *J East Asia Soc Transp Stud* 4:27–40
- Muraleetharan T, Adachi T, Hagiwara T, Kagaya S (2005) Method to determine pedestrian level-of-service for crosswalks at urban intersections. *J East Asia Soc Transp Stud* 6:127–136
- Bian Y, Ma J, Rong J et al (2009) Pedestrians' level of service at signalized intersections in China. *Transp Res Rec J Transp Res Board* 2114:83–89. <https://doi.org/10.3141/2114-10>
- (2013) Spatial variations in pedestrian and bicycle level of service (LOS) for infrastructure planning and resource allocation. In: *Second Conference on Green Streets, Highways, and Development*. ASCE
- Ye X, Chen J, Jiang G, Yan X (2015) Modeling pedestrian level of service at signalized intersection crosswalks under mixed traffic conditions. *Transp Res Rec J Transp Res Board* 2512:46–55. <https://doi.org/10.3141/2512-06>
- Transportation Research Board (2010) Highway Capacity Manual. National Research Council, Washington, DC

26. Das P, Parida M, Katiyar VK (2015) Analysis of interrelationship between pedestrian flow parameters using artificial neural network. *J Med Biol Eng* 35:298–309. <https://doi.org/10.1007/s40534-015-0088-9>
27. Bivina GR, Parida P, Advani M, Parida M (2018) Pedestrian level of service model for evaluating and improving sidewalks from various land uses. *Eur Transp Trasp Eur* 67:1–18
28. Das P, Parida M, Katiyar VK (2015) Characteristics of macro-level pedestrian movement for planning of pedestrian infrastructure. *Int J Earth Sci Eng* 8:904–908
29. Marisamynathan S, Lakshmi S (2016) Method to determine pedestrian level of service for sidewalks in Indian context. *Transp Lett* 7867:1–8. <https://doi.org/10.1080/19427867.2016.1264668>
30. Indian Road Congress (IRC 103) (2012) Guidelines for Pedestrian Facility. New Delhi
31. Indian Highway Capacity Manual (Indo-HCM) (2018) CSIR—Central Road Research Institute. New Delhi
32. Kao C, Chyu C-L (2003) Least-squares estimates in fuzzy regression analysis. *Eur J Oper Res* 148:426–435. [https://doi.org/10.1016/S0377-2217\(02\)00423-X](https://doi.org/10.1016/S0377-2217(02)00423-X)
33. Kim B, Bishu RR (1998) Evaluation of fuzzy linear regression models by comparing membership functions. *Fuzzy Sets Syst* 100:343–352
34. Azimi M, Zhang Y (2010) Categorizing freeway flow conditions by using clustering methods. *Transp Res Rec J Transp Res Board* 2173:105–114. <https://doi.org/10.3141/2173-13>
35. Fang FC, Elefteriadou L, Pecheux KK, Pietrucha MT (2004) Using fuzzy clustering of user perception to define levels of service at signalized intersections. *J Transp Eng* 129:657–663
36. Bhuyan PK, Rao KVK (2011) Defining level of service criteria of urban streets in Indian context. *Eur Transp* 49:38–52
37. Fang FC, Pecheux KK (2009) Fuzzy data mining approach for quantifying signalized intersection level of services based on user perceptions. *J Transp Eng* 135:349–358
38. Mohapatra SS, Bhuyan PK, Rao KVK (2012) Genetic algorithm fuzzy clustering using GPS data for defining level of service criteria of urban streets. *Eur Transp* 52:1–18
39. Chang YO, Ayyub BM (2001) Fuzzy regression methods—a comparative assessment. *Fuzzy Sets Syst* 119:187–203
40. Kao C, Chyu C (1987) A fuzzy linear regression model with better explanatory power. *Fuzzy Sets Syst* 126:401–409
41. Abdalla A, Buckley JJ (2007) Monte Carlo methods in fuzzy linear regression. *Soft Comput* 11:991–996. <https://doi.org/10.1007/s00500-006-0148-5>
42. Arabpour AR, Tata M (2008) Estimating the parameters of a fuzzy linear regression model. *Iran J Fuzzy Syst* 5:1–19
43. Wang H, Tsaor R (2000) Insight of a fuzzy regression model. *Fuzzy Sets Syst* 112:355–369
44. Hassanpour H, Maleki HR, Yaghoobi MA (2009) A note on evaluation of fuzzy linear regression models by comparing membership functions. *Iran J Fuzzy Syst* 6:1–6
45. Teodorovic Dusan (1999) Fuzzy logic systems for transportation engineering: the state of the art. *Transp Res Part A* 33:337–364
46. Zhang L, Prevedouros PD (2011) User perceptions of signalised intersection level of service using fuzzy logic. *Transportmetrica* 7:279–296. <https://doi.org/10.1080/18128601003667460>
47. Marisamynathan S, Vedagiri P (2015) A statistical analysis of pedestrian behaviour at signalized intersections. *Eur Transp* 57:1–18
48. Marisamynathan S, Vedagiri P (2014) modeling pedestrian delay at signalized intersections under mixed traffic conditions. Transportation research board 93rd annual meeting CD ROM. Washington, DC
49. Marisamynathan S, Vedagiri P (2016) A new approach to estimate pedestrian delay at signalized intersections. *Transport* 33(1):249–259. <https://doi.org/10.3846/16484142.2016.1158208>