

Abstract number: 113

Level of Service for Bicycle Through Movement at Signalized Intersections Under Heterogeneous Traffic Flow Conditions

Sambit Kumar Beura¹, N Kiran Kumar², P. K. Bhuyan³

¹ Research Scholar (Ph.D.), Department of Civil Engineering, National Institute of Technology Rourkela, Odisha, India 769008, Email: sambit.beura@gmail.com

² Former M. Tech. Student, Department of Civil Engineering, National Institute of Technology Rourkela, Odisha, India 769008, Email: kirannitturu@gmail.com

³ Assistant Professor, Department of Civil Engineering, National Institute of Technology Rourkela, Odisha, India 769008, Email: pkbtrans@gmail.com

Abstract:

Service qualities offered to the bicycle through movement at signalized intersections under heterogeneous traffic conditions have been modeled using attributes collected from 35 intersection approaches. Perceived satisfactions of 160 users were assessed and seven major attributes contributing to bicycle services were identified. A highly reliable model ($R^2 = 0.86$) was developed using step-wise regression analysis. Sensitivity analysis carried out on modelled variables reported that, main-street traffic volume (38%), width of main-street (17.6%), pavement conditions (15.9%) and average stopped time delay (15.2%) are by far the most important variables. Affinity Propagation clustering was used to define ranges of service categories (A-F).

Keywords: Signalized Intersection; Heterogeneous Traffic Flow; Bicycle Level of Service; Perception Survey; Step-wise Regression; Sensitivity analysis; Affinity Propagation Clustering.

1. Introduction

No fuel requirement, easy ridership, door-to-door mobility and several health related benefits have made bicycle mode to essentially contribute a significant percentage in total trips made around the world. Mass use of bicycles partly overcome several issues like traffic congestion, accident rate, shortage in petroleum fuel and air pollution level. Hence, transportation administrators are becoming more conscious these days to design bicycle friendly road networks to encourage bicycle use. However, in a developing country like India over 52 million people in urban areas are using this green mode of transportation with lots of difficulties due to absence of proper bicycle facilities. Transportation planners and engineers have not been provided with a suitable method to systematically assess the bicycle service qualities offered by existing road facilities under the influence of heterogeneous traffic flow conditions. Development of a bicycle model is highly essential which would help the administrative personnel to assess service qualities of present road infrastructure and subsequently recognize the degree of needs for improved facilities. Among all components of road networks, intersection is one of the most complex features of the transportation system and a highly formidable portion for the road users. Hence, this study has focused on analyzing the prevailing road intersection conditions in Indian mid-sized cities (population size 0.5-1 million) from a bicyclist's perspective and subsequently develop a highly reliable service prediction bicycle model. This type of model is commonly called as "Bicycle Level of Service" (BLOS) model as it is used to assess the quality of service or level of service (LOS) offered by a facility for bicycle use.

Data on traffic flow characteristics, geometric details and surrounding environments have been collected from 35 intersection approaches of four Indian mid-sized cities and are analyzed with respect to their goodness in satisfying bicycle users. Major aim of this study is to identify important intersections attributes contributing to bicycle service qualities at signalized intersection approaches and interlink them with perceived satisfaction levels of bicyclists. In this regard, satisfaction levels of approximately 160 users on studied intersection approaches have been assessed quantitatively through a perception survey. Six important attributes have been identified from the preliminary data analysis carried out using Pearson's correlation analysis. A relationship has been established among these important variables and perceived satisfactions of survey participants through a step-wise regression analysis. The prediction performance of the model was assessed in terms of several statistical parameters such as coefficient of determination

(R^2) and Nash–Sutcliffe model efficiency coefficient (E) etc. The developed BLOS model is highly consistent in its universal application and has a high R^2 of 0.86 with training data sets. The model has also been compared with other existing models to justify its distinctiveness.

2. Review of Literatures

Though modelling of bicycle service qualities at signalized intersections has not been explored under the influence of heterogeneous traffic flow conditions, several researchers have attempted the same under the influence of homogeneous traffic flow conditions in developed countries. Researchers have identified important attributes under prevailing conditions and have developed BLOS models mostly using regression analysis. Davis (1987) proposed the initial mathematical model named as, Bicycle Safety Index Rating (BSIR) model for assessing the bicycle service qualities offered by urban streets. This model comprises of two sub-models, Roadway Segment Index (RSI) and Intersection Evaluation Index (IEI) model. IEI model was proposed exclusively for determining BLOS at intersections which is a function of traffic volume, type of signalization, and some geometric factors such as: presence of left-turn lane and right-turn lane, number of through lanes, curb radii, and sight distance. Botma (1995) proposed methods to determine the BLOS for bicycle paths and bicycle–pedestrian shared paths. Both methods defined BLOS in terms of events: an event occurs when one user passes another user travelling in the same direction, or when one user encounters another user travelling from reverse direction. When events become more frequent, the LOS starts degrading from excellent towards worst. 2000 version of Highway Capacity Manual (HCM) adopted this methodology for exclusive and shared paths. Method proposed for assessing BLOS of on-street bicycle lanes is based on the number of events that varies based on variations in bicycle flow rate, average speed, and standard deviation of the speed. Method proposed for assessing BLOS at signalized intersection used controlled delay (s/bicycle) as the measure of effectiveness. Controlled delay estimated using uniform delay assumed there is no overflow delay at intersections.

Crider (2001) identified importance of conflicts, exposure and delay experienced by bicyclists for modeling of the bicycle through moment at signalized intersections. Landis *et al.* (2003) proposed a BLOS model for bicycle through movement at signalized intersection. Required qualitative data were obtained from perceptions of bicyclists who rode through studied intersections. The authors found that traffic volume, outermost through lane width, and the intersection crossing distance are the primary factors influencing the riding quality of bicyclists. Presence of a bicycle lane, paved shoulder, and curb radii were found to be statistically insignificant in the model estimation. Steinman *et al.* (2004) proposed a method to evaluate major design features contributing to bicyclists' crossings at a signalized intersection. The key features identified in this aspect are intersection crossing distance, roadway space allocation (i.e., crosswalks and bicycle lanes), corner radius dimension, and traffic signal characteristics. This analytical tool can be used to evaluate and improve BLOS at intersections by modifying the design and operational features. Carter *et al.* (2007) proposed Bicycle Intersection Safety Index (Bike ISI) model which can be used to prioritize intersection legs based on bicyclists' feelings of safety. Model proposed for bicycle through moment included several parameters such as: main-street and cross-street traffic volume, main-street speed limit, presence of turning vehicular traffic across the path of through bicyclists, number of right turn traffic lanes, presence of bicycle lane, signalization at intersection (yes or no) and on-street parking activities on main-street approach.

Rubins and Handy (2005) found that, bicycle crossing times at signalized intersections vary widely for each crossing distance, and longer crossing times should be provided to be ensured that 98%, or even 85%, of bicyclists will be able to safely cross an intersection. Shladover *et al.* (2009) used bicyclists' trajectories data to yield cumulative distributions of the crossing speeds of bicyclists who did not have to stop at the intersection, and the start-up times as well as final crossing speeds of the bicyclists who had to cross from a standing start. These data gave the timing information relative to the traffic signal which was then used to define recommended signal times to permit most bicyclists to safely cross wide arterial intersections. BLOS criteria proposed in HCM (2010) for signalized intersections considers several measures such as; width of outermost lane (or bicycle lane if present), traffic volume in the outermost lane and width of cross-street. Jensen (2013) proposed some models for quantifying bicyclists' satisfactions on roundabouts, signalized and un-signalized intersections and other crossings. The author identified the importance of bicycle facilities, roadway width, length of crossing, size of roundabout, traffic volume, waiting time and speed limit in BLOS at these facilities. Strauss *et al.* (2013) proposed a Bayesian modelling approach to study bicycle activity at signalized intersections. The authors concluded that, crosswalk length and presence of bus stops increase bicyclist injury occurrence and provision of a raised median decreases the same. Bicycle activities at intersections were found to increase with increase in employment, number of metro stations, commercial activities, length of bicycle facilities and the presence of schools within 50–800-meter distance of the intersection area.

After a thorough investigation on models developed for through moment at signalized intersections, several intersection attributes having potential effect on BLOS are summarized as: traffic volume on main-street and cross-street approach, number of through lanes, presence of left-turn and right-turn lane, curb radii, sight distance, delay, conflicts, exposure, outermost through lane width and the intersection crossing distance etc. However, these variables have different effect on bicycle service quality estimation under heterogeneous traffic flow conditions and differing riding behavior of the users. Hence, this study has examined a wide range of intersection attributes and identified the significant ones affecting the bicycle service qualities. Using those identified variables, a highly reliable BLOS model has been developed. Following section briefly discusses about the methodology followed in this study.

3. Study Methodology

This study has used step-wise regression analysis for development of the service prediction BLOS model. Ranges of service categories are defined using Affinity Propagation (AP) clustering technique. The underlying principles of these two techniques are briefly described in this section.

3.1 Step-wise regression analysis

Models developed using step-wise regression analysis being more intuitive in practice, this technique has found its prominent use in developing service prediction models in the field of traffic engineering (such as: bicycle and pedestrian models). The model was primarily developed in following three steps:

- Intersection attributes significantly ($p < 0.001$) affecting the bicycle through movement at signalized intersections were identified
- Several combined or transformed forms of these selected variables were developed and employed in the step-wise regression analysis to test the significance of their coefficients in developed models,
- Most suitable BLOS model satisfying several significance criteria was selected.

Equation 1 represents the mathematical expression for the relationship among the output and input variables (factors significantly affecting the bicycle through movement at signalized intersections).

$$BLOS_{pred} = a_1f(X_1) + a_2f(X_2) + a_3f(X_3) + \dots + a_nf(X_n) \quad (1)$$

Where, $BLOS_{pred}$ is the predicted BLOS score (output variable); $X_1, X_2, X_3, \dots, X_n$ are model input variables those can be an individual variable, combination of two or more variables, or transformed form of variables; and $a_1, a_2, a_3, \dots, a_n$ are the estimated coefficients of input variables.

3.2 Affinity Propagation (AP) clustering

AP is one of the recently introduced theoretic clustering techniques that simultaneously considers all data points as possible center point of the cluster, called as “exemplar” (Dueck and Frey 2007). Each message is sent to reflect the latest interest which is possessed by each data point, to be able to select another data points as their exemplar. Procedurally, this algorithm is operated using three matrices: one similarity matrix (s), one responsibility matrix (r), and one availability matrix (a). Outcomes are stored in a criterion matrix (c). This technique works in following steps in which these matrices are iteratively updated.

- Input similarity matrix $s(i,k)$: the similarity of point i to point k , where i and k respectively refer to the rows and columns of the associated matrix;
- Initialize the availabilities $a(i,k)$ to zero: $a(i,k)=0$;
- Update all responsibilities $r(i,k)$: $r(i,k) \leftarrow s(i,k) - \max_{k' \neq k} [a(i,k') + s(i,k)]$ (2)
- Update all availabilities $a(i,k)$: $a(i,k) \leftarrow \min \left\{ 0, r(k,k) + \sum_{i' \in \{i,k\}} \max \{ 0, r(i',k) \} \right\}$ for $k \neq i$ (3)
- Availabilities and responsibilities matrices are added to monitor the exemplar decisions. For a particular data point i ; $a(i,k) + a(i,k) > 0$ criteria is used to identify exemplars;
- If the decisions made in step-3 do not change for a certain numbers of iterations or if a fixed number of iterations is reached, then go to step-5; otherwise, go to step-1;

7. Assign other data points to the exemplars using the nearest assign rule (i.e. assign each data point to an exemplar to which it is highly alike).

3.3 Cluster validation measures

A cluster validation measure determines the suitable numbers of clusters for a set of data points and tests the quality of cluster results obtained. It evaluates and compares the partitions resulted from different algorithms or resulted from the same algorithm under different parameters. However, several measures proposed by past researchers are not very much accurate in their applications. Thus in present study, four indices discussed below are used altogether to validate clustering results.

3.3.1 Silhouette Index (SI)

Silhouette Index (SI) used to assess cluster results, is a composite index that reflects the compactness and separations of the groups (Rousseeuw 1987). SI for each data point i , designated as $SI(i)$, is calculated using Equation 4. The average value of all $SI(i)$ corresponding to all data points replicates the quality of groups obtained in the cluster analysis. A larger SI value indicates a better clustering result.

$$SI(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (4)$$

Where: $a(i)$ = average distance of a data point i to other data points in the same cluster, $b(i)$ = average distance of this data point i to other data points belonging to the nearest cluster.

3.3.2 Davies-Bouldin Index (DBI)

Davies-Bouldin Index (DBI) defined in Equation 5 (Davies and Bouldin, 1979) is a function of the ratio of the sum of intra-cluster scatter to inter-cluster separation. This ratio is small when obtained clusters are compacted and are far from each other. Hence, a smaller DBI value indicates better clustering result.

$$DBI = \frac{1}{c} \sum_{i,j=1}^c \max_{i \neq j} \left[\frac{D_c(i) + D_c(j)}{d_{ce}(i, j)} \right] \quad (5)$$

Where, $D_c = \sum_i \|X_a - C_i\| / N_i$ = intra-cluster distance calculated by using average of pair-wise distances from data points in the cluster to the cluster centroid; X_a = an arbitrary data point belonging to cluster i ; N_i = numbers of data points belonging to cluster i ; $d_{ce} = \|c_i - c_j\|$ = the inter-cluster distance or distance between centroids of two clusters; C_i = center of cluster i ; and C_j = center of cluster j .

3.3.3 Calinski-Harabasz Index (CHI)

Calinski-Harabasz Index (CHI) defined in Equation 6 (Caliński, Harabasz 1974) uses the quotient between the intra-cluster average squared distance and inter-cluster average squared distance. It is one of the best global statistical criteria used for cluster valuation (Milligan and Cooper 1985). A higher value of CHI indicates most separated clusters and the optimum number of clusters.

$$F = \frac{\left[\sum_{i=1}^n (X_i - \bar{X})^2 - \sum_k^c \sum_{i=1}^{n_k} (X_i - \bar{X}_k)^2 \right] / c - 1}{\left[\sum_k^c \sum_{i=1}^{n_k} (X_i - \bar{X}_k)^2 \right] / n - c} \quad (6)$$

Where: n = is total number of input data points; c = number of clusters; n_k = number of data points in cluster k ($k=1,2,\dots,c$); X_i = observation vectors for input data i ; and \bar{X}_k = centroid for cluster k .

3.3.4 Dunn Index (DI)

Dunn Index (DI) expressed in Equation 7 (Dunn 1974) is used to assess the quality of clusters obtained. A larger value of DI indicates the existence of well compacted and separated clusters.

$$D_{nc} = \min_{i=1, \dots, nc} \left\{ \min_{j=i+1, \dots, nc} \left(\frac{d(c_i, c_j)}{\max_{k=1, \dots, nc} (dia(c_k))} \right) \right\} \quad (7)$$

Where, $d(c_i, c_j) = \min_{x \in C_i, y \in C_j} d(x, y)$ = dissimilarity function between clusters c_i and c_j ; and $dia(C) = \max_{x, y \in C} d(x, y)$ = diameter of a cluster.

4. Site Selection

For developing an intersection BLOS model suitable for heterogeneous traffic flow conditions, the prime requirement is to include diversified intersection attributes collected from a sets of diversified conditions. Hence data has been collected from 35 signalized intersection approaches that well explains the desired variability in the operational characteristics of signalized intersections. Studied intersections are located in four Indian mid-sized cities namely; Bhubaneswar, Tirupati, Kurnool and Anantapur. Locations of these cities are shown in Figure 1. Bhubaneswar, the capital of Odisha State, is a center of economic and religious importance located in eastern part of the country. Tirupati is the seventh and Kurnool is the fifth most populous city of Andhra Pradesh State. Anantapur is another mid-sized city located in this Andhra Pradesh State, which is well connected to neighboring major cities with National Highways.

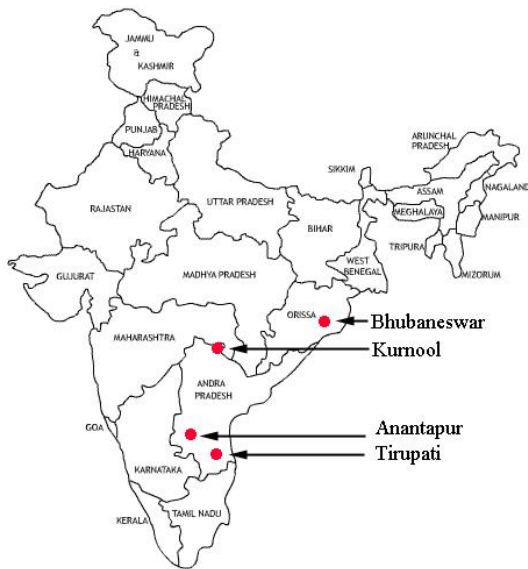


Figure 1: Location of studied cities in India.

Studied intersection approaches vary from one another with respect to geometric details, traffic composition and surrounding environments. Few intersections are located in residential areas, whereas others are located in semi-commercial and commercial areas. At these studied intersections, number of traffic lanes vary up to 4 lanes per direction on main-street approaches and up to 3 lanes per direction on cross-streets. Traffic volume on main-street varies from 200 to 3500 PCUs/h during peak hours. However, cross-streets carry relatively lower volume of traffic that varies from 150 to 2400 PCUs/h during peak hours. The average traffic speed on intersection approaches varies from 22 to 46 Km/h during peak hours. Intersections located at starting or end points of any road corridors are often characterized by authorized stops of intermittent public transits such as pick-up vans and 3-wheeler autos. All intersections are somewhat influenced from parking activities on approaching streets. Intersection approaches often have provisions of medians, parking lanes, curbs and gutter pans.

5. Data Collection

This section discusses about the procedure followed for data collection in the field. Three types of data are collected in this study such as physical features, traffic parameters, and user satisfaction ratings those are discussed below.

5.1 Physical features

Physical features include data of geometric details of intersection approaches, traffic control and facilities provided for bicycle use. Geometric measurements included during inventory survey are width of main-street, cross-street, medians on main-street and cross-street, parking lane, sidewalk, curb and gutter. In general, following attributes were collected as those were recognized by researchers as having potential impacts on bicyclists' perceived satisfactions.

- Traffic control,
- Number of intersection legs,
- Traffic movement (one-way or two-way),
- Number of traffic lanes on main-street and cross-street,
- Number of traffic lanes for a bicyclist to cross while making a right turn,
- Presence (yes or no) and width (m) of shared-use path on main-street approach,
- Intersection crossing distance (m),
- Presence of median islands (yes or no),
- Curb radii (m),
- On-street parking turn-over on main-street approach (high, medium or low),
- Street lighting (good, moderate or poor),
- Surrounding development type (residential, office or commercial),
- Right-turn-on-red allowance (yes or no),
- Sight distance (good, moderate or poor),
- Volume of turning vehicles across the path of through bicyclists (high, medium or low).

5.2 Traffic parameters

Traffic parameters those might have potential influence on bicyclists' perceived satisfactions were included in this study for investigation; such as: traffic volume on main street, traffic volume on cross street, volume of pedestrians, average stopped time delay and average traffic speed on main street. This study has used peak hour traffic volume to reflect the worst conditions a bicyclist encounters while using an intersection. Traffic flow at studied intersections were videotaped during expected peak periods of traffic flow, i.e. either during morning 8.30-10.30 AM or evening 4.30-6.30 PM. Collected videos were played in the institute laboratory and team members were employed for vehicular and pedestrian volume counts. Using running average method peak one hour was determined and peak hour volume was calculated. Traffic volume was expressed in Passenger Car Units per hour (PCUs/h) using the PCU conversion values recommended by Indian Road Congress (IRC-SP-41, 1994) for signalized intersections under heterogeneous traffic flow conditions. Average traffic speed on intersection approaches were calculated from video clips collected from segment area over a trap length of 30 meter. Average time taken by motor-vehicles to cross the 30-meter trap on any approach was extracted with an accuracy of 0.1 second and the average traffic speed was calculated.

5.3 User satisfaction scores (perceived BLOS scores)

One major aim of this study is the prediction of overall satisfaction levels perceived by bicyclists while crossing through signalized intersections. Hence an attempt has been made to estimate satisfaction ratings for studied intersection approaches from user perceptions. A perception survey was conducted in which 160 road users were participated to share their opinion. Participants represented a fair cross-section of age, sex, educational qualification, driving experiences and geographic origin. One video clip from each intersection (approximately 45 seconds in length) that expressed the average traffic volume, speeds of vehicles and other surrounding environment was shown to the participants on a wide screen in NIT Rourkela auditorium. The volume of speakers was set to approximate sound of real traffic flow. By creating a simulated environment, each participant was made to developed a feeling as if riding bicycle on field. Using a 6-point Likert scale varying from 1 (influence of the attribute is very much positive on BLOS) to 6 (influence of the attribute is very much negative on BLOS), participants rated their perceived satisfaction levels on intersection attributes mentioned in a questionnaire. The overall rating perceived by an individual named as

“perceived BLOS score” for an intersection approach was calculated by taking the average of all ratings obtained for the attributes of that particular approach. Hence a total of 160 perceived BLOS scores were obtained per intersection approach. This survey resulted in a total of 5,600 (35 approach×160 persons) effective perceived BLOS scores with a mean of score of 3.59 and standard deviation of 0.79. For checking the data sufficiency, Cochran’s sample size formula (Cochran 1977) was used and the allowed error in estimation of the mean perceived score (3.59) was calculated. The error in using this amount of data set is limited within 1% as estimated at 95% confidence level. This minimal error shows the data sufficiency in numbers of perceived BLOS scores collected in this study.

6. Variables selection

Using the perceived BLOS scores obtained for studied intersection approaches as output variable, Pearson’s correlation analysis was carried out to find out the intersection attributes important for BLOS model estimation. Pavement condition index (*PCI*) at the intersection and five other variables of main-street such as: road width per direction (*RW*), peak hour traffic volume (*PHV*), land use pattern (*LU*), on-street parking turn-over (*P*) and average stopped time delay (*D*) were found out to be important ($p < 0.001$). Correlations among these selected variables are shown in Table 1. It was observed that, these variables are not highly correlated with each other as correlations (*R* values) among them are far below 0.9 which indicates high correlation among input variables. Hence selected variables are able to contribute independently and significantly in the model building.

Table 1: Correlations among variables used in model fitting

Parameter	PHV	RW	PCI	LU	P	D
PHV	1.000	0.364	-0.074	0.223	0.295	-0.339
RW	0.364	1.000	0.438	0.088	0.097	-0.036
PCI	-0.074	0.438	1.000	-0.477	-0.517	-0.296
LU	0.223	0.088	-0.477	1.000	0.765	0.389
P	0.295	0.097	-0.517	0.765	1.000	0.112
D	-0.339	-0.036	-0.296	0.389	0.112	1.000

7. Model development and statistical tests

Suitable transformations and/or combinations were applied on above selected variables and a step-wise multi-variable regression analysis was carried out to develop required BLOS model. Several models were rejected those failed to satisfy required significance criteria and a best fit model shown in Equation 8 was selected and proposed.

$$BLOS_{pred} = 1.687 + 0.699 \times \ln\left(\frac{PHV}{RW}\right) - 0.529 \times PCI + 0.340 \times D + 0.226 \times (1 + LU) \times (1 + P) \quad (8)$$

Following sub-sections discuss about the significance of model input variables, quality of model fitting, prediction performance of the developed model, model validation, comparison of the model with existing ones and sensitivity analysis carried out to identify relative importance of model inputs.

7.1 Significance of modelled parameters

Table 2 shows coefficients of modelled parameters estimated using step-wise regression analysis. Other, statistics of model coefficients and modelled variables are also shown in this table. It can be observed that, all coefficients are statistically significant at more than the 95% level ($p < 0.05$). Hence all input variables are significantly contributing to the model. Two statistical parameters such as Tolerance and Variance Inflation Factor (VIF) were used to check multicollinearity among modelled parameters. A tolerance value of less than 0.2 indicates a potential problem (Menard 1995) and if the largest value of VIF is greater than 10 then there is a cause for concern (Bowerman and O’Connell 1990). Table 2 shows that all tolerance values are far above 0.2 and VIF values are far below 10. Hence multicollinearity is not affecting the selected variables used in regression analysis. Residual terms for any two observations should be independent or uncorrelated which was examined by using Durbin-Watson (DW) test. This test statistic varies between 0 (high positive correlation) and 4 (high negative correlation) and indicates residuals are not correlated if its value is 2. DW value obtained in this study is 1.8 which is very close to 2. These all tests conclude that, selected variables are very much significant for BLOS model development and the regression analysis carried in this study is not biased by collinearity among variables.

Table 2: Model coefficients, statistics and significance

Variable	Coefficient	Standard Error	<i>t</i> -statistic	Significance (<i>p</i> values)	Collinearity Statistics	
					Tolerance	VIF
<i>ln</i> (PHV/RW)	0.699	0.124	5.618	< 0.001	0.853	1.172
PCI	-0.529	0.212	-2.493	0.022	0.551	1.813
D	0.340	0.147	2.318	0.031	0.701	1.428
(1+LU) ×(1+P)	0.226	0.064	3.542	0.002	0.684	1.462
Constant	1.687	0.727	2.319	0.030	-	-

7.2 Quality of model fitting

The BLOS model reported in this paper is highly consistent in its universal application and has a high coefficient of determination value (R^2) of 0.86 with training data sets. The value of F -ratio (29.97) found through Anova test is significant ($p < 0.001$), which indicates that the model was fitted significantly to the training data set.

7.3 Model validation and assessment of prediction performance

Validation of the model was done using data of 10 intersection approaches (approximately 30% of total) those were reserved for validation purpose and were not used in model building. A graph (shown in Figure 2) was plotted taking average user perceived BLOS scores obtained on each approach in X-axis and predicted BLOS scores in Y-axis. Trend line produced in this graph shows a high R^2 value of 0.92 and a tangent angle of 44.29° with X-axis which is very close to 45° . These observations conclude that; the model significantly satisfies required criteria of validation.

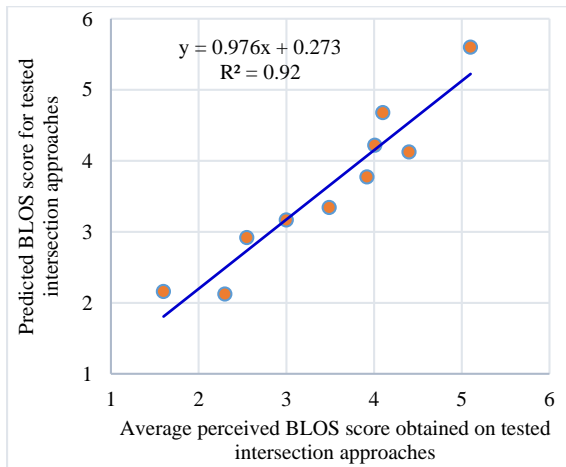


Figure 2: Validation of developed BLOS model.

Prediction performance of the developed model has been assessed in terms of several statistical parameters such as: coefficient of determination (R^2), Nash–Sutcliffe model efficiency coefficient (E), average absolute error (AAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and maximum absolute error (MAE). These parameters were calculated separately for training and validation data sets using perceived and predicted BLOS scores and results are shown in Table 3. As R^2 has been criticized by several researches as being a biased estimate, another parameter E defined in Equations 9a-9c (Nash and Sutcliffe 1970) was used in addition to test the prediction performance of the model. E compared perceived and predicted values of BLOS scores, and evaluated how good the BLOS model is able to explain total variance in the data.

$$E = \frac{E_1 - E_2}{E_1} \quad (9a)$$

$$E_1 = \sum_{i=1}^s \left(BLOS_{Perceived} - \overline{BLOS_{Perceived}} \right)^2 \quad (9b)$$

$$E_2 = \sum_{i=1}^s (BLOS_{Perceived} - BLOS_{Predicted})^2 \quad (9c)$$

Where, $BLOS_{Perceived}$ = perceived BLOS score; $\overline{BLOS_{Perceived}}$ = average of all perceived BLOS scores obtained for studied intersection approaches (3.59); $BLOS_{Predicted}$ = predicted BLOS score; and s = total number of intersection approaches used (i.e. 25 and 10 respectively in model training and model validation).

Table 3 shows that the model has got higher values of R^2 and E both with training data and validation data sets. Error measuring parameters such as: AAE, RMSE, MAPE, and MAE have got very lower values which indicate a strong prediction capability of the developed model.

Table 3: Prediction performance of developed intersection BLOS model

Data Set	R^2	E	RMSE	AAE	MAE	MAPE
Training	0.857	0.857	0.289	0.253	0.578	7.34%
Validation	0.919	0.878	0.354	0.313	0.578	10.61%

7.4 Comparison of developed BLOS model and existing models

Intersection Evaluation Index (IEI) model, and BLOS model proposed by Landis *et al.* (2003), Carter *et al.* (2007) model and HCM (2010) model for signalized intersections are generally followed in several countries. These models are primarily developed using data sets collected from homogeneous traffic flow conditions. Landis *et al.* (2003) has considered outside lane parameter and other models have considered number of through lanes as a major geometric factor of main approach. However, under heterogeneous traffic flow conditions road users hardly follow any lane discipline at intersections. Also, bicyclists often find gap to stand within motorized traffic stopped at intersections approach lanes. Hence the total width of the approach road was perceived to be an important geometric variable to the BLOS model. Hence the developed model considers total width and not number of lanes of the approach road. Existing models do not consider delay as a significant parameter. However, in present context average stopped time delay (s/bicycles) contributed significantly in the model development. Existing models have considered crossing distance and volume on cross-street as important variables. But in present context these variables had insignificant effect on BLOS compared to other important variables considered. Though ratio of green period to cycle length was expected to be an important attribute, but it was found to be highly correlated with average stopped time delay with correlation coefficient (R) of -0.78. These are some key factors which differentiate the BLOS model developed in this study under heterogeneous traffic flow conditions from existing models developed under homogeneous traffic flow conditions.

7.5 Sensitivity analysis

Sensitivity analysis was carried out using Equations 10a and 10b (Gandomi *et al.* 2003) to identify the relative importance of input variables used in the model building.

$$S_i = \frac{N_i}{\sum_{i=1}^n N_i} \times 100 \quad (10a)$$

$$N_i = f_{\max}(x_i) - f_{\min}(x_i) \quad (10b)$$

Where: S_i is the sensitivity of i^{th} variable calculated in percentage; N_i = difference between maximum value ' $f_{\max}(x_i)$ ' and minimum value ' $f_{\min}(x_i)$ ' of the predicted BLOS score over i^{th} input variable, those were calculated by putting the maximum and/or minimum values of i^{th} input variable and mean values of each remaining inputs in developed BLOS model ' $f(x)$ '; and n is the number of input variables.

S_i (%) values were calculated and all input variables were ranked in Table 4 to show their relative importance in the model fitting. It was observed that, traffic volume on main-street approach, width of main-street approach, pavement conditions and average stopped time delay at intersections are by far the most important variables for service quality estimation under heterogeneous traffic flow conditions.

Table 4: Relative importance of model inputs measured through sensitivity analysis

Parameter	PHV	RW	PCI	LU	P	D
S _i (%)	38.00	17.56	15.91	6.15	7.14	15.23
Rank	1	2	3	6	5	4

7.6 Stratification of BLOS scores into LOS categories

For stratifying predicted BLOS scores into LOS categories, the foremost requirement is determination of optimal numbers of clusters that is most suited for the data points (predicted BLOS scores) obtained in present context. After determining optimal numbers of clusters, the predicted scores were stratified into that number of LOS classes. The details of these activities are discussed below.

7.6.1 Determination of Optimal Number of Clusters

Predicted BLOS scores were analyzed using four cluster validation parameters discussed earlier such as: SI, DBI, CHI, and DI to determine suitable number of clusters that should be applied on the data points. As a thumb rule, it is always considered that lesser the numbers of clusters better is the classification results (when variation in validation parameters is marginal). Values of validation parameters are obtained for 2 to 7 numbers of clusters are shown in Table 5. Literature shows that, a larger SI value, a smaller DBI value, a larger CHI value and a larger DI value indicate a better quality of clustering result. Considering these criteria altogether, it can be determined from Table 5 that, for the data points obtained in present study, six numbers of clusters are most preferred. Hence the predicted BLOS score obtained for through movement of bicyclists of intersection approaches are clustered into six LOS classes (A-F) based on output of these validation measures; which has a resemblance to traditionally followed classes.

Table 5: Values of cluster validation parameters for different number of clusters (2-7)

Number of clusters	2	3	4	5	6	7
Silhouette index (SI)	0.533	0.584	0.590	0.609	0.613	0.586
Davies-Bouldin index (DBI)	0.654	0.469	0.401	0.411	0.389	0.416
Calinski-Harabasz index (CHI)	71.86	118.37	176.85	214.03	266.86	212.26
Dunn index (DI)	2.760	2.732	2.178	2.804	3.531	1.926

7.6.2 Ranges of LOS categories (A-F)

Affinity Propagation (AP) clustering technique was used to classify predicted BLOS scores into six levels of bicyclists' satisfaction (A-F), where 'A' designates excellent service quality and 'F' designates the worst. Figure 3 shows the results obtained through this clustering technique, in which ranges of service categories (A-F) are represented.

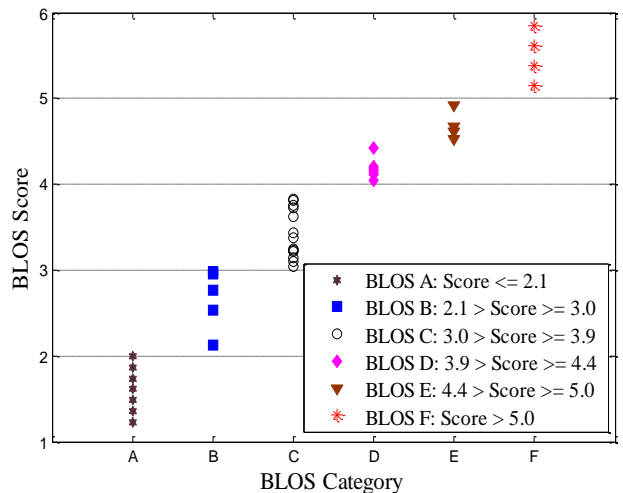


Figure 3: BLOS Categories (A-F) using Affinity Propagation (AP) clustering.

It was observed that, when model predicted BLOS score for a signalized intersection approach is below 2.1, it indicates that bicyclists riding in through movement are extremely satisfied. On the other end of this scale, BLOS score above 5.0 indicates that bicyclists feel extreme difficulties. Classification of BLOS scores into number of classes has passed through numerous statistical validation parameters discussed earlier and ranges have been defined with paramount care in the present context.

8. Findings and Conclusions

This study has identified the attributes significantly ($p < 0.001$) affecting the level of service or quality of service offered to bicycle through movement at signalized intersections under heterogeneous traffic flow conditions. Attributes of the main-street approaching to signalized intersections were observed to be dominant over attributes of the cross-street. Pavement condition index (*PCI*) at intersection area and five other variables of main-street such as road width per direction (*RW*), peak hour traffic volume (*PHV*), adjacent land use pattern (*LU*), on-street parking turnover (*P*) and average stopped time delay (*D*) of bicyclists were observed to be significantly contributing to the bicyclists' perceived satisfaction levels. Other parameters like intersection crossing distance, traffic volume on cross-street, pedestrian volume, type of signalization (e.g., actuated signals and semi-actuated signals) etc. were observed to have insignificant contribution in the model building. Though ratio of green period to cycle length was expected to be an important attribute, but it was found to be highly correlated with average stopped time delay with $R = -0.78$.

This research has developed a highly reliable BLOS model which has a high coefficient of determination (R^2) value of 0.86 with average observations. Both qualitative (perceived satisfactions of bicycle users) and quantitative (important attributes measured from sites) variables were incorporated in the development of this model. This study has deliberated unique variables those contribute bicyclists' satisfaction levels under heterogeneous traffic flow conditions. Researchers traditionally define ranges of LOS categories by taking overall perceived BLOS score as measure of effectiveness value (boundary between LOS 'C' and 'D') and further derive ranges of all classes at equal difference. This method never assesses similarity or dis-similarity that does exist among data points within a class or group of BLOS categories. Hence, this study has defined ranges of BLOS categories (A-F) in a more advanced way by using Affinity Propagation (AP) clustering technique and number of classes (six) adopted were also well supported by four cluster validation parameters. Hence this study has not only taken care of influencing variables in model fitting but also has given sufficient attention for the classification of BLOS scores into number of service categories.

The Bicycle Level of Service (BLOS) model developed and BLOS ranges defined in this study have got inclusive application efficiencies. It has been observed in this study that, approximately 69% of studied intersection approaches are offering above average quality of services (LOS A, B and C) and remaining 31% are offering below average kind of services (LOS D, E and F) to bicycle through movement. Wide approaches providing a good surface conditions and carrying traffic volume considerably below their practical capacities were observed to offer average or above average service levels. However, road intersections where service levels are below average in the prevailing conditions need geometric improvements with reduced volume of traffic flow. It is evident from this study that traffic volume on main-street approach that had a major contribution of 38% in the model building, is the most important parameter for BLOS offered to bicycle through movement. Hence provision of separate bicycle lane on main-street to minimize influences of motor-vehicles, is the most feasible solution to improve bicycle service qualities. Improvements in road and traffic management could able to reduce average stopped time delay which is another important parameter contributing approximately 15% to overall BLOS score of bicycle through movement at an intersection. The study findings would help transportation planners and engineers to systematically assess the quality of services offered by an intersection to the through movement of bicyclists. This, on the other hand, provides an insight on proper design of an intersection to accommodate both bicyclists and motorized traffic in better and safer ways.

Reference

1. Botma, H. (1995). Method to Determine Level of Service for Bicycle Paths and Pedestrian-Bicycle Paths. Transportation Research Record 1502, TRB, National Research Council, Washington, D.C., pp. 38-44.
2. Bowerman, B. L., and O'Connell, R. T. (1990). Linear statistical models: An applied approach, 2nd edn. Belmont, Duxbury, CA.
3. Caliński, T.; Harabasz, J. (1974). A dendrite method for cluster analysis, Communications in Statistics, Vol. 3, Issue 1, pp. 1-27. DOI: 10.1080/03610927408827101.

4. Carter, D., Hunter, W., Zegeer, C., Stewart, J., and Huang, H. (2007). Bicyclist intersection safety index. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2031, pp. 18-24. DOI: 10.3141/2031-03.
5. Cochran, W. G. (1977). *Sampling techniques*, 3rd edn, John Wiley & Sons, New York.
6. Crider, L. B., Burden, J., and Han, F. (2001). *Multimodal LOS-Point Level of Service Project-Final Report*. Department of Urban and Regional Planning, University of Florida, Gainesville, Florida.
7. Davies, D. L., Bouldin, D. W. (1979). A cluster separation measure, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-1, Issue 2, DOI: 10.1109/TPAMI.1979.4766909.
8. Davis, J. (1987). *Bicycle safety evaluation*. Auburn University, City of Chattanooga, and Chattanooga-Hamilton County Regional Planning Commission, Chattanooga, TN.
9. Dueck, D.; Frey, B. J. (2007). Non-metric affinity propagation for unsupervised image categorization, *IEEE 11th International Conference on Computer Vision*, pp. 1-8. DOI: 10.1109/ICCV.2007.4408853.
10. Dunn, J. C. (1973). A Fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, *Journal of Cybernetics* Vol. 3, No. 3, pp. 32-57. DOI: 10.1080/01969727308546046.
11. Gandomi, A. H.; Yun G. J.; Alavi A. H. (2013). An evolutionary approach for modeling of shear strength of RC deep beams. *Materials and Structures*, Vol. 46, pp. 2109-2119. DOI: 10.1617/s11527-013-0039-z.
12. *Highway Capacity Manual* (2000). Transportation Research Board, Washington, D.C., 1134 p.
13. *Highway Capacity Manual* (2010). Transportation Research Board, Washington, D.C., 1650 p.
14. *Indian Road Congress* (1994). *Guidelines for the Design of At-grade Intersections in Rural and Urban Areas*, IRC-SP-41, Indian Road Congress, New Delhi.
15. Jensen, S. U. (2013). *Pedestrian and Bicycle Level of Service at Intersections, Roundabouts and other Crossings*. 92nd Annual Meeting of the Transportation Research Board, TRB, Washington, DC, 19p.
16. Landis, B., Vattikuti, V., Ottenberg, R., Petritsch, T., Guttenplan, M., and Crider, L. (2003). Intersection level of service for the bicycle through movement. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1828, pp. 101-106. DOI: 10.3141/1828-12.
17. Menard, S. (1995). *Applied logistic regression analysis*. Sage university paper series on quantitative applications in the social sciences, Thousand Oaks, Sage, CA, pp. 7-106.
18. Milligan and Cooper (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, Vol. 50, Issue 2, pp 159-179. DOI: 10.1007/BF02294245.
19. Nash, J. E.; J. V. Sutcliffe. (1970). River flow forecasting through conceptual models part I-A discussion of principles. *Journal of Hydrology*, Vol. 10, No. 3, pp. 282-290. DOI: 10.1016/0022-1694(70)90255-6.
20. Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, *Journal of Computational and Applied Mathematics*, Vol. 20, pp. 53-65. DOI: 10.1016/0377-0427(87)90125-7.
21. Rubins, D. I., and Handy, S. (2005). *Bicycle Clearance Times: A Case Study of the City of Davis*. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1939, pp. 22-27. DOI: 10.3141/1939-03.
22. Shladover, S. E., Kim, Z., Cao, M., Sharafsaleh, A., and Li, J. (2009). *Bicyclist Intersection Crossing Times: Quantitative Measurements for Selecting Signal Timing*. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2128. DOI: 10.3141/2128-09.
23. Steinman, N., and Hines, K. (2003). *A Methodology to Assess Design Features for Pedestrian and Bicyclist Crossings at Signalized Intersections*. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1878, pp. 42-50, DOI: 10.3141/1878-06.
24. Strauss, J., Miranda-Moreno, L. F., and Morency, P. (2013). *Cyclist activity and injury risk analysis at signalized intersections: A Bayesian modelling approach*. *Accident Analysis & Prevention*, Vol. 59, pp. 9-17. DOI: 10.1016/j.aap.2013.04.037.