INVESTIGATING SERVICE PERFORMANCE OF SIGNALIZED INTERSECTIONS OPERATING UNDER MIXED TRAFFIC CONDITION

Suprava Jena, Corresponding Author
Research Scholar (Ph.D.)
Department of Civil Engineering
National Institute of Technology Rourkela, India 769008
Tel: +91-898-458-5857; Email: suprava728@gmail.com

Manaswinee Kar
Former M. Tech. Student
Department of Civil Engineering
National Institute of Technology Rourkela, India 769008
Tel: +91-768-391-6914; Email: manaswineekar@gmail.com

Prasanta Kumar Bhuyan
Assistant Professor, Department of Civil Engineering
National Institute of Technology Rourkela, India 769008
Tel: +91-661-2462313; Email: pkbtrans@gmail.com

Word count: 6,497 words text + (4 tables) x 250 words (each) = 7,497 words

Submission Date: August 1, 2018

ABSTRACT

This article proposes a new insight of modelling service quality offered by signalized intersections, which are the nodal focuses in a transportation network in developing countries. To achieve the objective of this research, a broad spectrum of geometrical, traffic operational, built-environmental and behavioral data sets are collected from 45 diversified signalized intersections with a widely varying driving environment through field investigations, videography techniques, and perception survey. Responses from around automobile drivers were gathered seeking socio-demographic information and overall satisfaction scores for respective approaches of intersections. Accordingly, the six parameters exerting significant influences on driver's satisfaction were highlighted by Spearman's correlation analysis. Exceptionally reliable, and less complex Automobile Level of Service (*ALOS*) models were formulated considering these six variables with the assistance of a unique and widely used artificial intelligence technique in particular, Multi-Gene Genetic Programming (MGGP) and Differential Evolution (DE). DE model displayed incredible likelihood efficiencies with high coefficient of determination (R^2) of 0.93 and 0.894 for training and testing datasets respectively. The sensitivity analysis showed that queue length plays a significant role in fixing ALOS standards of signalized intersections, having highest negative influence of 67.153%. Hence, optimizing traffic signalization timings and increasing effective green time for major approaches of an intersection in peak hours will significantly enhance the service quality of respective intersections. Similarly, other parameters are ranked in decreasing order of their relative importance to help the transportation administrators for making efficient resolutions for better infrastructural design.

Keywords: Signalized intersections, Automobile mode, Level of service, Multi-Gene Genetic Programming, Differential Evolution, Sensitivity analysis

INTRODUCTION

Over the generations, automobile has been the most dominant mode of road transportation, which progressively reshape patterns of living by increasing the accessibility and interconnectedness between different regions in developing nations. Hence, the demand for automobile mode of transportation is gradually increasing to manage the necessities of both personal and professional life. On the other hand, managing road networks has been dreadfully challenging due to limited space and resources. The road traffic in developing countries is heterogeneous in nature, which indicates a varied mix of fast and slow-moving vehicles traveling without proper lane discipline. Therefore, the models proposed in developed countries are not transferable to the contexts of developing countries, where the road geometric and traffic flow conditions are considerably different from the former. This is the fact due to which, urban street infrastructural developments in emerging countries like India is still far away from satisfactory irrespective of current improvements based on the applications of existing evaluation strategies. Approximately 83% of the world population resides in developing countries, but no standard handbook like Highway Capacity Manual (HCM) is available for the reference of highway authorities in these countries (*1, 2*). This article proposes a new insight of modelling service quality offered by signalized intersections, which are the nodal focuses in a transportation network, in urban Indian context. On the other hand, the studies carried out for determining signalized intersection LOS under heterogeneous traffic flow conditions have followed mostly the principles of HCM and quantified based on the ranges of travel time and delay. Whereas, the roles of various other road attributes and travelers' characteristics require a detailed investigation, which are commonly found in Indian traffic scenario. With the aim of justifying the existing approaches, this objective of this study is to proposed suitable Automobile Level of Service (ALOS) models to assess service quality provided by signalized intersections operating under heterogeneous traffic flow conditions.

To achieve the objective of the research, the first part is to investigate important service attributes, those influence the drivers' satisfaction level under highly heterogeneous traffic flow condition. To achieve the objective of this research, a broad spectrum of geometrical, traffic operational, built-environmental and behavioral data sets are collected from a widely varying driving environment through field investigations, videography techniques, and perception survey. Most of the researchers have used the conventional regression analysis, point systems, or probit and logit analysis to model the service levels of signalized intersections. However, Artificial Intelligence (AI) techniques are powerful alternatives to statistical techniques, which have better service prediction ability with drastically reduced processing time to achieves high level of accuracy. Therefore, two novel evolutionary algorithm of AI techniques namely, Multi-Gene Genetic Programming (MGGP) and Differential Evolution (DE), are implemented to develop a suitable ALOS model. The study also implemented a model sensitivity analysis, which have arranged the influential service attributes in order of their relative importance. Specific percentage values are determined for each input parameters based on their degree of importance on comfort level of automobile drivers. Based on this, several improvement strategies are outlined in this study. The proposed ALOS models developed though these analyses would assist the transport authorities to identify operational issues of existing street facilities, and to design a users' friendly transport system.

BACKGROUND LITERATURES

Several researchers, since the past few decades, have carried out extensive works regarding the performance assessment of signalized intersections based on user perception. Sutaria and Haynes carried out a user perception investigation to know drivers' opinions regarding LOS at intersections with signals. Delay was concluded as the highly influencing factor for defining LOS ranges (*3*). Pecheux *et al.* distinguished that LOS ranges based on delay at intersections with signals were not based straight away on surveys of user opinions (*4*). Zhang and Prevedouros developed a prototype to bring together delay and safety to obtain a widespread LOS indicator, i.e., delay and safety index for signalized intersections (*5*). The suggested methodology modelled the trade-off between safety and efficiency clearly and included both inter-vehicle and vehicle-to-pedestrian clashes connected with left turns. Zhang *et al.* (2007) analysed the responses of drivers to find how road users perceive difficulty in turning left at signalized intersections (*6*). Lee *et al.* devised a different methodology for assessing transportation service nature and facilities offered by signalized intersections utilizing fuzzy aggregation and a traditional consent analysis technique (*7*). Fang and Pecheux dealt with, how users framed an idea about the superiority of service at signalized intersections and how many levels of services motorists were capable of perceiving (*8*). Zhang and Prevedouros presented a method basing on fuzzy logic in order to decide LOS of signalized intersections which clearly accounted for user opinions (*9*). Delay was considered as the significant and one and only standard for deciding LOS of signalized intersection.

Chen *et al.* broadly studied urban arterial performance evaluation having principal attention on the determination of average commute time (*10*). Saha *et al.* offered an improvised model for estimating delay at signalized intersections with the prevailing traffic flow conditions being highly heterogeneous since the prototypes established based on the homogeneous traffic conditions yielded inaccurate outcomes for developing nations, where the traffic is extremely varied with practically no lane discipline (*11*). Ban *et al.* studied the ways to find out real time queue lengths utilizing intersection travel times accumulated from mobile traffic sensors at signalized intersections (*12*). Chang *et al.* presented a simple procedure to calculate queue length on any approach of a signalized intersection (*13*). This methodology had a minimal set of data especially flow, occupancy, cycle length, and detector setback, as compared to the existing techniques.

From the reviews of literatures, it is concluded that, delay and travel time are not the only parameters influencing riding quality of drivers. The driving environment at signalized intersections are influenced by several in-built environmental elements like, un-authorized on-street parking turnover, vending activities, interruptions due to slow movement of non-motorized traffic and oppositely moving encounters etc., which are not addressed in the previous studies. Ignoring drivers' behavior, any implementation would face a possibility of strong rejection. In light of this, the present study has examined impact of wide range of road attributes under mixed traffic flow condition to justify the existing approaches of evaluating offered service quality at signalized intersections.

SELECTION OF STUDY AREA

The elementary requirement for a thorough investigation of urban driving environment includes the preparation of a wide-ranging and well-diversified database. The data sufficiency and data diversity leads to the development of a well-generalized service prediction model for the context of interest. Thus, both of these principles are given due consideration in this study for carrying out a comprehensive analysis of ALOS criteria. To satisfy the above requirements of site selection, required quantitative as well as qualitative data sets were collected from 178 approaches of 45 signalized intersections from two Indian cities namely, Kolkata and Rourkela. In the group of 45 signalized intersections, 34 are 4-legged intersections and the remaining11 are

3-legged or T-shaped intersections. The diversities in site selection includes varying number of lanes, presence and width of various geometrical elements, pavement surface smoothness, varying traffic volume, land use pattern (residential, commercial or office area), etc. All these intersections inspected in the present study have actuated and semi- actuated traffic signals. The behavior and composition of traffic along with the road infrastructures are also different from each other in these cities. Moreover, the selected signalized intersections are chosen for ideal representation of heterogeneity in Indian traffic scenario.

DATA COLLECTION

The accuracy in data collection is highly desirable to reduces the likelihood of errors and to maintain novelty of the research. From the literatures, the basic information was gathered about the limitations of applying former methods to evaluate urban street service quality under highly heterogeneous traffic flow condition. To achieve the objective of the research in a best possible way, both quantitative and qualitative parameters affecting driver's comfort level at signalized intersection have been considered in data collection process. These variables were extracted approach wise in an excel sheet, which serves as a database for development of ALOS model to evaluate the service quality provided by signalized intersections operating under mixed traffic flow condition. The timings of data collection were chosen during both morning (between 8.00-11.00 AM) and evening peak hours (between 4.00-7.00 PM), to reflect probably the worst traffic conditions encountered by any drivers. Traffic flow data were not collected on weekends and holidays as the traffic volume are comparatively low on these days.

Extraction of quantitative service attributes

To outline most anticipated QOS parameters for a comprehensive ALOS model, variations of drivers' level of satisfaction with respect to the existing transportation facilities have been exposed to a lot of research and discussion. So that the complications associated to each occurrence can be reasonably well defined. The term quantitative parameter refers to the road geometrical, traffic operational and built-environmental variables considered in this research, that were quantified by in-situ investigations and videotaping. Along with the literature survey, a pilot questionnaire survey was also conducted to gather basic information related to the Quality of Service (QOS) attributes affecting driver's comfort level on urban street signalized intersections. Based on the findings of pilot survey, all feasible quantitative variables are listed out with their corresponding units in Table 1. to collect the datasets from selected signalized intersections, operating under highly heterogeneous traffic flow condition.

The road dimensional or geometric parameters like effective road width, width of median (if present), shy distance etc. were recorded using a measuring tape and expressed in meters. Effective road width (W_e) can be defined as the available road width for free movement of traffic. Excluding un-authorized on-street parking activity and vendor encroachment respectively. In-situ investigations were carried out to examine the pavement surface quality, road markings, camber/cross slope presence or absence of median etc.

High-resolution video cameras fitted on tripod stands were used to directly record the traffic flow parameters like, volume and composition of different types of vehicles in the traffic mix during the period of data collection. The recorded video clips were played on the monitor and each category of vehicle volume per respective approaches of intersections was counted separately for peak 15-minute period. Running average method was used to decide one peak hour of traffic volume (*PHV*) among the hours of rush conditions. The conversion factors mentioned for different types of vehicles in Indian Road Congress (IRC): 106 were multiplied with the classified traffic count to express the extracted peak hour through traffic volume in terms of Passenger Car Units per hour (PCU/h) (*14*). On street pedestrian volume was also calculated from the recorded video clips.

TABLE 1 Extracted QOS parameters collected from Signalised intersections

Sl.	QOS parameters	Units		
1	No. of lanes	in number		
2	No. of Intersection legs	in number		
3	Effective road width	in meter		
4	Pavement condition index	Rating scale: 1-5		
5	Control Delay	in seconds		
6	Queue length	in number		
7	Cycle length	in seconds		
8	Effective green time	in seconds		
9	Red clearance	in seconds		
10	Motorized traffic volume	in PCU/h		
11	Non-motorized traffic volume	in PCU/h		
12	On street pedestrian volume	in number		
13	Proportion of vehicles arriving on green time	in PCU/h		
14	Capacity	in PCU/h		
15	Composition of heavy vehicles	% in total traffic		
16	Shy distance in meter	in meter		
17	Direction of traffic movement	One-way		1.0
		Both way		0.5
18	Presence of median barrier	Absent		0.0
		Present	Width < 500 mm	0.5
			Raised or Width ≥ 500mm	1.0
19	Presence of camber	Absent		0.0
		Present	Somewhat irregular	0.5
			Proper	1.0
20	Presence of grade separated sidewalk	Absent		0.0
		Present		1.0
21	Presence of separate bike lane	Absent		0.0
		Present		1.0
22	Land use pattern	Residential (No/low commercial density)		0.0
		Mixed (medium commercial density)		0.15
		Commercial (high commercial density)		0.3
23	On-street parking turnover	Absent (< 5%)		0
		Low (5-25%)		0.13
		Moderate (25-45%)		0.27
		High (> 45%)		0.4
24	Frequency of interruptions due to frequently stopping public transits ahead	Presence of bus pull-out lane		0.0
		Low		0.15
		Moderate		0.3
		High		0.45
		Very high		0.6
25	Average time interval to face an encounter	in minutes		

Though, delay measurement directly from the field is possible, yet, it is a quite tedious process. Control delay (*d*) is a portion of total delay that results from the type of control at the intersection. From the extensive literature survey conducted on delays at signalized intersections, it is found that cycle length, green ratio (E_g/C) and volume to capacity ratio (v/c) significantly influence control delay at the intersections. Some associate parameters were also extracted at a prior to assist the estimation of control delay. Effective green time (E_g) is the time during which a given traffic movement or set of movements may proceed at saturation flow rate. Cycle length (C) is the time in seconds, that takes a signal to complete one full cycle of indications. It is noted as the time interval between the starting of green for one approach till the next time the green for that approach starts. Capacity (c) as defined by the HCM, is the maximum hourly rate at which persons or vehicles can be reasonably expected to traverse a point or a uniform segment of a lane or roadway during a given time period under prevailing roadway, traffic and control conditions. Capacity of each approach was determined by playing the videos of arrival and departure, which were recorded using high resolution video cameras installed at upstream and downstream side of the signalized intersections. Saha *et al.* (2017) have proposed a more accurate and justifiable approach for estimating control delay under heterogeneous traffic condition taking the effect of all the above mentioned parameters. Control delay (*d*), mainly comprises of three components, namely, uniform delay (d_1), oversaturation or incremental delay (d_2) and the stopped delay (d_3).

$$d = d_1 + d_2 + d_3 \tag{1}$$

Where,

- 'd_1' is the uniform delay of vehicles entering into an intersection, which is a function of C, X (which is $=v/c$) and E_g/C as shown in equation (2).

$$d_1 = \frac{0.5 \times C \times (1 - \frac{E_g}{C})^2}{(1 - X \times \frac{E_g}{C})} \tag{2}$$

- 'd_2' accommodates the variation in arrival of vehicle platoons at different points of signal cycle, especially during green time. d_2 depends on the Platoon ratio (R_p), which is calculated by the ratio of percentage of vehicles arriving during green to percentage of green time.

$$d_2 = 6.23 - 15.35 \times R_p \tag{3}$$

- 'd_3' is zero here, as no vehicles are left after the cycle is completed.

Mathematically, control delay at signalised intersections is calculated using equation (4).

$$d = 6.23 + \frac{0.5 \times C \times (1 - \frac{E_g}{C})^2}{(1 - X \times \frac{E_g}{C})} - 15.35 \times R_p \tag{4}$$

Queue length (Q_l) is the distance of vehicle, stopped farthest from the STOP line during the cycle as a result of the display of a red signal indication. The back-of-queue size depends on the arrival pattern of vehicles and on the number of vehicles that do not clear the intersection during the previous cycle. Q_l on any selected approach is estimated by considering a longitudinal trap, which is extended from stop line of the intersection to the end of queue, in order to mark the entry and exit of vehicles at the intersection.

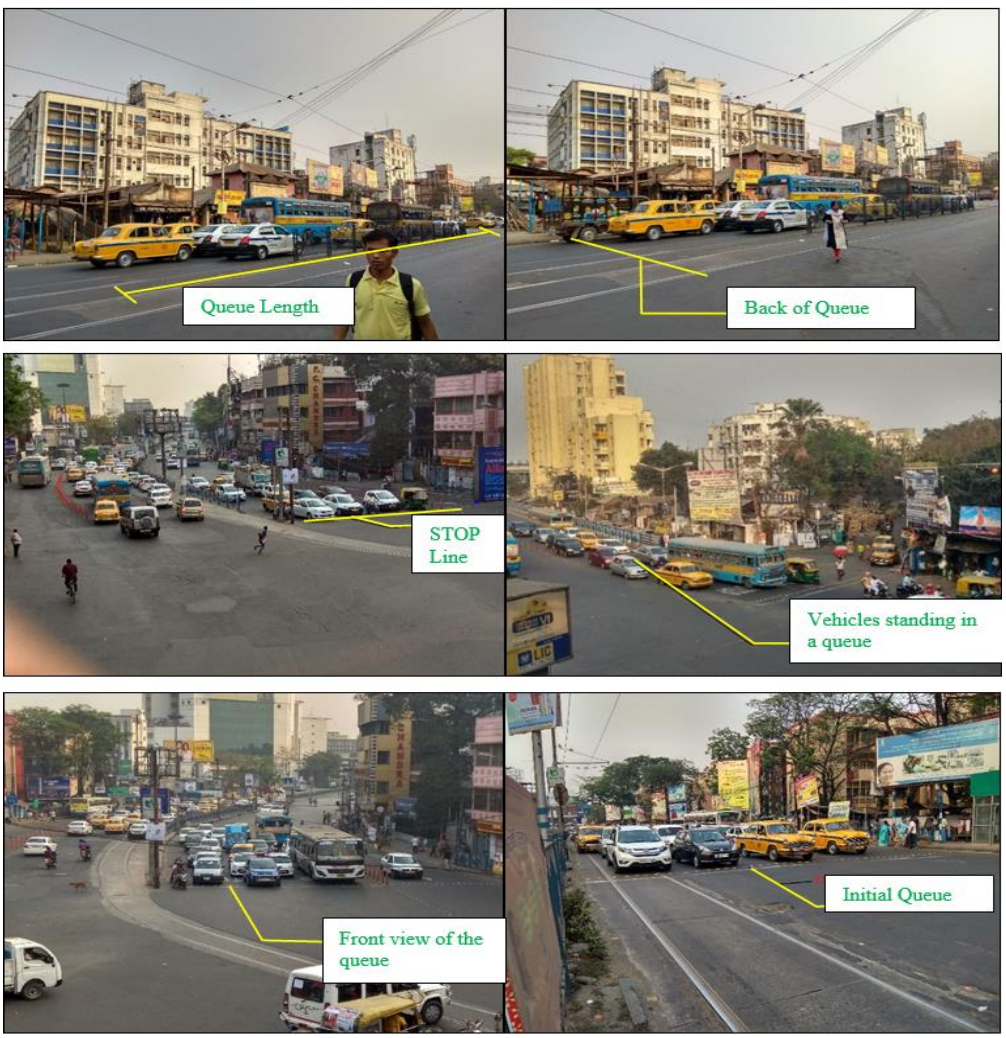


FIGURE 1 Different parameters considered for Queue length estimation

Q_l for a particular approach of an intersection are expressed in terms of number of vehicles previously standing in a queue along with those just arriving within a five seconds interval, excluding those have already crossed the STOP line during that time period (*2*). Figure 1 presented a schematic diagram of intersection, illustrating queue length.

These obstructions in driving originated from few other sources, like on-street parking turnover, intensity of commercial activities in the adjoining areas, frequent stopping of public transits, slow movement of non-motorized vehicle, interaction with encounters etc. are examined with reference to the adjustment factors, proposed by Jena *et al.* (*15*)

Extraction of qualitative service attributes

This research presents a qualitative study on automobile users' response pattern to assess the provided transportation service quality at signalized intersection operating under heterogeneous traffic flow conditions. The whole questionnaire was divided in to two phases. 'Section A' is prepared to record the demographic information of survey participants and 'Section B' to rate the influencing parameter (with respect to different dimension of service attributes) affecting overall satisfaction level of drivers of, while driving along the road. People of varying socio-economic categories were included in "Traveler's intercept survey" to judge how the real-time response of drivers varies with respect to different road and traffic conditions. The automobile drivers are either orally interviewed about the trip quality on the spot or fill the survey form given to them based on their stated preference immediately after driving on a particular road. The drivers were well explained about the purpose of this study, and requested to answer their stated preferences on the prepared questionnaire on a Likert scale ranging from 1= highly dissatisfied to 6= highly satisfied.

The real time perception of 3775 drivers about the road attributes are extracted on an excel sheet. The average of all responses with respect to different dimension of service quality were taken to determine the Overall satisfaction of each driver on respective approaches of an intersection. Approximately 53% and 47% of the drivers are male and female respectively. From the classification done with respect to age, young population comprise of 39 %, middle- aged are found to be 45 % and elderly people constitute 16% of the total population considered for survey. However, drivers of age less than 18 years were excluded from this survey due to lack of enough experience to give proper judgement about the road infrastructure and traffic conditions. With respect to driving experience, 39% have an experience of 2-5 years, 29% have an experience of 5-15 years while 32% have greater than 15 years of driving experience. 87% of the drivers are employed, while, rest are unemployed. From the survey, 32% drivers are found to be using bikes, 30 % cars, 16% LCVs, and 22% autos. All of the trips were intended for various purposes like official work, business, school, marketing, etc. Considering the contextual diversity in the data sets in developing ALOS model, it is anticipated that these study findings will be universally applicable to forecast the service quality of signalized intersections in developing countries.

MODELLING APPROACHES

Although statistical modelling strategies are to a great extent utilized for the predictive models, AI techniques for model building, is quite advantageous in the way of solving nonlinear and complex problems. These are capable of tolerating imprecision, uncertainty, and partial truth to achieve tractability and robustness on simulating human decision-making behavior with drastically reduced processing time and high accuracy under low supervision as compared to other algorithms. Therefore, the current work has introduced the application of two novel evolutionary algorithms, to develop a suitable ALOS model for solving the existing problems with better reliabilities.

Multi-Gene Genetic Programming (MGGP)

MGGP is an extension of Genetic Programming (MGGP), also known as symbolic regression. which creates exceptionally compacted prescient conditions (*16*). It naturally advances computer programs to create prescient models without indicating their structures in advance. Once the initial populace is ready, the objective function surveys the forecasting capacity of each parental quality and returns a fitness value for it. In the current workspace, the root mean-square error (RMSE) of observed and anticipated estimations of the output variable is utilized as the objective function. Unless the issue is so little and straightforward, the greater part of the initial genes shows extraordinary poor fitness esteems. Consequently, *GP* makes a posterity populace of better-fitted people by executing different hereditary proprietors (reproduction, crossover, and mutation) on the initial genes. This progression is known as generation-1. In the propagation or reproduction process, better-performing people of the underlying population are straightforwardly duplicated to the posterity populace. In the crossover process, better-performing posterity qualities are made by trading the subtrees of any two parental qualities or genes chosen in extent to their fitness value. In the mutation process, a haphazardly chosen node of the parent is supplanted with a component of the terminal set. Here, the functional node is supplanted with a functional element and the terminal node is supplanted with a terminal component or element.

MGGP utilizes mathematical functions to set up the best configuration of input variables for the expectation of the yield variable with high exactness. These mathematical functions do different blends as well as changes of the input variables to make them very efficient for the forecast of model yield. At generation-1, the recently generated posterity populace replaces the underlying one. The posterity populace again experiences the hereditary activities to make another populace of better-fitted posterity people (generation-2). This iterative procedure proceeds over numerous generations and ends when either the threshold fitness value or the greatest of the number of generations is accomplished. Consequently, the best-fit individual showed up at any generation defines the yield of the GP formalism. The MGGP formalism is additionally done in comparative way with the exception of that, every individual in MGGP is a linear amalgamation of at least two trees of GP associated with weights, which are alluded as 'genes'. Therefore, this extemporized procedure is named as multi-gene GP or MGGP. The combination of two or more genes occurs as follows to estimate the output variable (ALOS):

$$y = a_0 + \sum_{j=1}^{n} a_j g_j = a_0 + \sum_{j=1}^{n} a_j \times F[X, f^{'}(X)] \quad (5)$$

Where,

a_0 is the bias parameter, n represents the number of genes (g) in the target expression, a_j is the weight or linear coefficient of j^{th} gene $g_j = F\ [X, f'\ (X)])$, F represents the model function, X represents the vector of influencing variables, and f' is a functional element selected from the functional set. In MGGP, a_j and a_0 parameters are estimated by using the ordinary least squares method, i.e., by minimizing the sum of squared errors between actual and predicted outputs.

Differential Evolution (DE)

DE Algorithm, introduced by Storn and Price (*17*), is a stochastic, population-based optimization algorithm evolved to optimize real parameter, real valued functions. DE is a method that optimizes a problem by trying to improve a candidate solution through several iterations with regard to a given measure of quality. Through several iterations, DE makes a population of candidate solutions, which converges to an optimum of the function. It is a variant of GA and proposed to solve global optimization problem over continuous space. DE is utilized for the sake of real-valued functions having multiple dimensions but does not utilize the slope of the problem that is getting optimized. This implies that DE does not need the optimization problem to be differentiable, while classic optimization methods such as gradient descent, quasi-newton methods need it. Thus, DE maximizes or minimizes problem by upholding a populace of candidate solutions and generating new candidate solutions on combining current ones as per the unpretentious formulae, and then keeping the candidate solution with a best score on impending optimization problem. In this manner, the optimization problem is preserved like a black box which merely gives a standard of quality.

A basic variation of DE algorithm works on the principles of a populace of candidate solutions, which are known as 'agents'. The agents are stimulated around in the exploration-space through mathematical formulae, which are quite simple to conglomerate the positions of current agents from the populace. Suppose, the new position of an agent is an improvement, then, it is acknowledged and it forms a portion of the populace, else, the new position is simply prohibited. The process is iterative in nature and by doing so it is anticipated, that a satisfactory solution will eventually be revealed.

Mechanism of DE

The DE algorithm uses three major operators similar to GAs, such as: mutation, crossover and selection operators. However, DE depends heavily on mutation as a primary search mechanism to distinguish it from the GA. It is similar to GA except that, the candidate solutions are not considered as binary strings or chromosome, but usually as real vectors. One key aspect of DE is that the mutation step size is dynamic. That means, it adjusts to the configuration of the population and will tend to zero, when it converges. The predicted '$ALOS_{sig}$' can be estimated using the following equation:

$$ALOS_{sig} = a_0 + \sum_{i=1}^{n} a_i x_i \tag{6}$$

Where, a_0 is the bias parameter, n represents the number of input variables in the target expression, a_i is the linear coefficient of 'i^{th}' variable x_i.

RESULT AND ANALYSIS

This section deals with the identification significant input variables having primary influences on ALOS of signalized intersections. The procedure followed to develop a reliable ALOS model for signalized intersections is also explained in this section, where data sets of 124 approaches were utilized for model training, and the remaining 54 approaches were kept for the purpose of model testing. The aggregate database was divided in such a way that the properties of two groups of data must have the wide variation similar to the sample size. The prediction performance of these models was assessed in terms of several statistical parameters to decide the better one in the present context.

TABLE 2 Descriptive statistics and Spearman's correlation among variables

Variables	Correlation with $ALOS_{sig}$	Significance
Peak hour traffic volume per effective road width *(PHV/W_e)*,	-0.822	0.000
Control delay *(d)*,	-0.496	0.000
Pavement condition index *(PCI)*,	0.377	0.000
Queue length, *(Q_l)*,	-0.609	0.000
Interruptions due to non-motorized traffic(O_{NV})	-0.329	0.000
Oppositely moving encounters (O_E)	-0.281	0.000

Correlation (ρ value) among input variables

Variables	PHV/W_e	d	PCI	Q_l	O_{NV}	O_E
PHV/W_e	1.000	0.481	-0.226	0.661	-0.042	0.159
d	0.481	1.000	0.001	0.275	0.106	0.095
PCI	-0.226	0.001	1.000	-0.226	0.022	-0.115
Q_l	0.661	0.275	-0.226	1.000	-0.029	0.262
O_{NV}	-0.042	0.106	0.022	-0.029	1.000	-0.016
O_E	0.159	0.095	-0.115	0.262	-0.016	1.000

Descriptive statistics

Variables	Range	Mean	Standard Deviation
PHV/W_e	(28.57-1053) in PCU/h/m	228.186	208.931
d	(0-14.75) in second	7.673	2.9702
PCI	2-5	3.75	0.735
Q_l	0-14	7.12	3.687
O_{NV}	0-0.5	0.0404	0.1044
O_E	0-0.5	0.1837	0.1026
$ALOS_{sig}$	1.4-5.6	4.3002	0.922

Prediction of $ALOS_{sig}$ with the help of MGGP model architecture

The development of ANN model was carried out by taking six predictor variables (i.e. *PHV/W_e, S_{avg}, PCI, P, LU* and *HF*) as input neurons and one response variable (i.e. *ALOS_score*) as output node. In the current study, the *MGGP* algorithm is executed utilizing *MATLAB* to evaluate the model parameters (a_j and a_0). The best *ALOS* prototype demonstrate for the present study was acquired with, a populace size of 1000 at 150 generations, g_{max}, D_{max}, p_r, p_c and p_m values of 5, 5, 0.02, 0.84 and 0.14 individually. It demonstrates two arrangements of ALOS models through green-and blue shaded specks or dots. Here, the green dots are the arrangement of non-dominated models (better-performing or superior models), and the blue dots are the arrangement of dominating ones (inferior models). The curve of non-overwhelmed models is called as the "Pareto front". Along these lines, each point on the Pareto front signifies the present issue from which the best one is to be chosen. All models on the Pareto front predominantly differ from each other regarding their likelihood capacities and expressional complexities. The location of the best model on the Pareto front is demonstrated utilizing an arrow as shown in Figure 2 (a). The best MGGP model was made out of five genes. Every gene is an element of certain arrangement of input factors. Consequently, the fundamental difference between these genes is that they exhibit various nonlinear alterations of a specific set of factors. Presenting the structures of individual genes, their scientific articulations are determined as under:

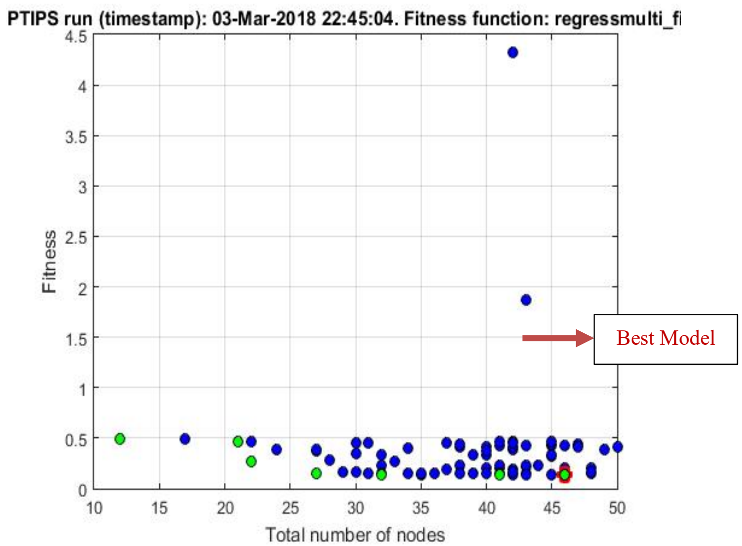
Gene-1= $7.574*x_1-3.787*x_2-3.787*x_3-3.787*x_4-3.787*(x_5)^4-3.787*(x_5)^2+0.9045$

Gene-2= $-1.126*x_6$

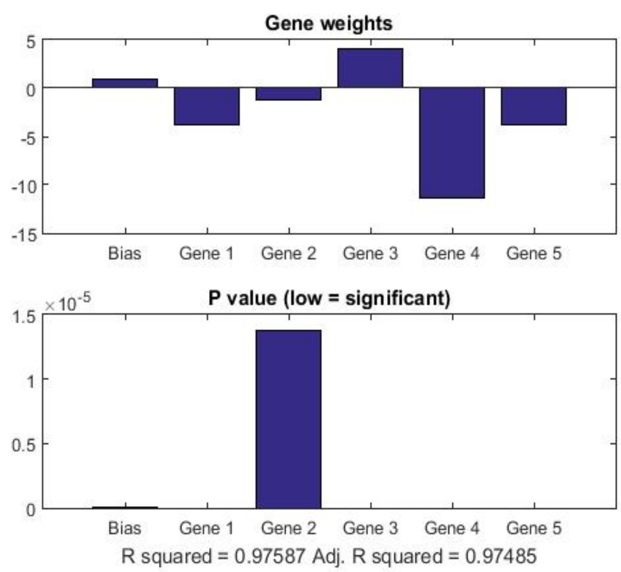
Gene-3= $4.074*x_3+8.148*(x_5)^2+3.71$

Gene-4= $-11.34*x_1-11.34*(x_5)^6$

and Gene-5= $3.758*x_1+3.758*x_2+3.758*x_4-3.758*x_5-3.758*(x_5)^6$



(a)



(b)

FIGURE 2 (a) Population of evolved models in terms of their complexities and fitness along with the best model: (b) MGGP based ALOS model showing gene weights and *p*-values

These genes don't represent an *ALOS* model individually, rather each one of them are joined together based on equation (5) to build an ALOS model. Specifically, the mathematical formulations of individual genes were multiplied with their comparing weighs and acquired outcomes were summed up with the bias term to develop the final ALOS model. The estimations of the bias term (a_0) and weights of five qualities (a_1, a_2, a_3, a_4 and a_5) were evaluated at 95% confidence level ($p < 0.05$). Statistical significances (p-values) of these coefficients are introduced in Figure 2(b). As seen, the most astounding p-esteem is found for gene-2, which is under 1.5 ×10-5. Incorporating all these genes in equation (5), the mathematical interpretation of the *MGGP*-based *ALOS* model is inferred and exhibited in equation (7).

$$ALOS_{sig} = 0.287x_3 - 0.029x_2 - 0.008x_1 - 0.029x_4 - 3.758x_5 - 1.126x_6 - 3.787(x_5)^4 - 15.1(x_5)^6 + 4.361(x_5)^2 + 4.615 \tag{7}$$

Prediction of *ALOS*$_{sig}$ with the help of DE model architecture

By implementing the DE algorithm coded in MATLAB, the coefficient of each input variable and the constant term were estimated for signalized intersections. Many trials with different lower and upper bounds for coefficient of predictor variables and constant term were changed to achieve a reliable model with better R^2 value and lesser RMSE. The values of particle best obtained from the DE algorithm is represented in Table 3.

TABLE 3 Particle best results obtained from DE model for signalized intersections

Variables	Coefficients	Particle Best Values	Standard error of the Estimates	Significance
Constant	a_0	2.75	0.131	2.04E-86
PHV/W_e	a_1	-0.0031	0.000	2.94E-57
d	a_2	0.0307	0.007	0.000134
PCI	a_3	0.32	0.029	1.06E-09
Q_l	a_4	-1.181	0.007	0.011405
O_{NV}	a_5	-0.854	0.192	3.57E-49
O_E	a_6	-0.8511	0.198	3.43E-11
Model summary				
Number of observations	***R***	***R*^{2}**	**Standard error**	**Significance**
124	0.966	0.930	0.257	2.41209E-93

The mathematical interpretation of the DE model to determine $ALOS_{sig}$ is shown as:

$$ALOS_{sig} = 2.75 - 0.0031 * \frac{PHV}{W_e} + 0.0307 * d + 0.32PCI - 1.181 * Q_l - 0.854 * O_{NV} - 0.851 * O_e \quad (8)$$

Prediction performance of different ALOS models

To determine which model is the best one among the developed models; Rank Index (RI) was proposed by Cai et al. and later a modified rank index (MRI) was proposed by Beura and Bhuyan *(18, 19)*. The performance measuring parameters are chosen from both of them.

The prediction performances of ALOS models proposed for the assessment of signalized intersections are compared in Table 4, for both training data sets and testing data sets. Based on various statistical parameters of MRI, the lowest MRI of '5' is assigned to DE modelling approach. Therefore, DE model is concluded to have highly precise prediction efficiency within both training and testing data sets to predict $ALOS_{sig}$. of different approaches of signalized intersections, as the prediction efficiency of MGGP model is comparatively inferior than DE. It is found out from Figure 3, that the model is well validated with R^2-values for training and testing datasets of 0.93 and 0.894 respectively.

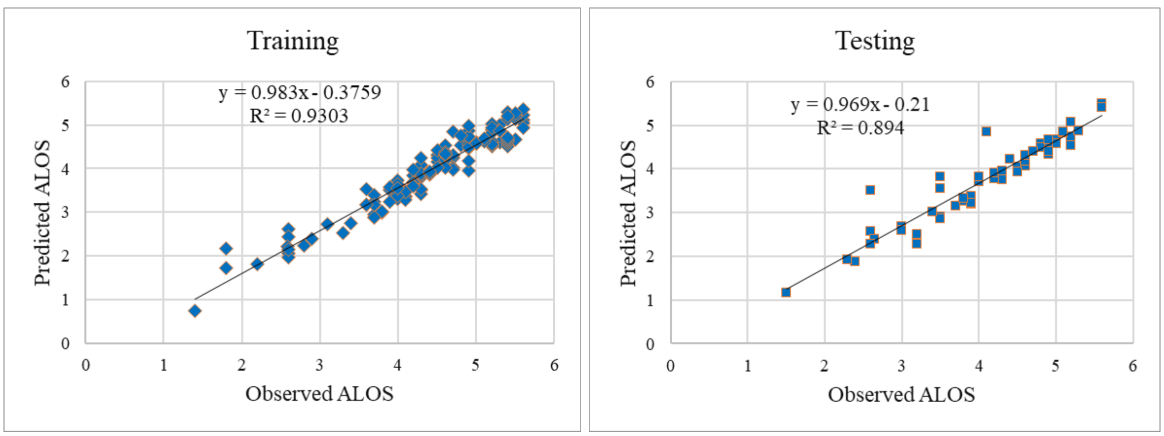


Figure 3 Validation of DE model for signalized intersections

Defining ranges of ALOS Classes (A-F)

In this investigation, the apparent estimated $ALOS_{sig}$ obtained for signalised intersections were classified on a six-point scale (1– 6), where '1' represents the worst and '6' represents the best driving environment perceived by the users. By considering the symmetry of six ALOS classes, the mean estimation of the ALOS scores is 3.5, which corresponds to the boundary between service categories "C" and "D". Thus, overall symmetrical partitions were made on the two sides of 3.5 to define the threshold values for service categories 'A– F'. If the range of $ALOS_{sig}$ is > 5.1, >4.33 to ≤5.1, >3.5 to ≤4.33, >2.67 to ≤3.5, >1.84 to ≤2.67 and ≤ 1.84; then the corresponding ALOS categories will be A (Excellent), B (Good), C (Average), D (Below average), E (Poor) and F (Very poor) respectively.

Model Comparison

An example problem has been discussed to help in understanding the genuine implementation of the suggested ALOS model in the field. DE model has been chosen in this discussion, since it is more efficient than all other models in terms of complexity as well as reliability to predict the $ALOS_{sig}$. The followings are the observations collected from the field for one-side through

movement of an approach of Hudco Crossing. Such as: PHV/W_e = 132.157 PCU/h/lane, d = 10.38 seconds, PCI = 3, Q_l = 3, O_{NV} = 0 and O_E = 0.1.
Using equation (8) of, the values of $ALOS_{sig}$ is calculated as 4.3.

Similarly, the observations for the other side through movement of the same approach are: PHV/W_e = 236.9857 PCU/h/lane, d = 7 seconds, $PCI = 3$, Q_l = 9 PCUs, O_{NV} = 0 and O_E = 0.4. So, the $ALOS_{sig}$ is calculated as 3.6.

Therefore, predicted $ALOS_{Overall}$ score for the intersection can be averaged out to be 3.95. From the field investigation, the perceived ALOS scores for both the sides of the approaches are 4.7 and 3.5 respectively. Hence, the $ALOS_{Overall}$ score for the intersection can be averaged out to be 4.1. By referring to the ALOS ranges, the predicted LOS category for the intersection is designated as ALOS category 'C'. Similarly, the predicted $ALOS_{sig}$ anticipated by each approach of all other signalized intersections shows that, around 13%, 40%, 28%, 10%, 6%, and 2% of the investigated approaches of signalized intersections are offering service categories of 'A', 'B', 'C', 'D' 'E' and 'F' respectively for automobile through movement.

TABLE 4 Model comparison using MRI

Model			MGGP		DE	
Data Type			**Training**	**Testing**	**Training**	**Testing**
Best fit Calculation	R^2		0.975	0.875	0.93	0.89
	E		0.60	0.73	0.64	0.76
	R_1		2		1	
Error Measuring Parameters	RMSE		0.60	0.49	0.51	0.46
	AAE		0.50	0.43	0.46	0.41
	MAE		0.96	0.95	0.95	0.93
	R_2		2		1	
$ALOS_{Pr}$/ $ALOS_{Act}$	Mean		0.84	0.88	0.89	0.91
	Sigma		0.082	0.098	0.073	0.096
	R_3		2		1	
Cumulative Probability	Probability	P_{50}	0.87	0.89	0.89	0.904
		P_{90}	0.95	0.98	0.97	1
	R_4		2		1	
20% Accuracy	Log-Normal		87.54	88.32	88.92	89.55
	Histogram		89.2	90	90.32	91
	R_5		2		1	
Overall Rank	MRI		10		5	
	Final Rank		**2**		**1**	

Improvement strategies for signalized intersections

The variation in drivers' satisfaction level at signalized intersections with respect to percentage change in the input parameters was tested using a model sensitivity analysis (*20*). Based on the obtained S_i values, 'Q_l' is found to be the most important variable, having the highest relative importance of 67.153% on comfort level of drivers. Therefore, it plays a significant role in fixing the ALOS standards of signalized intersections in the present context. Optimization of traffic signalization timings and increasing effective green time for major approaches of an intersection in peak hours will significantly enhance the service quality of respective intersections. Similarly, PHV, W_e, PCI, d, O_{NV}, and O_E are ranked second, third, fourth, fifth, sixth and seventh important parameters with relative importance of 16.598%, 7.047%, 3.899%, 1.84%, 1.734% and 1.728%

respectively. Referring to the rank of variables, several strategies are made for improvement of driving condition at signalized intersections.

- The negative influence of PHV/W_e can be mitigated through the widening of existing carriageway width, removal of road side obstructions for the provision of extra driveway width, and construction of paved shoulders beyond the outermost lane for pedestrian movement. Besides, the street vending and parking activities on sides of the road should be minimized by imposing strict rules by the city authorities.
- The pavement surface requires regular maintenance for smooth ride of vehicles along the road ways, which in turn increase the drivers' comfort level at signalized intersections.
- The intersections need to have adaptive traffic signal control to adjust the timing of their green light cycles to improve the current traffic conditions at signalized intersections. It is a very efficient way to move and manage through traffic stream to create a completely new timing sequence, that is customized to current conditions.
- There must be provision of separate bicycle lanes to avoid the delay in through movement of mainstream vehicles. Besides, the illegal movement of vehicles or pedestrians, oppositely coming on the sides of drive ways should be restricted by imposing strict rules by the city authorities to give rise to better service provision at signalized intersections.

CONCLUSION

This section briefly summarizes all the research practices carried out to evaluate service quality and to address the improvement issues of signalized intersections. The present research has justified the existing approaches, by developing new ALOS models taking a broad spectrum of geometrical, traffic operational, built-environmental and behavioral data sets are collected from a widely varying driving environment through field investigations, videography techniques, and perception survey. PHV/W_e, d, PCI, Q_l, O_{NV} and O_E are the resulting significant variables found out from Spearman's correlation, which have significant influence on drivers' riding quality at signalized intersections. DE model is found out to be suitable in defining ALOS scores of signalized intersections with R^2 values of 0.93 and 0.89 for training and testing datasets respectively. The resulting $ALOS_{sig}$ anticipated by approaches of intersections shows that, only 13% of the studied segments are labelled under ALOS category "A". Most of the intersections are observed to provide inferior service quality, which have vital issues related to the provided service quality. The relative importance of each service attribute is determined with the help of model sensitivity analysis, and ranked in order of specific percentage value based on their degree of relative importance. Q_l plays a significant role in fixing ALOS standards of signalized intersections, which have highest negative influence of 67.153%. Hence, optimizing traffic signalization timings and increasing effective green time for major approaches of an intersection in peak hours will significantly enhance the service quality of respective intersections. Similarly, other parameters are ranked in decreasing order of their relative importance. Based on the sensitivity analysis report, several improvement strategies are outlined in this study, which will help the transport authorities to identify operational issues of existing street facilities, and to design a users' friendly transport system with better driving environment.

The proposed model offers some new insights of traveler satisfaction level under highly heterogeneous traffic flow conditions and got scope of wide application in developing countries like India. However, traffic characteristics are somewhat different for developed countries and users' perceived satisfaction scores may vary accordingly. Likewise, this approach needs some modifications by iteratively changing the input parameters in varying road and traffic conditions for effective assessment of the service quality of urban street infrastructures in a global scenario.

REFERENCES

1. *Highway Capacity Manual*. Transportation Research Board, Washington, D.C., No. 1134, 2000.
2. *Highway Capacity Manual*. Transportation Research Board, Washington, D.C., No. 1650, 2010.
3. Sutaria, T. and J. Haynes. (1977). Level of service at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, No. 644, 1977.
4. Pecheux, K. K., M. T. Pietrucha, and P. P. Jovanis. User perception of level of service at signalized intersections: Methodological issues. *Transportation Research Circular E-C018: Proceedings of the Fourth International Symposium on Highway Capacity,* 2000, pp. 322–335.
5. Zhang, G., J. Chen, and J. Zhao. Safety performance evaluation of a three-leg unsignalized intersection using traffic conflict analysis. *Mathematical Problems in Engineering*, 2017. 10.1155/2017/2948750.
6. Zhang, L., A. R. Archilla, and P. D. Prevedouros. Ordered probit modeling of user perceptions of protected left-turn signals. *Journal of transportation engineering*, Vol. 133(3), 2007, pp. 205–214.
7. Lee, D., T. G. Kim, M. Pietrucha. Incorporation of transportation user perception into evaluation of service quality of signalised intersections. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2027, 2007, pp. 9-18. https://doi.org/10.3141/2027-02.
8. Fang, F. C. and K. K. Pecheux. Fuzzy data mining approach for quantifying signalized intersection level of services based on user perceptions. *Journal of Transportation Engineering*, Vol. 135(6), 2009, pp. 349–358. 10.1061/(ASCE)0733-947X(2009) 135:6(349).
9. Zhang, L. and P. D. Prevedouros. User perceptions of signalised intersection level of service using fuzzy logic. *Transportmetrica,* Vol. 7(4), 2011, pp. 279–296. 10.1080/18128601003667460.
10. Chen, P., J. Sun, and H. Qi. Estimation of delay variability at signalised intersections for urban arterial performance evaluation. *Journal of Intelligent Transportation Systems*, Vol. 21(2), 2017, pp. 94–110. 10.1080/15472450.2016.1216319.
11. Saha, A., S. Chandra, and I. Ghosh. Delay at signalized intersections under mixed traffic conditions. *Journal of Transportation Engineering, Part A: Systems*, Vol. 143(8), 2017. 10.1061/JTEPBS.0000070.
12. Ban, X. J., P. Hao, and Z. Sun. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transportation Research Part C: Emerging Technologies*, Vol. 19(6), 2011, pp.1133–1156. 10.1016/j.trc.2011.01.002.
13. Chang, J., M. Talas. and S. Muthuswamy. Simple methodology for estimating queue lengths at signalized intersections using detector data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2355, 2013, pp. 31–38. 10.3141/2355-04.

14. Indian Road Congress: Guidelines for capacity of urban roads in plain areas. *IRC code of practice*, 106, 1990.
15. Jena, S., A. Chakraborty, and P.K. Bhuyan. Performance assessment of urban streets addressing improvement issues for automobile mode of transport. *Transportation Research Record: Journal of the Transportation Research Board*, 2018. https://doi.org/10.1177/0361198118782761.
16. Searson, D. P., D. E. Leahy, and M. J. Willis. GPTIPS: An Open Source Genetic Programming Toolbox from Multi-Gene Symbolic Regression. *International Multi-Conference of Engineers and Computer Scientists*, Vol. I, Hong Kong, 2010.
17. Reddy, M. J. and D. N. Kumar. Multiobjective differential evolution with application to reservoir system optimization. *Journal of Computing in Civil Engineering*, Vol. 21(2), 2007, pp.136–146. 10.1061/ (ASCE) 0887-3801(2007)21:2(136).
18. Cai, G., S. Liu, L. Tong, and G. Du. Assessment of direct CPT and CPTU methods for predicting the ultimate bearing capacity of single piles. *Engineering Geology, Elsevier*, Vol. 104, 2009, pp. 211-222.
19. Beura, S. K., and P. K. Bhuyan. Modeling quality of bicycle accommodations on urban road segments using functional networks and multivariate adaptive regression splines techniques. *Transportation Research Board's 96th Annual Meeting,* Washington D.C, 2017, 20p.
20. Gandomi, A. H., G. J. Yun, and A. H. Alavi. An evolutionary approach for modeling of shear strength of RC deep beams. *Materials and Structures*, Vol. 46, 2013, pp. 2109-2119. http://dx.doi.org/10.1617/s11527-013-0039-z.