

- the dry zone of Sri Lanka. *Int. J. Primatol.*, 2012, **33**(6), 1382–1405.
48. Campos-Arceiz, A., The ecological role of Asian elephants as agents of seed dispersal in human-dominated landscapes. Doctoral dissertation, The University of Tokyo, 2009.
  49. Shoshani, J., Understanding proboscidean evolution: a formidable task. *Trends Ecol. Evol.*, 1998, **13**(12), 480–487.
  50. Babweteera, F., Savill, P. and Brown, N., *Balanites wilsoniana*: regeneration with and without elephants. *Biol. Conserv.*, 2007, **134**(1), 40–47.
  51. Krefting, L. W. and Roe, E. I., The role of some birds and mammals in seed germination. *Ecol. Monogr.*, 1949, **19**(3), 269–286.
  52. Traveset, A. and Verdu, M., A meta-analysis of the effect of gut treatment on seed germination. In *Seed Dispersal and Frugivory: Ecology, Evolution, and Conservation*, CABI, 2002, vol. 22, pp. 339–350.
  53. Prasad, P. and Nautiyal, A. R., Physiology of germination in *Bauhinia*: involvement of seed coat in inhibition of germination in *B. racemosa* Lam. seeds. *Seed Sci. Technol. (Switzerland)*, 1996, **24**(2), 305–308.
  54. Traveset, A., Effect of seed passage through vertebrate frugivores' guts on germination: a review. *Perspect. Plant Ecol. Evol. Syst.*, 1998, **1**(2), 151–190.
  55. Weaver, S. E. and Cavers, P. B., The effects of date of emergence and emergence order on seedling survival rates in *Rumex crispus* and *R. obtusifolius*. *Can. J. Bot.*, 1979, **57**, 730–738.
  56. Fernando, P. and Leimgruber, P., Asian elephants and seasonally dry forests. *The Ecology and Conservation of Seasonally Dry Forests in Asia*, Smithsonian Institution Scholarly Press, Washington, DC, 2011, pp. 151–163.
  57. Richardson, D. M., Allsopp, N., D'antonio, C. M., Milton, S. J. and Rejmánek, M., Plant invasions – the role of mutualisms. *Biol. Rev.*, 2000, **75**(1), 65–93.
  58. Bal, P., Nath, C. D., Nanaya, K. M., Kushalappa, C. G. and Garcia, C., Elephants also like coffee: trends and drivers of human–elephant conflicts in coffee agroforestry landscapes of Kodagu, Western Ghats, India. *Environ. Manage.*, 2011, **47**(5), 789–801.

**ACKNOWLEDGEMENTS.** We thank the Department of Wildlife Conservation in Sri Lanka for permission to conduct this research. We also thank Krishantha Sameera de Zoysa and others who assisted with field work, and Ringling Bros (Center for Elephant Conservation) for funds.

Received 17 October 2018; revised accepted 30 September 2019

doi: 10.18520/cs/v118/i4/648-654

## Reliability modelling on drivers' decision during the yellow phase of a signal intersection

Sabyasachi Biswas<sup>1,\*</sup> and Indrajit Ghosh<sup>2</sup>

<sup>1</sup>Department of Civil Engineering, National Institute of Technology, Jamshedpur 832 014, India

<sup>2</sup>Department of Civil Engineering, Indian Institute of Technology Roorkee, Roorkee 247 667, India

**During the yellow phase starts of a traffic signal, the drivers are unable to take quick decisions whether to stop or cross the signal-controlled intersection. This dilemma zone (DZ) can cause a mix-up among drivers during the yellow phase and may lead to accidents. In the present study we use a DZ prediction model for analysis. In this study, approach speed, acceleration/deceleration of different vehicle categories, distance to stop line, offset of yellow time, and presence or absence of countdown timer are considered as the main factors influencing the model. In order to address several drawbacks associated with traditional regression-based models, a kriging-based surrogate model has been developed to explore the drivers' behaviour during the yellow phase.**

**Keywords:** Dilemma zone, kriging model, traffic intersection, yellow phase drivers.

IN India, the highest number of road casualties occur at signal intersections. According to a Ministry of Road Transport and Highways (MoRTH), Government of India report road accidents account for 37.8% of total accidents<sup>1</sup>. Previous studies observed that the dilemma zone (DZ) is one of the most significant parameters affecting the safety at an intersection<sup>2,3</sup>.

A driver needs to take quick decision based on his experience while approaching the signalized intersections during the initial yellow phase. He needs to decide whether to cross the intersection or halt before the stop line. The virtual area immediately before the stop line is the DZ<sup>4</sup>, where drivers get confused whether to cross or stop at the intersection. In addition, drivers have a small response time to react. This results in red-light violations (RLVs) leading to safety problems at the intersections during the yellow phase.

To reduce problems associated with the DZ, development of drivers' decision prediction models has gained huge attention in the last few decades. The evolution of intelligent transportation systems (ITSs) has led to the effectiveness of these models. The system of sensors can be used in these models to accurately predict the drivers' decision at the intersections and prevent accidents like rear-end and angular collisions.

\*For correspondence. (e-mail: sabyasachibiswas01@gmail.com)

This study was conducted to identify the factors affecting drivers' behaviour during the yellow phase by considering the presence of various vehicle types. Also, in developing countries like India, there is less adherence towards lane discipline, which proliferates the difficulty in prediction modelling. In this study, a surrogate modelling technique is used to examine the factors influencing the drivers' decision-making process. A surrogate technique develops an approximate model by linking the inputs and outputs. It is an indirect technique to determine the desired output at unknown inputs. The main advantage of the surrogate model is its computational efficiency. In the present study, the surrogate model that has been used for studying drivers' decision in the yellow phase is based on the kriging approach. To the best of our knowledge, there is no previous study dealing with the uncertainty of the drivers' decision model using kriging approach.

Gazis *et al.*<sup>4</sup> developed a deterministic model to express the kinematic DZ based on drivers' decision during the yellow phase. Zegeer<sup>5</sup> reported another method (DZ II) to calculate DZ at signalized intersections. Several studies have been conducted on driver behaviour at the signalized intersections by taking into account many factors<sup>6-9</sup>. Some studies reported that few drivers utilize the yellow phase to change the DZ at signal-controlled intersections<sup>8,10,11</sup>. The presence of signal countdown timer (SCT) is another major factor, which influences the DZ<sup>12-16</sup>. The flashing green encourages drivers to make a decision, on whether to cross or halt at the stop line<sup>17</sup>. Overall, the uncertainty of drivers' behaviour is quite challenging to compute in traditional DZ models. Some researchers developed a stop-go model considering the dynamic nature of the DZ at signalized intersection<sup>2,3,18-21</sup>. The literature review shows that different methods can be used to explain the drivers' behaviour at offset of yellow phase. Among them, regression techniques are the most frequently used methods<sup>2,3,8,17,22</sup>. However, Hurwitz *et al.*<sup>23</sup> studied the drivers' behaviour during the yellow phase using the fuzzy logic technique. Biswas and Ghosh<sup>2</sup> conducted a comparison study on the stop-go model using various methods, namely logistic regression, ANN and fuzzy logic technique. They proposed a new methodology to combine the two best models (based on their prediction accuracy) by the weighted average technique.

It is well-known that various factors affect the drivers' decision-making process at the signalized intersections, namely speed, acceleration/deceleration, time to stop line (TTS), distance to stop line (DTS), age of the driver, presence of countdown timer, type of vehicle, type of intersection, duration of yellow phase, green time ratio, RLV, driver reaction time, etc.<sup>3</sup>. Additionally, earlier studies mostly concentrated on the development of prediction modelling. Only a few of them analysed drivers' decision model under heterogeneous traffic conditions<sup>2,3</sup>.

Furthermore, regression technique is found to be the most common method used for this purpose. The regression-based stop-go model is easy to understand and hence, suitable for on-field applications. However, for these regression-based studies, it is simply assumed that the dependent variable is linearly associated with the independent variables. In this process, the least square method is used to minimize the sum of the square of the variance between predicted and observed values. Consequently, the present study aims at developing a stop-go prediction model which addresses several issues associated with the regression model and appropriate for mixed traffic situation.

Eight different approaches at six typical four-legged signalized intersections were selected for data collection based on the criteria suggested by earlier studies<sup>8,24,25</sup>. It was observed that last-to-go and first-to-stop vehicles during the yellow phase at the intersections need to be considered for DZ analysis. So, all the vehicles were segregated into different types in the heterogeneous traffic stream. Six intersections with a large number of passenger cars and two-wheelers (either last-to-go or first-to-stop vehicles) posted speed limit 50 km/h from Delhi and Chandigarh, India were considered in this study (Table 1).

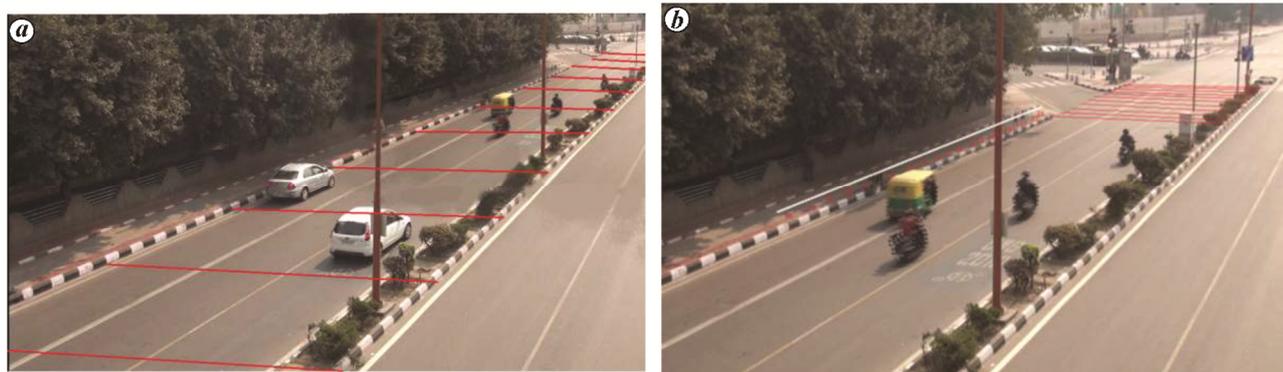
Initially, the videography technique was utilized for data collection at selective approaches of each of the intersections. At some sites, it was very difficult to capture a clear view of the approaches to cover at least 100 m upstream from the zebra crossing due to the presence of big trees, visual obstacles and security reasons. So, based on suitability, one camera was installed at a certain vantage point to capture the traffic movement, traffic signal changes and exact location of the vehicles at the approaching lane. It was observed that the yellow phase varied between 4 and 5 sec at all the selected intersections during peak as well as non-peak hours (data were mostly collected during 10 am to 3 pm). Six speed traps were virtually marked on the selected approaching road, each at 10 m interval. The second camera captured the traffic movement from different angles along with the SCT at the signalized intersections.

The video-recorded data were brought to the laboratory for extraction and further analysis. Accordingly, drivers' behaviour during the yellow phase and the each speed of approaching vehicle were mined from the video data.

In order to measure the approach speed, the approach road was divided into speed traps by virtual marking of 10 m interval (Figure 1a). The approach speed (km/h) was measured using the travel time of individual vehicles to cross the speed trap. The displacement-time relationship was used to calculate the acceleration/deceleration of the vehicles. These 10 m speed traps were additionally segmented into minor traps of 1 m interval to measure the DTS more precisely (Figure 1b), as recommended by previous studies<sup>2,8</sup>. On the other hand, TTS has been

**Table 1.** Detailed description of the selected intersections

Intersection ID	Location	Target approaches	Target movements	Approach width (m)	Cycle length (s)	Green time (sec)	Yellow time (sec)
Intersection 1 (IS1)	Lodhi Road and Bhisma Pitamah Marg	North-bound	Through and right turning	9	180	35	5
Intersection 2 (IS2)	Bhishma Pitamah Marg and Kendriya Vidyalaya Lodhi Road	West-bound	Through and right turning	9.6	165	45	5
Intersection 3 (IS3)	Manimajara and Chandigarh-Panchkula Road	West-bound	Through and right turning	14.6	135	35	4
Intersection 4 (IS4)	Chandigarh Rail Station Road and Chandigarh-Panchkula Road	West-bound	Through and right turning	13.6	149	29	4
Intersection 5 (IS5)	Zakir Chowk intersection	East-bound	Through and right turning	9	200	45	5
Intersection 6 (IS6)	Zakir Chowk intersection	West-bound	Through and right turning	9	200	45	5
Intersection 7 (IS7)	Pruthvi Raj Chowk second intersection	East-bound	Through and right turning	10.2	270	80	5
Intersection 8 (IS8)	Pruthvi Raj Chowk second intersection	North-bound	Through and right turning	7	270	45	5



**Figure 1.** *a*, Speed measurement trap on approach road at 10 m interval. *b*, Distance to stop line measurement trap on the same approach road at 1 m interval.

measured from the video. TTS is the time instance during yellow phase when the vehicular wheel touched the stop line at selected approach. In this context, the last vehicle to cross the intersection during the yellow phase in a particular signal cycle is known as the last-to-go vehicle. However, the very first vehicle which can stop before the stop line after the onset of the yellow phase in each signal cycle is defined as the first-to-stop vehicle. In this study, we only considered the last-to-go and first-to-stop vehicles to understand the drivers' behaviour during the yellow phase<sup>8</sup>. Moreover, among the 1468 vehicles selected for analysis during the yellow phase, 96% were either cars or two-wheelers (65% cars and 31% two-wheelers respectively). The drivers' decision model has been developed based on 1000 data, while the model was validated by 468 external data.

Various parameters were obtained from the available data and detailed statistical analysis was carried out (Table 2).

It was observed from previous studies that logistic regression method was most extensively used to predict the drivers' decision during the yellow phase at signalized intersections<sup>2,3,8,17</sup>. In this regard, a binary logistic regression that can well formulate the drivers' decision at the onset of the yellow phase based on all influencing parameters, can be represented as

$$\ln \left[ \frac{P_i}{1 - P_i} \right] = \varepsilon + \alpha_i X_i, \tag{1}$$

where  $P_i$  denotes the chance of drivers crossing the signalized intersection during the yellow phase,  $\varepsilon$  is a constant,  $X_i$  is the influencing parameters of drivers' decision during the yellow phase, namely approaching speed of the vehicles, acceleration/deceleration, DTS, TTS, type of vehicle and SCT and  $\alpha_i$  are the corresponding coefficients of the influencing parameters.

**Table 2.** Analysis of the various influencing parameters

Parameters			Stopping vehicles						Passing vehicles				Comparison of
	K-S test	P-value	Cumulative value at percentiles				K-S test	P-value	Cumulative value at percentiles				stopping and
			15	50	85	95			15	50	85	95	passing vehicles
Approach speed	0.32	0	11	20.4	33.2	39.3	0.078	0	12.1	25	41.2	52.94	11.21
Acceleration/ Deceleration	0.891	0.701	1.4	2.79	4.4	5.57	0.76	0.61	-2.9	-0.3	2.1	4.41	9.16
Distance-to-stop line (DTS)	0.821	0.571	20	40.2	58	69.5	0.097	-	3.8	21.87	43.7	58.38	-9.13
Time-to-stop line (TTS)	0.96	0.62	1.7	2.41	2.6	4.34	0.93	0.525	0.73	2.83	4.56	5	10.07

#t-Test value at 99% confidence level.

It was observed that all these (above mentioned) independent influencing factors have significant impact on drivers' behaviour during the yellow phase. Therefore, logistic regression-based drivers' decision model was developed by considering these factors (at 95% confidence).

In the present study, we have developed a stop-go model at signalized intersections using kriging technique. Kriging has the following advantages compared to other conventional approaches like regression methods:

- (i) The functional form of kriging consists of two parts, one is the polynomial trend and the other is the Gaussian process. The Gaussian process minimizes the spatial distance between two points according to an autocorrelation function<sup>26</sup>.
- (ii) The kriging prediction is stochastic in nature, i.e. the variance of prediction is obtained at any prediction point. It should be noted that the prediction variance is a local error measure and hence significantly improves the approximation accuracy.
- (iii) The mean kriging estimator interpolates the data, i.e. the variance is zero at these points. Consequently, the kriging estimator is obtained as an unbiased linear combination of the observations with minimum variance and with significantly small dataset compared to the other conventional techniques<sup>27</sup>. To the best of our knowledge, no previous studies have used the kriging technique model drivers' decision of stop and go at signalized intersections.

In the present study, a prediction model was developed based on drivers' stop or go decision. Consequently, a linear relationship was established to minimize variance (prediction errors) for the unknown dataset. In this context, kriging is a well-known technique to optimize the prediction error. Wang and Kockelman<sup>28</sup> predicted annual average daily traffic using the kriging method.

In the present study, drivers' decision prediction at the signalized intersections can be expressed as follows

$$Y_0(x) = \sum_{i=1} a_i \mathbb{F}_i(x), \tag{2}$$

$$Y_0(x) = a_i \mathbb{F}_i + \varepsilon, \tag{3}$$

where  $Y_0(x)$  is an unknown function representing drivers' decision at the signalized intersections,  $\mathbb{F}_i(x) = [\mathbb{F}_1(x), \dots, \mathbb{F}_m(x)]^T \in \mathfrak{R}^m$  a known function in the regression model,  $a_i$  are the regression coefficient vectors and  $\mathbb{F}_i(x)$  denote the basic functions in the model.  $\varepsilon$  is expected to be a Gaussian stationary process and it can be estimated by eq. (4).

$$\text{Cov} = \sigma^2 \xi(x_i, x_j), \quad i, j = 1 \dots n, \tag{4}$$

where  $\xi(x)$  is a correlation function and  $\sigma^2$  is the process variance.

$$\xi = \begin{pmatrix} \xi(x_i^1, x_j^1) & \dots & \xi(x_i^1, x_j^n) \\ \vdots & \ddots & \vdots \\ \xi(x_i^n, x_j^1) & \dots & \xi(x_i^n, x_j^n) \end{pmatrix}. \tag{5}$$

The best linear unbiased prediction estimation,  $\hat{y}$  of the response  $y(x)$  at random values of  $x$ , is defined by eq. (6). The outcomes at the sample points are expressed as  $Y = \{y_1(x), y_2(x), \dots, y_n(x)\}$ .

$$\hat{y}(x) = \mathbb{F}^T(x)a + \xi^T \omega. \tag{6}$$

In eq. (6),  $\xi^T \omega$  is an interpolation of the residuals part in the regression model and  $\mathbb{F}^T(x)a$  is the function part of the model. Thus, all predicted data would be accurately anticipated. Where  $\omega$  is defined by the eq. (8). Now  $\xi^T(x)$  is a vector indicating the relationship between an indefinite set of values  $x$  and all known datasets for the model

$$\xi^T(x) = \{\xi(x_i, x_j^1), \dots, \xi(x_i, x_j^n)\}^T, \tag{7}$$

$$\omega = \xi^{-1}(Y - \mathbb{F}\hat{a}). \tag{8}$$

From the eqs (9) and (10), the unknown factors  $\hat{a}$  and  $\sigma^2$  can be determined

$$\hat{a} = (\mathbb{F}^T \xi^{-1} \mathbb{F})^{-1} \mathbb{F}^T \xi^{-1} Y, \tag{9}$$

$$\hat{\sigma}^2 = \frac{1}{n} (Y - \mathbb{F}\hat{a})^T \xi^{-1} (Y - \mathbb{F}\hat{a}), \tag{10}$$

where  $\mathbb{F}$  is the column vector containing the values of  $\mathbb{F}(x)$  evaluated at each sample value.

Various kriging methodologies have their own correlation functions, which leads to an accurate surrogate. Previous studies have used different correlation functions<sup>26</sup>

$$\xi(x_i, x_j) = \xi_n(\theta_n, x_i - x_j), \tag{11}$$

where  $\theta_n$  are the unknown correlation parameters, and  $(x_i^n - x_j^n)$  signifies the distance between two data points.

Several correlation functions, i.e. Gaussian, exponential, linear and generalized exponential correlation function have been used in this study. The expressions of the correlation functions are given below.

Gaussian:

$$\xi = \exp \left[ -\sum_{n=1} \theta_n |x_i^n - x_j^n|^2 \right]. \tag{12}$$

Exponential:

$$\xi = \exp \left[ -\sum_{n=1} \theta_n |x_i^n - x_j^n| \right]. \tag{13}$$

Linear:

$$\xi = \max \{0, 1 - \theta_n |x_i^n - x_j^n|\}. \tag{14}$$

Generalized exponential:

$$\xi = \exp \left[ -\sum_{n=1} \theta_n |x_i^n - x_j^n|^{\theta_{n+1}} \right], \quad 0 < \theta_{n+1} \leq 2. \tag{15}$$

In this study, maximum likelihood for  $\theta_n$  was computed as

$$\log \mathcal{L}(\theta, x_1, x_2 \dots x_n) = \sum_{i=1}^n \log f(x_i | \theta), \tag{16}$$

$$\log \mathcal{L} = \frac{n}{2} \log(2\pi) + \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log(|\xi|) + \frac{1}{2\sigma^2} (Y - \mathbb{F}\hat{a})^T \xi^{-1} (Y - \mathbb{F}\hat{a}). \tag{17}$$

In order to simplify the eq. (17), derivatives are taken with respect to  $\hat{a}$  and  $\sigma^2$ . Further, it was considered that the eq. (17) is equal to zero. The simplified equation is

$$\log \mathcal{L}(\theta_n) = \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log(|\xi|). \tag{18}$$

In the random process,  $\varepsilon$  was obtained from the proposed model, which helps improve the model predictions. Root mean square error (RMSE) was estimated as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \tag{19}$$

where  $y_i$  and  $\hat{y}_i$  are the vectors of the actual and predicted values respectively.

The drivers' decision model was externally validated using data collected from different intersections.

In the present study, data collected from the eight signalized intersections were analysed. It was found that the level of heterogeneity was similar in all of them. However, the vehicular composition of traffic was found to vary. In IS1, there was a higher percentage of cars, whereas in IS3 there was almost equal proportion of cars and two-wheelers. Similarly, there was total absence of auto rickshaws in IS2 during the yellow phase. These factors lead to more versatility in the collected data. Table 3 shows that the range for DTS and TTS is notably distinct for passing and stopping vehicles. On the other hand, the range for speed and acceleration overlaps.

The logistic regression based drivers' decision model to predict the drivers' behaviour during the offset of yellow phase is given in eq. (1). The probability of a drivers' decision to cross an intersection is anticipated using various influencing parameters. Table 3 presents the calculated model coefficients for the influencing parameters.

It was observed that the logistic model is usually recognized by models' chi-square value (for this model it is 9.153), odds ratio, Cox and Snell  $R$  square value and Nagelkerke  $R$  square value (Table 3). In order to interpret the coefficient of various influencing factors (Table 3) indicate that approach speed, acceleration or deceleration, DTS, TTS and vehicle type had significant impact on drivers' decision at the offset of the yellow phase. Gates and Noyce<sup>8</sup> have made similar observations on the drivers' decision model. However, SCT was not found to be significantly related to the drivers' decision during the yellow phase<sup>22</sup>. This proposed model indicates that crossing probability of the drivers increases with increase in

**Table 3.** Calculated coefficient for various influencing factors of drivers’ decision model based on logistic regression

	Coefficient ( $\alpha$ )	SE	Wald test	df	Sig.	Exp ( $\alpha$ ) or odds ratio
Model variables <sup>5</sup>						
DTS	-0.014	0.010	2.979	1	0.0016	1.014
TTS	-0.502	0.124	16.293	1	0.0000	1.652
Approach speed	0.027	0.008	10.490	1	0.0010	1.027
Acceleration/deceleration	0.106	0.050	4.500	1	0.0034	0.899
Signal countdown timer	0.097	0.250	0.151	1	0.698	1.102
Vehicle type	0.589	0.217	7.361	1	0.007	1.801
Constant	-1.829	0.536	11.648	1	0.001	0.161
-2 log likelihood	Cox and Snell <i>R</i> square		Nagelkerke <i>R</i> square			
524.584 <sup>a</sup>	0.518		0.618			

<sup>5</sup>Where approach speed of the vehicles (continuous variable), acceleration/deceleration (continuous variable), DTS (continuous variable), TTS (continuous variable), type of vehicle (binary variable, car = 1, two-wheeler = 0) and SCT (if SCT was on = 1, if SCT was off = 0).

**Table 4.** Proposed model validation dependent on correlation functions

Correlation functions	Field	Model prediction			Percentage accuracy*
		Go	Stop	Total	
Logistic regression	Go	<b>230</b>	34	264	87.12
	Stop	95	<b>109</b>	204	53.43
	Total	325	143	468	
	Overall accuracy		RMSE-0.92		72.43
Gaussian	Go	<b>251</b>	13	264	95.07
	Stop	3	<b>201</b>	204	98.52
	Total	254	214	468	
	Overall accuracy		RMSE-0.163 $R^2 = 0.93$		96.58
Exponential	Go	<b>231</b>	33	264	87.5
	Stop	10	<b>194</b>	204	95.09
	Total	241	227	468	
	Overall accuracy		RMSE-0.30 $R^2 = 0.8374$		90.81
Linear	Go	<b>217</b>	47	264	82.19
	Stop	11	<b>193</b>	204	94.60
	Total	228	240	468	
			$R^2 = 0.8357$		87.61

\*The number of accurate drivers’ decision prediction by the Kriging model.

$$\text{Percentage accuracy in model prediction} = \frac{\text{Number of accurate predictions by the model}}{\text{Number of actual drivers' decision}} \times 100.$$

approach speed and acceleration of the vehicles. The overall precision of the logistic regression-based drivers’ decision model was 72.43% under the heterogeneous traffic condition. However, the stopping probability anticipation on the vehicles during the yellow phase was poor in the present prediction model (53.43%).

Previously, several stop-go models were developed incorporating combinations of different parameters. The base kriging approach is typically used in cases where uncertainties of the stopping probability are involved.

The proposed model has a good *R*-square value (0.93) and low RMSE value (0.163) showing the goodness-of-fit for the model. Several correlation functions were used in this study, though the best results were obtained with the Gaussian correlation function (Table 4). The overall accuracy of the kriging model was 96.58% with high precision in both stop and go anticipation based on 468 external data, which were not used in model development. Table 4 gives the combination of the all four prediction matrices obtained from logistic regression method and

Kriging methods (with three various functions). Also, prediction for the stopping vehicles was equally good in Kriging base prediction models with three different functions. While, the passing decision predictions were comparatively low in the kriging models with exponential and linear correlation functions. Table 4 also shows the outcomes from the kriging models with Gaussian, exponential and linear functions.

Similarly, Gates and Noyce<sup>8</sup> show that the regression based driver’s decision model overall prediction was 82% during the validation process by 61 external data. Yang *et al.*<sup>20</sup> concluded that the fuzzy logic (82.5%)-based drivers’ decision model performed (in terms of accuracy) comparatively better than drivers’ decision by logistic regression (80.5%). Similar outcomes were found by Biswas and Ghosh<sup>2</sup>. They concluded that ANN and fuzzy logic technique-based drivers’ decision models perform better than the logistic regression model. However, this outcome contradicts the conclusion drawn by Shen and Wang<sup>14</sup>.

Figure 2 illustrates that the possibility of crossing increases as approach speed increases, TTS decreases, DTS decreases and acceleration increases.

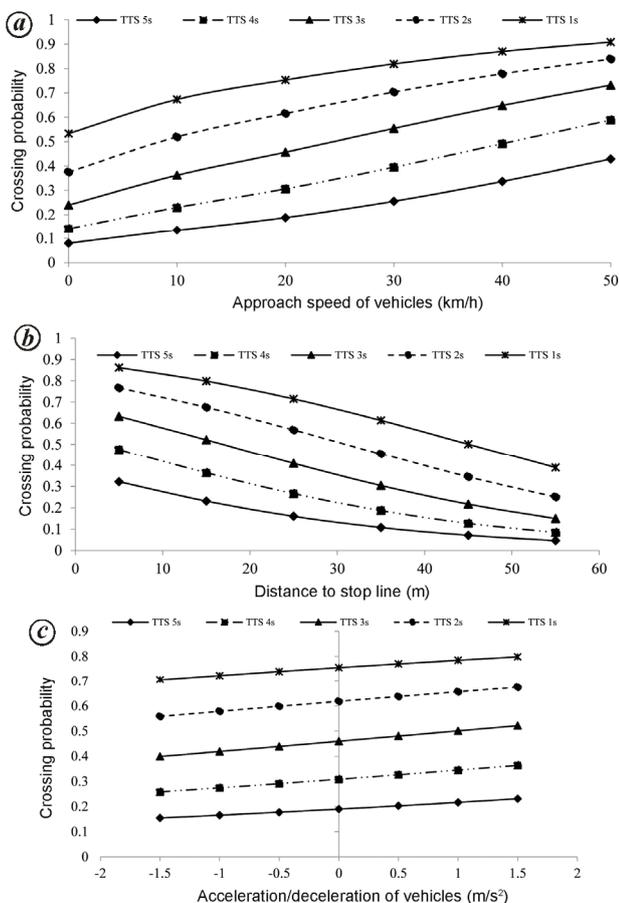


Figure 2. Variation of drivers’ decision model for crossing vehicles.

probability) of the crossing probability of DTS are found to be 17.3, 30.5, and 45.0 m, when TTS is 3, 2, and 1 s respectively. On the other hand, 0.5% of crossing probability is reached when the approach speed of vehicles is found to be 8.5, 24.0 and 40.5 km/h and TTS is 2, 3 and 4 s respectively.

The main aim of this study was to understand of the characteristics of drivers’ behaviour in the DZ under mixed traffic condition. In this regards, regression-based stop-go models are most commonly used for this purpose. While, a surrogate model using kriging approach has been utilized to develop drivers’ stop-go decision model due to several limitations of the regression model. Cars were found to be the most predominant vehicle category. It is interesting to note that characteristics of stopping are significantly different from crossing vehicles. The study compared the three most cited and widely used kriging correlation function models, with the same dataset.

The kriging method was highly accurate in predicting the stopping as well as moving vehicles. Also, correlation function is one of the major attributes in the prediction model. Three different correlation functions were used in this model. The Gaussian correlation function provides the best outcomes among them. Several other influencing parameters can be identified and added in the model to improve the prediction accuracy. Further research to develop surrogate models can be done utilizing other techniques apart from the kriging model.

The present study has been carried out in major cities in India, where non-motorized vehicles are restricted on major roads. On the other hand, smaller city roads have a high proportion of non-motorized vehicles. Hence, their effect along with the drivers’ gender and age may influence the drivers’ decision at signalized intersections, and these need to be explored in the future.

1. MoRTH, Road Accidents in India. Transport Research Wing, Ministry of Road Transport and Highway, Government of India, New Delhi, 2017.
2. Biswas, S. and Ghosh, I., Modeling of the drivers’ decision-making behavior during yellow phase. *KSCE J. Civ. Eng.*, 2018, **22**, 1–13.
3. Pathivada, B. K. and Perumal, V., Analyzing dilemma driver behavior at signalized intersection under mixed traffic conditions. *Transp. Res. Part F*, 2019, **60**, 111–120.
4. Gazis, D., Herman, R. and Maradudin, A., The problem of the amber signal light in traffic flow. *Oper. Res.*, 1960, **8**, 112–132.
5. Zegeer, C. V., Effectiveness of Green Extension Systems at High Speed Intersections, Lexington, KY, USA, 1977.
6. Parsonson, P. S., Roseveare, R. W. and Thomas, J. M., Small area detection at intersection approaches. *J. Transp. Eng.*, 1974, **44**, 8–17.
7. Papaioannou, P., Driver behaviour, dilemma zone and safety effects at urban signalized intersections in Greece. *Accid. Anal. Prev.*, 2007, **39**, 147–158.
8. Gates, T. J. and Noyce, D. A., Dilemma zone driver behavior as a function of vehicle type, time of day, and platooning. *Transp. Res. Rec. J. Transp. Res. Board*, 2010, 2149, Trans, 84–93.

9. Elmitiny, N., Yan, X., Radwan, E., Russo, C. and Nashar, D., Classification analysis of driver's stop/go decision and red-light running violation. *Accid. Anal. Prev.*, 2010, **42**, 101–111.
10. Wei, H., Li, Z., Yi, P. and Duemmel, K. R., Quantifying dynamic factors contributing to dilemma zone at high-speed signalized intersections. *Transp. Res. Rec. J. Transp. Res. Board*, 2011, **2259**, 202–212.
11. Rakha, H., El-Shawarby, I. and Setti, J. R., Characterizing driver behavior on signalized intersection approaches at the onset of a yellow-phase trigger. *IEEE Trans. Intell. Transp. Syst.*, 2007, **2259**, 630–640.
12. Chiou, Y. C. and Chang, C. H., Driver responses to green and red vehicular signal countdown displays: safety and efficiency aspects. *Accid. Anal. Prev.*, 2010, **42**, 1057–1065.
13. Long, K., Han, L. D., and Yang, Q., Effects of countdown timers on driver behavior after the yellow onset at Chinese intersections. *Traffic Inj. Prev.*, 2011, **12**, 538–544.
14. Shen, J. and Wang, W., Effects of flashing green on driver's stop/go decision at signalized intersection. *J. Cent. South Univ.*, 2015, **22**, 771–778.
15. Ma, W., Liu, Y. and Yang, X., Investigating the impacts of green signal countdown devices: empirical approach and case study in China. *J. Transp. Eng.*, 2010, **136**, 1049–1055.
16. Sharma, A., Vanajakshi, L., Girish, V., and Harshitha, M. S., Impact of signal timing information on safety and efficiency of signalized intersections. *J. Transp. Eng.*, 2012, **138**, 467–478.
17. Köll, H., Bader, M. and Axhausen, K. W., Driver behaviour during flashing green before amber a comparative study. *Accid. Anal. Prev.*, 2004, **36**, 273–280.
18. Easa, S. M., Reliability-based design of intergreen interval at traffic signals. *J. Transp. Eng.*, 1993, **119**(2), 255–271.
19. Li, K., Dong, S., Sun, J. and Yu, X., Study on the influence of signal countdown device on traffic safety of intersections. In IEEE International Conference on Measuring Technology and Mechatronics Automation, Zhangjiajie, Hunan, China, 2009, pp. 602–606.
20. Yang, Z., Tian, X., Wang, W., Zhou, X. and Liang, H., Research on driver behavior in yellow interval at signalized intersections. *Math. Probl. Eng.*, 2014, **2014**, 1–8.
21. Sharma, A., Bullock, D. and Peeta, S., Estimating dilemma zone hazard function at high speed isolated intersection. *Transp. Res. Part C*, 2010, **19**, 400–412.
22. Biswas, S., Ghosh, I. and Chandra, S., Influence of signal countdown timer on efficiency and safety at signalized intersections. *Can. J. Civ. Eng.*, 2017, **44**, 308–318.
23. Hurwitz, D. S., Wang, H., Knodler, M. A., Ni, D. and Moore, D., Fuzzy sets to describe driver behavior in the dilemma zone of high-speed signalized intersections. *Transp. Res. Part F*, 2012, **15**, 132–143.
24. Bonneson, J., Nevers, B., Zegeer, J., Nguyen, T. and Fong, T., Guidelines for quantifying the influence of area type and other factors on saturation flow rate. Texas Department of Transportation, The Texas A&M University System College Station, Texas, USA, 2005.
25. Biswas, S., Chakraborty, S., Ghosh, I. and Chandra, S., Saturation flow model for signalized intersection under mixed traffic condition. *Transp. Res. Rec.*, 2018, 0361198118777407.
26. Mukhopadhyay, T., Chakraborty, S., Dey, S., Adhikari, S. and Chowdhury, R., A critical assessment of kriging model variants for high-fidelity uncertainty quantification in dynamics of composite shells. *Arch. Comput. Methods Eng.*, 2016, **24**(3), 1–24.
27. Kaymaz, I., Application of kriging method to structural reliability problems. *Struct. Saf.*, 2005, **27**, 133–151.
28. Wang, X. and Kockelman, K. M., Application of the dynamic spatial ordered probit model: patterns of land development change in Austin, Texas. *Pap. Reg. Sci.*, 2009, **88**, 345–365.

Received 22 November 2017; revised accepted 9 October 2019

doi: 10.18520/cs/v118/i4/654-661