Transportation Research Part F 15 (2012) 132–143

Contents lists available at SciVerse ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Fuzzy sets to describe driver behavior in the dilemma zone of high-speed signalized intersections

David S. Hurwitz [a,*], Haizhong Wang [b], Michael A. Knodler Jr. [b], Daiheng Ni [b], Derek Moore [a]

[a] *School of Civil and Construction Engineering, Oregon State University, 220 Owen Hall, 1501 SW Campus Way, Corvallis, OR 97331, United States*

[b] *Department of Civil and Environmental Engineering, University of Massachusetts Amherst, 130 Natural Resources Road, Amherst, MA 01003, United States*

ARTICLE INFO

Article history:
Received 29 June 2010
Received in revised form 19 September 2011
Accepted 28 November 2011

Keywords:
Dilemma zone
Signalized intersections
Safety
Driver behavior

ABSTRACT

The Type II dilemma zone describes a segment of road on the approach to a signalized intersection where, if occupied by a motorist presented with the circular yellow indication, is likely to result in a motorist having difficulty deciding to stop at the stop line or proceed through the intersection. This phenomenon results in increased frequency of three failure conditions: rear-end collision at the stop line (excessive deceleration rates), the more severe right-angle crashes in the intersections, and left-turn head-on collisions (both resulting from incorrect estimates of clearance time). A more effective boundary definition for Type II dilemma zones could contribute to the safe design of signalized intersections. The prevailing approaches to dilemma zone delineation include the consideration of the vehicle's travel time to the stop line or the driver's likelihood of stopping at a particular distance from the stop line. The imprecision of the driver's perception of speed and distance suggest that fuzzy logic may contribute to the identification of the Type II dilemma zone boundaries. A fuzzy logic (FL) model was constructed and validated from driver's empirically observed behavior at high-speed signalized intersections. The research resulted in an increased understanding of the phenomenon which, when applied to the timing of signals and the placement of vehicle detection, can improve the overall safety of signalized intersections.

Published by Elsevier Ltd.

1. Introduction

The Type II dilemma zone describes a segment of road on the approach to a signalized intersection where, if occupied by a motorist presented with the circular yellow indication, is likely to result in a motorist having difficulty deciding to stop at the stop line or proceed through the intersection. This phenomenon results in increase frequency of three failure conditions: rear-end collision at the stop line (excessive deceleration rates), the more severe right-angle crashes in the intersections (incorrect estimates of intersection clearance time), and left-turn head-on collisions (incorrect estimates of clearance time). It should be noted that the Type II dilemma zone is not the only cause of these three failure conditions. Alternative causes are related to disobedience of the yellow and red indication or distracted drivers not detecting the yellow or red indications. The Type II dilemma zone conflict is believed to substantively contribute to the overall safety of signalized intersections, and particularly at high-speed signalized intersections where the severity of these types of crashes is worse due to the kinematics of the impulses exerted at higher speeds. Unfortunately, national standards have yet to be implemented to address this issue.

* Corresponding author. Tel.: +1 541 737 9242; fax: +1 541 737 3052.
E-mail address: david.hurwitz@oregonstate.edu (D.S. Hurwitz).

1369-8478/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.trf.2011.11.003

This research seeks to improve the way in which Type II dilemma zone boundaries have been previously identified. With improved boundary identification, more optimal change interval timing practices may result. This work contributes toward the information necessary for development of a standard procedure for change interval timing.

The Manual on Uniform Traffic Control Devices (MUTCDs) is the generally accepted authority on the application of traffic signs, signals, and pavement markings within the United States (MUTCD, 2003). The MUTCD provides information on the meaning, sequencing order, and ranges of duration for the circular yellow and red indication, but appropriately does not provide guidance on specific timing practices, since signal timing exists outside the influence of the MUTCD (MUTCD, 2003).

Given the lack of a design standard for calculation of change or clearance intervals, several approaches have been adopted by different agencies across the country. Arguably the most common approach in use is by the Institute of Transportation Engineers (ITEs), which developed a recommended calculation that accounts for grade of approach roadway, perception–reaction time of driver, deceleration rate of vehicle, velocity of approaching vehicle, length of vehicle, and the width of the intersection as described by Roess, Prassas, and McShane (2004) and in the Traffic Engineering Handbook (1999) and Institute of Transportation Engineers (1999).

Several alternative practices to the ITE recommended calculations have also been employed to address change and clearance interval timing. For intersections with relatively level approaches, some authorities calculate the yellow change interval as one tenth the operating speed of the approach vehicles, with an arbitrary red clearance interval of 1 or 2 s. Other agencies may apply the same change and clearance duration to roads of similar functional classification or closely grouped intersections as described in the Institute of Transportation Engineers' Traffic Signal Clearance Indication Course Material (2004).

As evidenced in the above, the lack of a national standard for the timing of the change and clearance intervals has resulted in a variety of strategies being implemented across the county. This lack of uniformity, when considered in conjunction with the wide variety of vehicle and user composition interacting on today's roadways, requires an updated examination of the dilemma zone issue. More specifically, there is a critical need to establish a consistent definition within this document for the purposes of understanding and accurately describing the nature of the dilemma zone conflict (Gates, Noyce, & Larauente, 2007; Urbanik & Koonce, 2007). Since Type II dilemma zones are important to the timing of change and clearance intervals, a more accurate definition of the Type II dilemma zone would contribute significantly towards improvements in change and clearance interval timing practices.

This research initiative seeks to improve upon our ability to identify the boundaries of the Type II dilemma zone region as well as the range of driver behavior within the Type II dilemma zone. To accomplish this task, fuzzy logic (FL) is proposed as an analytic tool. FL is applied to empirically observed dilemma zone interactions at high-speed signalized intersections to build a model of the location of the Type II dilemma zone. This proposed model is then validated against a separate set of empirically observed driver behaviors. Finally, the potential of the validated model is considered as a means to improve the design of high-speed signalized intersections.

2. Literature review

2.1. Defining two types of dilemma zone problems

The development of successful design solutions to transportation problems, or any other complex system, can be greatly hindered by poor problem identification. Such has been the case in the diagnosing of dilemma zone issues at signalized intersections. It is critical that a common definition be established if this traffic safety issue is to be adequately addressed. This document, building on previously established terminology (Gates et al., 2007; Urbanik & Koonce, 2007), will refer to 2 general classes of dilemma zone conflicts (Types I and II).

The Type I dilemma zone was first referenced in the literature by Gazis, Herman, and Maradudin (1960). It describes the possibility that a motorist presented with a yellow indication while approaching a signalized intersection will, due to the physical parameters of the situation, be unable to safely pass through the intersection or stop prior to the stop line. This scenario is the result of poor intersection design associated with errors in signal timing and detector placement. Several site-specific characteristics can contribute to these errors, including but not limited to, the distribution of vehicle type in the traffic stream, the grade and operating speed of the approach, and the available stopping sight distance.

It was not until 1974 that the Type II dilemma zone was formally identified in a technical committee report produced by the Southern Section of ITE (Parsonson, 1974). The Type II dilemma zone has also been termed an "indecision zone" which reflects the dynamic and probabilistic nature of Type II dilemma zone (Gates et al., 2007). The Types I and II dilemma zones are both depicted in Fig. 1.

2.2. Existing approaches to dilemma zone description

Several attempts have been made to quantify the location of the Type II dilemma zone. In 1978, Zegeer and Deen defined the boundaries of the Type II dilemma zone in terms of driver decision making. They identified the beginning of the zone as occurring at the position where 90% of drivers stopped and the end of the zone as occurring where only 10% of the drivers stopped (Zeeger & Deen, 1978). This definition agreed with the previous works of May (1968) and Herman, Olson, and Rothery (1963). Chang, Messer, and Santiago (1985) tried to define the boundaries in terms of the travel time to the stop

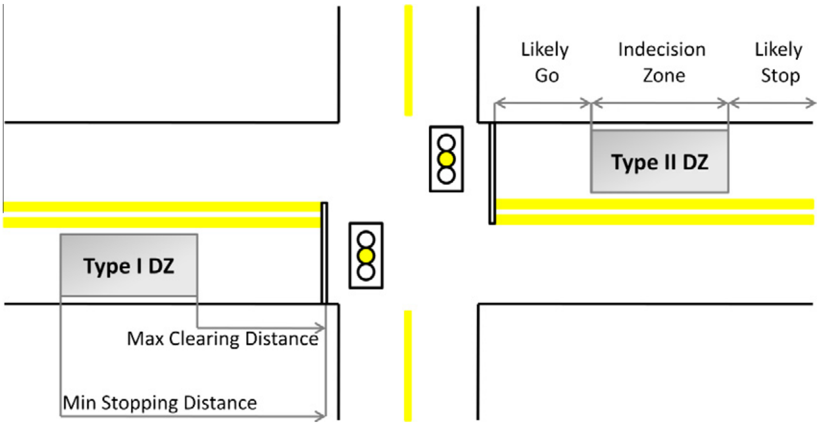

Fig. 1. Types I and II dilemma zone diagram.

line as estimated from the onset of the yellow indication. The research found that 85% of drivers stopped if they were 3 s or more back from the stop line while almost all drivers continued through the intersection if they were two seconds or less from the stop line (Chang et al., 1985). Supporting examples of defining the Type II dilemma zone in relation to the stop line can be seen in the works of Webster and Elison (1965) and Bonneson, McCoy, & Moen, 1994. Based on several of the previously mentioned studies it has been concluded that, as a rule of thumb, the Type II dilemma zone generally exists in the area between 5.5 s and 2.5 s from the stop line as measured from the onset of the yellow indication (Bonneson, Middleton, Zimmerman, Charara, & Abbas, 2002).

The three crash situations associated with dilemma zones are abrupt stops leading to rear-end crashes, failure to stop leading to right-angle crashes, and left-turn head-on crashes from incorrectly judging the clearance distance. On average, right-angle crashes tend to result in more serious injuries, therefore more emphasis is typically placed on their prevention. As the approach speeds of the intersecting roadways increase so too does the severity of the collisions, which is one reason why an added emphasis is placed on dilemma zone issues at high-speed signalized intersections.

2.3. Comments on existing dilemma zone definitions

The currently accepted definition of the Type II dilemma zone relies on the measure of a vehicles time to the stop line measured from the onset of the circular yellow indication with approximate boundaries reported from 2.5 to 5.5 s. Existing research suggests that this definition captures the two most critical factors affecting the driver behavior upon exposure to the circular yellow indication, vehicle position and vehicle approach speed. Additionally, the time to stop line measure provides a metric for identifying the approximate location of the Type II dilemma zone on the approach to a signalized intersection. This definition does not however account for the imprecision with which the driver perceives measures such as speed and distance at the instant the circular yellow is presented. It also does not provide additional information as to how driver behavior varies within the Type II dilemma zone region. For these reasons alternative boundary identification strategies should be considered.

3. Methodology

3.1. Fuzzy set theory for generic model development

Several research efforts have focused on the use of fuzzy sets or fuzzy logic (FL) as a tool for modeling vehicle interactions with traffic control devices such as signalized intersections. As an analytical tool, FL allows for the modeling of imprecise information. Driver decision making at a signalized intersection requires the estimation of vehicle position relative to the stop line, the speed and acceleration/deceleration capabilities of the vehicle, and the duration of the current indication. These quantities are continuously approximated by the approaching driver, and are therefore ideal fuzzy sets to be modeled with FL. (Kuo, Chen, & Hwang, 1996) used FL as a new mechanism for the calculation of change and clearance intervals. The new approach results in the determination of dynamic values for change and clearance intervals based on specific intersection and geometric conditions. Recently, (Rakha, El-Shawarby, & Setti, 2007) performed a field study involving 60 participants to characterize the driver behaviors at the onset of yellow indication. They considered the uncertainty and anxiety in this decision-making process which was quantified in the choice situation by Yager and reviewed by Rakha et al. (2007) given as:

$$A = 1 - \int_{x}^{\alpha_{\max}} \frac{1}{|A_\alpha|} d\alpha \tag{1}$$

where A is the level of uncertainty, and A_α is the number of alternatives whose choice probability is greater than α. In the case of driver choices within a dilemma zone, there are only two alternatives: either stop or go (Rakha et al., 2007). The quantification of uncertainty is reduced to:

$$A = 1 - \max(P_s, P_G) + \frac{1}{2}\min(P_s, P_G) \tag{2}$$

In which P_s and P_G are the possibility of stopping or going.

3.2. Formulation of Type II dilemma zone problem using Type 2 fuzzy set

In this paper, driver behavior in Type II dilemma zone is modeled as an indeterminacy phenomenon which essentially has two aspects: uncertainty and vagueness. The uncertainty mainly comes from different groups of drivers with varying indecisive driving behaviors during the change interval when approaching high-speed intersections. The indecisiveness involves two choices: either to stop at the stop line or continue through the intersection.

The vagueness phenomenon, different from the uncertainty, is an instance of second-order uncertainty arising when trying to group objects with a certain property. We claim that the fuzzy set of Type 2 theory is a reasonable mathematical description of the vagueness phenomenon. Type 2 fuzzy sets, sets with fuzzy grades of membership, generalize the Type 1 fuzzy set in order to handle more uncertainty in the decision-making process.

There are three primary approaches in the literature which describe how to establish the fuzzy-set membership function namely: declarative approach, computational approach, and modelization approach (Novak, 2006). A fuzzy set is a group of objects with a continuous grade of membership (Zadeh, 1965) characterized by a membership function ranging between 0 and 1. It can be constructed by assigning a membership value to each object in the interval of $[0,1]$. Membership values indicate the degree to which an object belongs to a fuzzy set. Let X denote the vehicle position during the change interval when approaching the high-speed intersection. With the decrease of the distance to the stop line, the driver behavior in the decision-making choice process varies. The drivers have to make a decision on either to stop at the stop line or proceed through the yellow light. The originally simple driver behavior becomes complex at a particular distance during the onset of the change interval because the driver may only know the duration of change interval approximately.

A fuzzy set is characterized by its membership function as seen in:

$$\mu_A(x) : X \to [0,1] \tag{3}$$

The fuzzy set A can be defined as a set of ordered pairs as seen in:

$$A = \{(x, \mu_A(x)) | x \in \chi\} \tag{4}$$

Membership functions are usually formulated with idealized representation of straight lines. For practical purposes, triangular, trapezoidal, or Gaussian membership functions are utilized as idealized shapes. Specific to our case, we use triangular membership function.

3.3. Fuzzy boundary identification of dilemma zone at signalized intersection approach

3.3.1. Data collection methods

The data collection procedure included completing intersection inventories at each of the experimental locations, capturing video of vehicles interacting with the onset of the yellow indication, and conducting automated speed studies at the location of the advanced detector data. The collection of speed and video data were critical because individual vehicle speed and position are believed to impact the potential for conflicts during clearance intervals.

As with many experiments that incorporate field observation, the identification of adequate experimental sites was of crucial importance. Vermont Agency of Transportation (VTrans) engineers led the selection of the test sites based upon their knowledge of the operational and safety characteristics of the Vermont state highway system. Both major approaches of the following intersections, located in the municipalities of Berlin and Rutland, were included in the experiment:

- Route 62 at Paine Turnpike (eastbound and westbound approaches).
- Route 62 at Airport Road (eastbound and westbound approaches).
- Route 62 at Berlin Road (eastbound and westbound approaches).
- Route 7 at North Shrewsbury Road (northbound and southbound approaches).
- Route 7 at Route 103 (northbound and southbound approaches).

An intersection inventory was completed to help adequately describe some of the relevant geometric characteristics of each individual intersection approach. Aspects such as horizontal and vertical curvature, grade, clear zones, adjacent land use, and presence of guard rails were considered in the selection of appropriated data collection locations but not included in the modeling associated with this research initiative.

Observations of intersection operations and driver behavior were also conducted through the collection of video data. Cameras were unobtrusively mounted (15–20 ft off the ground) on a variety of fixed structures (500–600 ft back from the stop line) near the roadside. The cameras were oriented to face towards the signal heads on each major intersection approach. This system allowed for the clear identification of vehicle position and signal phase from a single location for a period of up to 4 h between tape changes. Fig. 2 depicts the installation of one such camera setup.

In order to effectively use the 8 mm video tapes to accurately identify the position of the vehicle at the onset of the solid yellow indication, the tapes were digitized and measurement points were transposed onto the digital files. The video file was then used in the dilemma zone and driver behavior analysis. Fig. 3 shows a still frame of a completed digital video file overlaid with 50 ft intervals extending back from the stop line for several 100 ft.

Once the 8 mm video tapes were digitized with the measurement zones in place, they were burned to CDs so that multiple researchers were able to reduce the data into Excel© spreadsheets simultaneously. A team of trained researchers collaborated on the reduction of the overall database. As a part of the training component, researchers reviewed the same video file to ensure consistent results across researchers. In addition, random files were watched by multiple researchers in an effort to ensure consistency and validation of the research findings.

3.3.2. Data collection results

Approximately 510 h of video-taped observation were collected across all 10 high-speed intersection approaches. Of this 510 h sample, approximately 75 h of video were reduced representing approximately 15% of the overall sample with an approximate range of 5–15 h per intersection approach. Reduced observations yielded a sample size of approximately 1900 vehicles which interacted with the change interval while approaching one of the signalized intersections from either direction on the main line.

Samples of the speed, position, and driver behavior data collected in the field from two locations (Rte 7 at N. Shrewsbury (SB), and Rte 62 at Paine Tpke (WB)) are presented in this section. Specifically, the cumulative frequency of approach speeds collected at the location of the advanced detector for a duration of 72 h and the cumulative frequency of vehicles choosing to stop in relation to their distance to the stop line at the onset of the yellow indication. Fig. 4 shows an empirical cumulative distribution collected from Route 62 at Paine Turnpike.

3.3.3. Generic membership function formulation

The development of a fuzzy membership function is an important step in evaluating fuzzy system applications. It has been a challenging issue to generate suitable membership function in generic terms by different fuzzy construction rules.

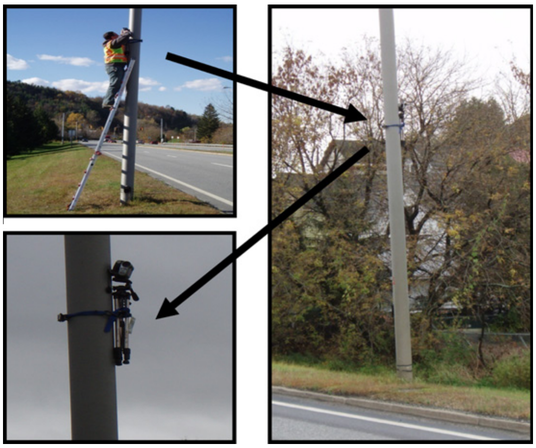

Fig. 2. Example of typical video camera installation.

Fig. 3. Digitized video with measurement zones.

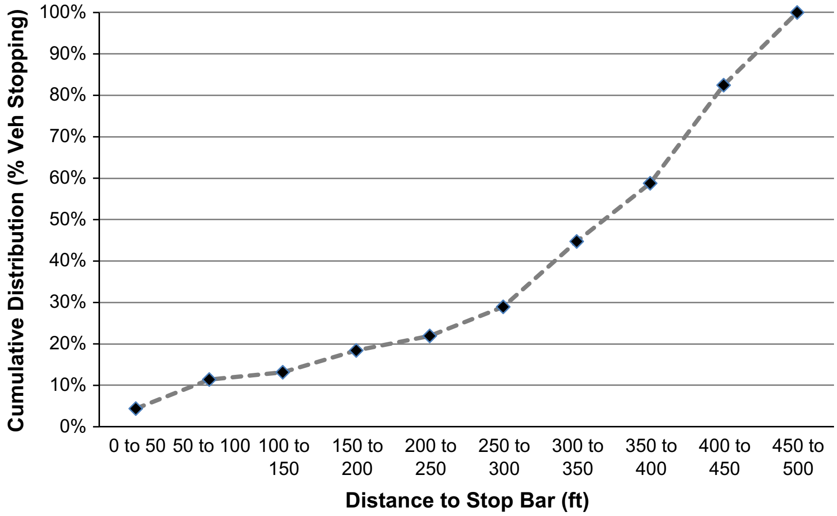

Fig. 4. Cumulative distribution of driver stopping behavior: Route 62@Paine Turnpike (WB).

There are many research efforts focusing on generating fuzzy rules automatically such as genetic algorithms, simulated annealing and Kalman filtering techniques. The shape of membership function, i.e. trapezoidal, triangular, or sigmoid, can be determined through a heuristic process. The generation of generic membership function can be realized by learning strategies using training data. A detailed description of generating membership function by different fuzzy construction rules is beyond the scope of this paper. For more information, interested readers are referred to (Makrehchi, Basir, & Kamel, 2003).

3.3.4. Example of membership function formulation for specific location

A fuzzy set can be fully defined by its membership function. How to determine the membership function is usually the first question that needs to be addressed. The adoption of any particular shape of membership function is often dependent on its applications. At most times, the assumption for fuzzy control problems is a linear membership function in a triangular shape because of its mathematical simplicity and easy implementation.

Fuzzy logic is a multi-valued logic with truth represented by a value on the closed interval, where 0 is equated with the classical false value and 1 is equated with the classical true value. Values in (0, 1) indicate varying degrees of truth. In the problem of Type II dilemma zones, the membership function of fuzzy subsets of vehicle position to the stop line is developed in Table 1. It is the construction of such bounding distributions that lead to validation. The concerned problem domain is 600 ft, and it is divided into 12 subsets with each interval 50 ft. These subsets are categorized by close distance (possible to go), medium distance (greatest uncertainty) and far distance (possible to stop) using different membership functions. For the close distance case, if the vehicle position to stop line is less than or equal to 50 ft, then it is definitely close with membership equal to 1. From 50 ft to 300 ft, the membership is given by the linear function as can be seen from Fig. 5. As the vehicle position to the stop line increases, the membership to classify its closeness decreases. When the vehicle position to stop line is greater than 300 ft, the membership function becomes 0. The same logic is applied to the medium distance and far distance. To clarify, a membership given by the membership function is not a probability, and it is not simply a quantitative variable at the interval level.

The membership function is a graphical representation of the magnitude of participation of each input variable. A weighting factor is associated to each of the input variables by membership function which describes the membership grade of the

Table 1
Fuzzy subsets and membership function for vehicle-position (VP).

Fuzzy subsets	Membership function
Close distance (possible to go)	$f(VP) = \begin{cases} 1.0 & VP \leqslant 50 \\ 1.2 - (\frac{1}{250})VP, & 50 < VP \leqslant 300 \\ 0 & 300 < VP \end{cases}$
Medium distance	$f(VP) = \begin{cases} 0 & VP \leqslant 50 \\ -0.33 + (\frac{1}{150})VP, & 50 < VP \leqslant 200 \\ 2.33 - (\frac{1}{150})VP, & 200 < VP \leqslant 350 \\ 0 & 350 < VP \end{cases}$
Far distance (possible to stop)	$f(VP) = \begin{cases} 0 & VP \leqslant 300 \\ -2.0 + (\frac{1}{150})VP, & 300 < VP \leqslant 450 \\ 1 & 450 < VP \end{cases}$

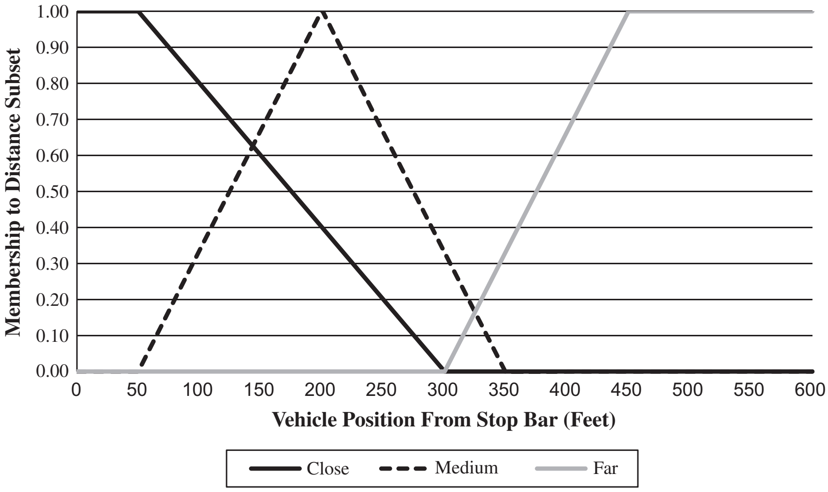

Fig. 5. The membership function of vehicle position.

elements in the fuzzy set. Take the membership function from the vehicle position to stop line in Fig. 5 as an example, the triangular membership function is formulated.

In the development process, we divided the vehicle position to stop line into three categories: close, medium, and far. The essence of fuzzy variable is its indeterminacy, which means the description of close, medium, or far is not a fixed number but an interval such as the interval we define to delineate close is (0,300), similarly (50,350) and (300,600) for medium and far distance. Obviously, there are some other combinations to describe the fuzziness, but the idea regarding how the fuzzy membership function is created remains unchanged.

3.3.5. Validation of membership function

In the domain of fuzzy set theory, validation is usually defined as bounding a fuzzy set by lower and upper functions which enclose fuzzy membership functions (Lodwick, Jamison, & Newman, 2001; Moore & Lodwick, 2003). The ways that validation can be used involving two situations: an uncertain construction of a membership function or known underlying membership function.

Given a fuzzy variable A, with associated membership function $\mu_A(x)$, we say A is enclosed if there exist bounding functions $p_A(x)$ and $n_A(x)$ satisfying $n_A(x) \leqslant \mu_A \leqslant p_A(x), \forall x$ and for all valid membership functions $\mu_A(x)$ of A. Given any fuzzy variable A with any associated valid membership function $\mu_A(x)$, the fuzzy variable A is validated if there exists a sequence of functions $p_k(x)$, and $n_k(x)$ that enclose $A(n_k(x) \leqslant \mu_A(x) \leqslant p_k(x),\ forall\ k)$ such that $p_k(x) \to \mu_A x\ and\ n_k(x) \to \mu_A(x)$.

It is obvious that membership functions $p(x) = 1$ and $n(x) = 0$ are enclosures for any fuzzy member and hence validate every real fuzzy variable. We seek the tightest bounding membership function. More advanced interval validation methods can be employed through constructing enclosures and bounding algorithms but it is not the main focus of this paper. Interested readers are directed to (Lodwick et al., 2001) for more information. The idea is that we can bound the results given probability distributions or membership functions.

3.3.6. Model formulation and validation

It is the intention of this research effort to incorporate the empirically generated fuzzy sets described in Table 1 as an input for a probabilistic model which will output the likelihood of a driver to stop at the stop line when presented the solid yellow indication. The idea of building a model to describe the probability of a vehicle stopping is not new, several attempts have been made using a variety of factors in a binary logistic regression model (Zhixia, Heng, Qingyi, & Zhuo, 2010; Gates et al., 2007; Chang et al., 1985). However, our approach of combining a fuzzy set of vehicle position data as a determinate of stopping probability is unique.

The model development and selection process was very much guided by the nature of this research problem. We are concerned with modeling the Type II dilemma zone problem as an indeterminacy problem using a probabilistic model that predicts driver stopping probability according to fuzzy inputs. To go or to stop is a decision which has to be made in a Type II dilemma zone. It is essentially a choice behavior which can be discrete or continuous. We are formulating this problem using a continuous approach. Generally, binary logistic models are appropriate to predict a dependent variable on the basis of continuous or categorical independents to assess the impact of independent control variables. Thus, a binary logistic regression model is adapted to model drivers' stopping probability in which "distance to stop line" is the only independent variable. The model takes the form presented in:

$$\begin{aligned} P_i(stop) &= \tfrac{1}{1+e^{-\beta_i}} \\ \beta_i &= a + b_0 Z_0 \end{aligned} \tag{5}$$

where p_i is the stopping probability of ith driver , β_i is a linear combination of multiple factors such as vehicle position etc. Z_0 is the vehicle's yellow onset distance.

The above model can be calibrated using field-observed trajectory data by performing regression analysis. The calibrated model can therefore be viewed as ground truth data when evaluating the fuzzy logic model of driver's stopping probability. The result of the regression analysis is summarized in Tables 2–4.

Table 2 shows the determination of the constants a and b_o for Eq. (5). It was determined that $a = -3.916$ and $b_o = 0.014$ with statistical significance ($P < 0.001$).

Table 3 shows the calculation of the R^2 values for the binary logistic regression. The Nagelkerke R^2 (0.581) is appropriate for this model. In many circumstances, desirable R^2 might be as high as 0.8 or 0.9. However, when dealing with field experimentation it can be difficult to achieve R^2 as high.

According to the classification analysis in Table 4, the distance to stop line is statistically significant by have a p-value less than 0.05. Therefore, the binary logistic regression model after calibration can be expressed with the following constants for a and b_o:

$$\begin{aligned} P_i(\text{stop}) &= \tfrac{1}{1+e^{-\beta_i}} \\ \beta_i &= -3.916 + 0.014Z_0 \end{aligned} \tag{5 updated}$$

On the other hand, when using the fuzzy membership value as the only independent variable, the driver's stopping probability can be represented by the following binary logistic model in:

$$\begin{aligned} P_i(\text{stop}) &= \tfrac{1}{1+e^{-\beta_i}} \\ \beta_i &= a + b_0 FMV \end{aligned} \tag{6}$$

where FMV is the fuzzy membership value from the defined membership function by vehicle's yellow-onset distance. The result of the regression analysis for Eq. (6) is summarized in Tables 5–7.

Table 2
Variables in the equation.

		B	SE	Wald	df	Sig.	Exp(*B*)
Step 1[a]	DIST	.014	.001	90.404	1	.000	1.014
	Constant	−3.916	.430	82.759	1	.000	.020

[a] Variable(s) entered on step 1: DIST.

Table 3
Model summary.

Step	−2 Log likelihood	Cox and Snell R^2	Nagelkerke R^2
1	255.190[a]	.432	.581

[a] Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Table 4
Classification table.

Observed		Predicted		
		Go = 0, Stop = 1		Percentage correct
		0	1	
Step 1				
Go = 0, Stop = 1	0	164	20	89.1
	1	34	102	75.0
Overall Percentage				83.1

* The cut value is .500.

Table 5
Variables in the equation.

		B	SE	Wald	df	Sig.	Exp(*B*)
Step 1[a]	FMV	−6.397	.737	75.368	1	.000	.002
	Constant	2.129	.273	60.684	1	.000	8.410

[a] Variable(s) entered on step 1: FMV.

Table 6
Model summary.

Step	−2 Log likelihood	Cox and Snell R^2	Nagelkerke R^2
1	266.783[a]	.411	.553

[a] Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Table 7
Classification table.

Observed		Predicted		
		Go = 0, stop = 1		Percentage correct
		0	1	
Step 1	0	151	33	82.1
Go = 0, stop = 1	1	15	121	89.0
Overall percentage				85.0

* The cut value is .500.

Table 5 shows the determination of the constants a and b_o for Eq. (6). It was determined that $a = -6.397$ and $b_o = 2.129$ with statistical significance ($P < 0.001$).

Table 6 shows the calculation of the R^2 values for the binary logistic regression. The Nagelkerke $R^2 = 0.553$ is appropriate for this model.

After calibration using regression analysis, Eq. (6) can be updated with constants for a and b_o to be the following expression, where FMV is identified to be a statistically significant variable.

$$P_i(stop) = \frac{1}{1+e^{-\beta_i}} \qquad \beta_i = 2.129 - 6.397FMV \tag{6 updated}$$

Fig. 6 shows an empirical cumulative distribution from Route 7 at North Shrewsbury, in which the input variable to the membership function is the distance to the stop line. The membership function outputs a fuzzy membership value which is factored as an input to the binary logistic regression model. A proper choice of the behavioral parameter α to the stopping probability model will reproduce the logistic relationship.

Fig. 7 shows the ultimate result of the calibrated binary logistic regression produced by the updated Eq. (6). This function is validated in Fig. 7 by comparing the ground truth (GT) and the fuzzy membership result (FMR). The GT probability function describes the actual observed stopping behavior of drivers related to the vehicle position at the onset of the CY indication from the data presented in Fig. 6, while the FMR result predicts the driver stopping behavior based on the Fuzzy Set prediction. A visual inspection of the comparison of the two models clearly shows that the FMR is very similar to the GT.

3.3.7. Comparison to previous results

As an additional measure of validation for our model, which uses fuzzy logic to characterize uncertainty, was compared to previous work. There was some limitation as to the number of previous research efforts that could be directly compared with

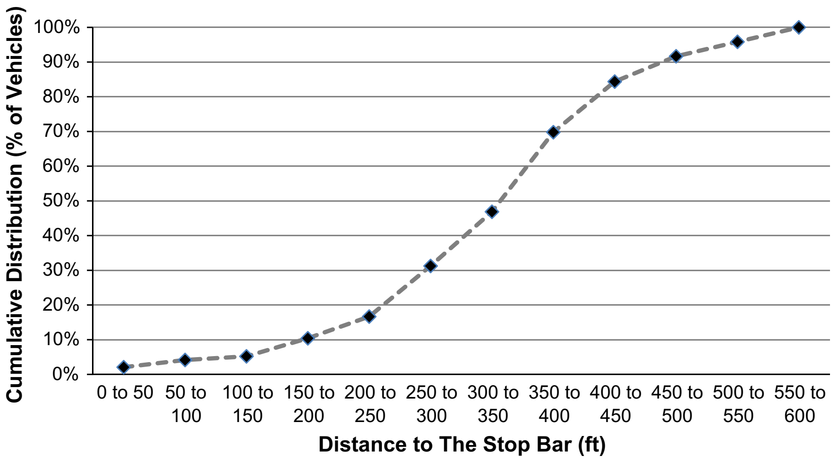

Fig. 6. Cumulative distribution of driver stopping behavior: Route 7@North Shrewsbury (SB).

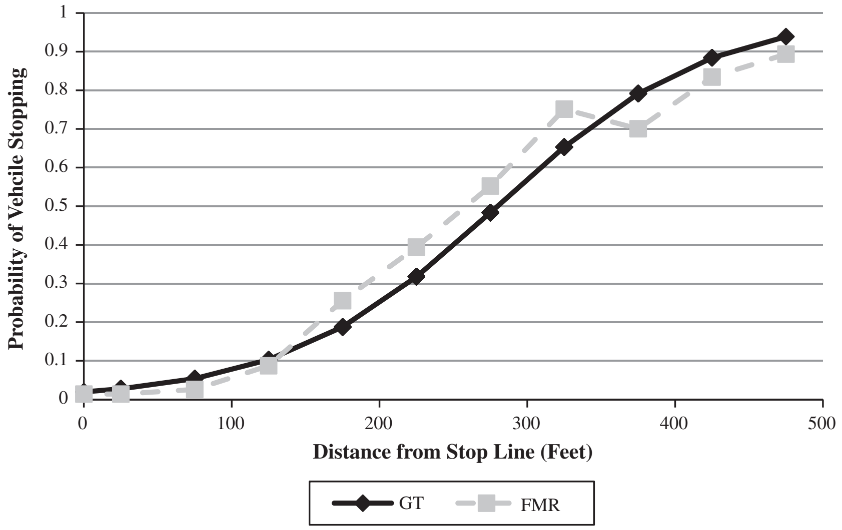

Fig. 7. Comparison of vehicle stopping probability based on its position at the onset of the CY for the ground truth and the fuzzy logic model result.

the model developed herein as a result of the need for consistent input variables (a critical aspect of comparison). However, the work of Rakha et al. (2007), previously discussed in this manuscript presents one model which can be directly compared. Fig. 8 displays the probability functions for running and stopping as presented from Raka et al. as well as from our FMR.

As presented in Fig. 8, the probability function developed by Rakha et al. terminates at approximately 360 ft from the stop line. The original Rakha et al. study measured distances in meters, but was converted to feet here to allow for a more direct comparison of the results. While our work carries driver behavior out to 475 ft. This difference can be attributed most directly to the operating speed of the approaches. Rakha et al. directed subjects on a closed course to travel at 45 mph while our naturalistic approaches operated at an 85th percentile speed of 57.5 mph. The higher speed results in a shift of the Type II dilemma zone appearing further away from the intersection.

3.3.8. Conclusion

To summarize, the authors have established that it is feasible to use fuzzy logic to delineate dilemma zones and by doing so some of the uncertainty involved in the delineation can be effectively captured. When applied to a specific intersection such an approach allows a better match to the "true dilemma zone" over the long run perceived by most drivers. This fuzzy approach was developed from a sample of field data collected in Vermont and validated off a second independent set of data also acquired in Vermont. When compared to the previous work of Rakha et al. the model is reasonably consistent.

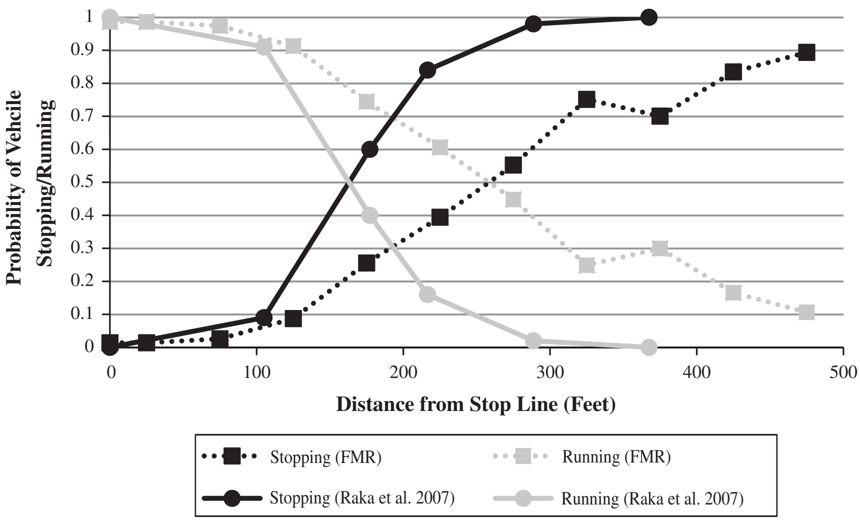

Fig. 8. Comparison of vehicle stopping/running probability based on its position at the onset of the CY for the fuzzy logic model result and Raka et al. (2007).

4. Contribution

This work supports, and builds upon previous research (Kuo et al., 1996) where similar membership functions were created. The distance to stop line membership function developed by Kuo et al. is comprised of three similar fuzzy subsets with slightly varying thresholds used to define the functions. The vehicle position membership function developed in this research builds upon previous research by capturing critical driver behavior patterns at higher speed locations where the consequences of driver failure are more severe.

This current research develops a binary logistic regression model for drivers stopping probability in a fashion similar to the work of Gates et al. (2007); however, the input to the model is achieved though a fuzzy subset which appears to result in a desirable prediction of driver behavior while being significantly less data dependent than many previous modeling efforts. The implications of the ability to predict driver behavior, while requiring less by way of data input has the potential to translate into the ability to model significantly more intersection approaches thus developing a better understanding of the entire phenomena.

Additionally, one model developed by Rakha et al. was compared to the FMR in Fig. 8 and is very similar in shape, although constructed for an intersection approach on a test track with a stipulated approach speed more than 10 mph less than the naturalistic observations used in the development of the FMR. Our work contributes to Rakha et al.'s effort as our model is provided for higher speed approaches where the consequences of Type II dilemma zones are even more severe.

There are three primary failure conditions resulting from the two types of dilemma zone scenarios. This research effort concerns itself specifically with the two of these failures (rear-end collision on the approach and right-angle collision in the intersection) associated with through traffic in a Type II dilemma zone. The fuzzy boundary identification of the Type II dilemma zone has the potential to improve safety at high speed signalized intersections by accurately predicting where drivers will have an increased difficulty in deciding how to react at the onset of the circular yellow indication. With this knowledge traffic engineers can ensure that signal timing and detection practices to improve the operations of high-speed signalized intersections. Future research work is planned to further test how this model can help to achieve this ultimate goal.

Disclosure

The authors have no conflict of interests associated with the research included in this manuscript.

Role of the funding source

The initial data collection for this research initiative was funded by the Vermont Agency of Transportation (VTrans). VTrans representatives participated in the identification and selection of the study locations and provided information on the existing conditions at the locations (i.e. grades of approach roads, signal timings, etc.). They have been made aware of the submission of this document and were permitted the opportunity to review and comment on the manuscript before submission.

References

Bonneson, J. A., McCoy, P. T., Moen, B. A. (1994). *Traffic detector design and evaluation guidelines (Report TRP-02-31-93)*. Lincoln, Nebraska: Department of Roads.

Bonneson, J. A., Middleton, D. R., Zimmerman, K. H., Charara, H. A., & Abbas (2002). *Intelligent detection-control system for rural signalized intersection (0-4022-2)*. Texas Transportation Institute.

Chang, M. S., Messer, C. J., & Santiago, A. J. (1985). Timing traffic signal change intervals based on driver behavior. *Transportation Research Record, 1027*, 20–30.

Gates, T. J., Noyce, D. A., & Larauente, L. (2007). Analysis of dilemma zone driver behavior at signalized intersections. In *Paper presented at the meeting of TRB*, Washington, DC.

Gazis, D. C., Herman, R., & Maradudin, A. (1960). The problem with the amber signal light in traffic flow. *Operations Research, 8*(1), 112–132.

Herman, R., Olson, P. L., & Rothery, R. W. (1963). Problem of the amber signal light. *Traffic Engineering and Control, 5*, 298–304.

Institute of Transportation Engineers (2004). *Traffic signal clearance indication: Course material*, Washington, DC.

Kuo, K. Y., Chen, Y. J., & Hwang, R. C. (1996). Calculation of the change and clearance intervals of traffic signal by fuzzy logic system. In *Paper presented at the meeting of joint conference of international computer symposium*, Kaohsiung, Taiwan.

Lodwick, W. A., Jamison, K. D., & Newman, F. D. (2001). Extension of interval validation methods to fuzzy set theory. In *Paper presented at the North American Fuzzy Information Processing Society*.

Makrehchi, Masoud, Basir, Otman, Kamel, Mohamed (2003). Generation of fuzzy membership function using information theory measures and genetic algorithm. *Lecture notes in computer science, fuzzy sets and systems – IFSA 2003* (Vol. 2715). Springer Berlin/Heidelberg.

May, A. D. (1968). Clearance interval at flashing systems. *Highway Research Record, 221*, 41–71.

Moore, R., & Lodwick, W. (2003). Interval analysis and fuzzy set theory. *Fuzzy Sets System, 135*(1), 5–9.

Manual on Uniform Traffic Control Devices (2003). *Federal Highway Administration*, US Department of Transportation, Washington, DC.

Novak, V. (2006). Fuzzy sets as a special mathematical model of vagueness phenomenon. In B. Reusch (Ed.), *Computational intelligence, theory and applications* (pp. 683–690). Dortmund, Germany: Springer Berlin Heidelberg. doi:10.1007/3-540-34783-6_66.

Parsonson, P. S. (1974). *Small area detection at intersection approaches: A section technical report*. Washington, DC: Institute of Transportation Engineers.

Rakha, H., El-Shawarby, I., & Setti, J. R. (2007). Characterizing driver behavior on signalized intersection approaches at the onset of a yellow-phase trigger. *IEEE, 8*(4), 630–640. doi:10.1109/TITS.2007.908146.

Roess, R. P., Prassas, E. S., & McShane, W. R. (2004). *Traffic engineering* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Traffic Engineering Handbook, 1999. *Institute of Transportation* (5th ed.). Washington, DC: Engineers.

Urbanik, T., & Koonce, P. (2007). *The dilemma with dilemma zones*. <http://www.oregonite.org/2007D6/paper_review/A4_Urbanik_Paper.pdf>.
Webster, F. V. & Elison, P. B. (1965). *Traffic signals for high-speed roads* (RLL Technical Paper 74), Crowthorne, Berkshire, England.
Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*, 338–353.
Zeeger, C. V., & Deen, R. C. (1978). Green-extension systems at high-speed intersections. *ITE Journal*, 19–24.
Zhixia, L., Heng, W., Qingyi, A., & Zhuo, Y. (2010). Empirical analysis of drivers' yellow stopping behaviors associated with dilemma zones. In *Paper 10-3315 presented at the annual meeting of TRB*, Washington, DC.