ARCHIVES OF TRANSPORT *ISSN (print):* *0866-9546*
Volume 44, Issue 4, 2017 ***e-ISSN (online):*** *2300-8830*
DOI: *10.5604/01.3001.0010.6163*

DEVELOPMENT OF A STATISTICALLY-BASED METHODOLOGY FOR ANALYZING AUTOMATIC SAFETY TREATMENTS AT ISOLATED HIGH-SPEED SIGNALIZED INTERSECTIONS

Remigiusz M. Wojtal[1], Laurence R. Rilett[2]

[1] Cracow University of Technology, Faculty of Civil Engineering, Cracow, Poland
[2] University of Nebraska-Lincoln, Lincoln, USA

[1]e-mail: rwojtal@pk.edu.pl
[2]e-mail: lrilett@unl.edu

Abstract: *Crashes at isolated rural intersections, particularly those involving vehicles traveling perpendicularly to each other, are especially dangerous due to the high speeds involved. Consequently, transportation agencies are interested in reducing the occurrence of this crash type. Many engineering treatments exist to improve safety at isolated, high-speed, signalized intersections. Intuitively, it is critical to know which safety treatments are the most effective for a given set of selection criteria at a particular intersection. Without a well-defined decision making methodology, it is difficult to decide which safety countermeasure, or set of countermeasures, is the best option. Additionally, because of the large number of possible intersection configurations, traffic volumes, and vehicle types, it would be impossible to develop a set of guidelines that could be applied to all signalized intersections. Therefore, a methodology was developed in in this paper whereby common countermeasures could be modeled and analyzed prior to being implemented in the field. Due to the dynamic and stochastic nature of the problem, the choice was made to employ microsimulation tools, such as VISSIM, to analyze the studied countermeasures. A calibrated and validated microsimulation model of a signalized intersection was used to model two common safety countermeasures. The methodology was demonstrated on a test site located just outside of Lincoln, Nebraska. The model was calibrated to the distribution of observed speeds collected at the test site. It was concluded that the methodology could be used for the preliminary analysis of safety treatments based on select safety and operational measures of effectiveness.*

Key words: *Traffic control devices, Traffic safety, Traffic signals, Traffic speed, Traffic analysis, Calibration, Validation, Simulation models.*

1. Introduction

In recent years, many transportation agencies have considered implementing safety treatments at high speed, isolated intersections. In one ITE study, 20 potential safety treatments or engineering countermeasures were identified (ITE, 2003) as shown in Table 1. Given the large range of geometric and operating conditions at these types of intersections, it would be impossible to develop a set of guidelines that could be utilized for all situations. Therefore, traffic agencies must examine each intersection with respect to its specific characteristics. To address this issue, the current paper describes a methodology for analyzing the safety and efficiency metrics associated with various safety countermeasures at a particular intersection.

Given the nature of the problem, transportation agencies are not able to conduct some types of field experiments on various safety countermeasures at signalized intersections. For example, an agency could not implement an Advance Warning System (AWS) at a site and then turn it on and off to analyze critical safety measures of effectiveness (e.g., crash rates). In this situation, the only option is to either examine the success of countermeasures at other sites, model the different countermeasures, or some combination of these two approaches. The paper will focus on the second option where a model is used to analyze the options and the recommendations are made based upon the results. Note that no matter what option is chosen it would not preclude the transportation agency from conducting a before/after study to study the effectiveness of the countermeasure.

Because traffic demand at intersections varies over time (e.g., hour of the day, day of the week, day of the year, etc.), analytical macroscopic models may not be appropriate, as they are not designed to handle the dynamic and stochastic nature of signalized intersections. For the current study, the decision was made to employ a traffic microsimulation tool for analyzing the studied countermeasures, since such models are better suited to modeling the complexities of the problem.

Table 1. Engineering safety countermeasures

Objective	**Treatment**
For improving visibility	Placement and number of traffic signal heads Size of sign display Line of sight
For increasing likelihood of stopping	Signal ahead signs Advanced-warning flashers Rumble strips Left-turn signal sign Pavement surface condition
For eliminating the need to stop	Unwarranted signals Intersection design change Flash mode
For improving traffic signal conspicuity	Redundancy LED signal lenses Back-plates Strobe lights
For addressing intentional violations	Signal optimization Signal-cycle length Yellow-change interval All-red clearance interval Dilemma zone protection

The model was calibrated to empirical speed distributions measured at four test sites in Nebraska. From these analyses, one model was selected that was acceptable, from a statistical point of view, for all four test sites (Wojtal, 2012). This paper demonstrates the methodology that was developed for the selection of safety treatments at a single test site in Nebraska. It should be noted that, while the methodology was used at all four test sites and similar results were found. However, space limitations preclude a detailed discussion of the results from all four sites.

It is important to note that current microsimulation models are not adequately robust to sufficiently model all proposed safety countermeasures—particularly those that are related to driver behavior and characteristics (e.g., monitoring of individual drivers) and vehicles (e.g., limits on vehicle size and speed). Therefore, the authors decided to narrow the scope of the current study to countermeasures that are well-suited for modeling using available microsimulation models. Note that, by definition, these countermeasures relate to operational improvements to an intersection. Based on a literature review, two engineering countermeasures were selected: an Advance Detection System (ADS) and an Advance Warning System (AWS). The comprehensive methodology developed in this paper was generic in nature, and it is hypothesized that other countermeasures may also be analyzed when more sophisticated microsimulation models are developed.

This paper first provides an overview of the statistical methodology used to conduct operational and safety analyses on the select safety treatments. This methodology consists of three phases. The first phase identifies the safety treatments examined in this paper. The second phase selects an appropriate traffic microsimulation model and includes a brief description of the extensions to the microsimulation logic required to facilitate the modeling of the safety treatments chosen in phase one. In the third phase a description of the safety measures of effectiveness that were chosen is also provided. Lastly, the methodology is illustrated on a rural, isolated, high-speed intersection test site to demonstrate how the methodology could be applied. The test site was located at the intersection of US Highway 77 and Pioneers Boulevard in Lincoln, Nebraska and its outline is shown on Fig. 1 (Wojtal, 2012). This test site will be used to motivate the discussion throughout the paper.

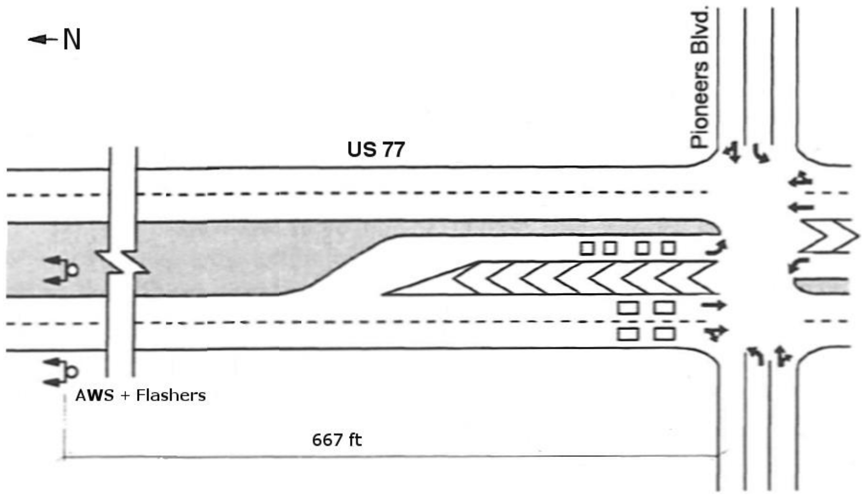

Fig. 1. Test site configuration – SB Approach at US Highway 77 & Pioneers Blvd

2. Methodology

A schematic of the proposed methodology is presented in Fig. 2. The preliminary step was to identify the potential safety treatments. It is important to note that the set of potential safety treatments should include a base, or, "do-nothing" case.

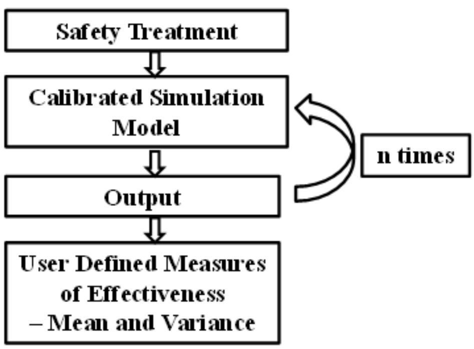

Fig. 2. Generic methodology for the analysis of traffic safety

A VISSIM simulation model, which was calibrated using empirical data obtained from the test site, was used to model the safety treatment (Wojtal, 2012). For each of the safety treatments considered the model was run *n* times (i.e., 5-10), and safety and operational metrics (e.g., measures of effectiveness) were output. These measures of effectiveness were selected by the modeler, and could have included any number of efficiency and safety-related metrics, including queue length, delay, number of vehicles in the dilemma zone, etc. Measures of central tendency (e.g., mean) and measures of dispersion (e.g., variance) were estimated from the results, and were used to make statistical inferences among the treatments. The number of runs, *n*, was a function of the accuracy desired by the modeler. Each of these steps are discussed in detail in the following sections.

2.1. Safety treatments

As was shown in Table 1, there are more than 20 safety treatments that could be considered for adoption at high-speed signalized intersections. In this paper, three operations-based safety treatments were selected for analysis:

- No safety treatment (NST);
- Advance detection system (ADS);
- Advance warning system (AWS).

The first case, NST, is a "do-nothing" scenario, and was used for comparison. The second treatment, ADS, has been used to provide dilemma zone protection for high-speed approaches at isolated signalized intersections. The system detects approaching vehicles and, when appropriate, extends the green interval to enable drivers to safely traverse an intersection without needing to decide whether or not to stop. In essence, the ADS attempts to reduce the number of times that a driver has to choose to stop or go. In summary, the goal is to reduce the number of times a dilemma zone situation may occur. Green time extension is typically of the order of three to seven seconds. Note that drivers are unaware of when an ADS is active, and when green time has been extended. Based on this fact, driver behavior can be modeled in a manner similar to that of the NST scenario. A complete description of ADS may be found elsewhere (Appiah et al., 2011; McCoy and Pesti, 2002). AWS was the third countermeasure to be examined. AWS provides information to drivers, via flashing signal heads and warning signs, regarding whether or not they should be prepared to stop as they approach a signalized intersection. In this paper, it was assumed that the AWS scenario also included an operational ADS system. The flashing signal head(s) were activated at a predetermined time (i.e., between five and eight seconds, depending on the location of the flashers) prior to the termination of the green interval. In this situation, the microsimulation tool needed to be able to model driver reactions to the flashing signal. A complete description of AWS can be found in extant literature (Appiah et al., 2011; McCoy and Pesti, 2002; Park et al., 2015; Wojtal, 2012).

2.2. Calibration of the microsimulation model

The VISSIM microsimulation model was used in this paper and was calibrated to and validated against data from a test site located at the intersection of US Highway 77 and Pioneers Boulevard. This intersection was located in a rural area approximately five miles south of Lincoln, Nebraska and was outfitted with an AWS developed by the Nebraska Department of Roads (NDOR) (McCoy and Pesti, 2002; Wojtal, 2012). Empirical data gathered from this location consisted of traffic volume, traffic composition, traffic speed (used in the calibration procedure), and waiting times on the minor approaches (used in the validation procedure). A brief synopsis of the procedure is provided below, and a complete description of the approach can be found elsewhere (Wojtal, 2012).

First, a VISSIM microsimulation model of the signalized intersection was developed, including road geometry and traffic signal timing. Subsequently, the model was adjusted so that the select safety treatments could be tested. To effectively model the AWS, a special algorithm needed to be incorporated into the VISSIM model used in this study (Wojtal, 2012). Specifically, the warning sign had to be added to the simulation and the drivers behavior, once the sign was activated, had to be modeled accurately based on empirical data.

Once the microsimulation model of the signalized intersection was complete, the model was calibrated to empirical data collected from the test site (Wojtal, 2012). The goal of the calibration procedure was to identify the "best" set of driving behavior parameters, where the best set was that which provided statistically acceptable results and had the lowest difference, as measured by Mean Absolute Percentage Error (MAPE). The MAPE was calculated using parameters of observed and simulated speed distribution (mean, median, mode, standard deviation and kurtosis) (Wojtal, 2012). A genetic algorithm was used to conduct the calibration. Because the objective of the methodology was to indicate the most effective safety treatment, it was decided to calibrate the model to the distribution of observed speeds, rather than to the mean of the speed distribution. The speed distribution was selected as a criterion because it was a parameter that effectively characterized the nature of traffic, and was a common measure of safety. The calibration was performed on nineteen VISSIM parameters, which included car-following, lane changing, desired speed distribution, and signal control parameters. The initial set of VISSIM parameters used in the calibration was identified and selected based on engineering judgment and a review of salient literature. The model was successfully calibrated, in that the various observed and simulated metrics were not statistically different at the 5% significance level (Wojtal, 2012).

Once the signalized intersection model was successfully calibrated, a validation procedure was performed to determine whether the model performed adequately. In this case, empirical data pertaining to waiting time on the minor approaches were compared to the simulation model output. Note that these empirical data were not used during the calibration procedure. It was determined that the calibrated model behaved appropriately and could be used for further analysis at the test site (Wojtal, 2012). The geographic transferability of the model within Nebraska confirmed similar work (Essa and Sayed, 2015).

2.3. Measures of effectiveness

Most microsimulation models do not provide output on numbers of crashes or crash rates due to the nature of their internal logic. For example, VISSIM does not allow vehicles to collide, so the number of crashes for a given simulation would be impossible to obtain. In this situation, surrogate safety measures, that attempt to gauge the safety of a facility, are utilized (Gettman and Head, 2003a; Gettman and Head, 2003b). Basic surrogate safety measures proposed in the literature for intersections include minimum time to collision, delay, maximum speed of two vehicles during conflict, maximum difference in the speed of vehicles during conflict, travel time, approach speed, percent stops, queue length, stop-bar encroachments, red light violations, percentage of left turns, spot speed, speed distribution (Gettman and Head, 2003a; Gettman and Head, 2003b; Liu et al., 2006), and the number of vehicles in the dilemma zone (Huang and Pant, 1994; Machiani and Abbas, 2015; Perkins and Bowman, 1986). In general, measures of traffic conditions, such as delays or queues, are not related directly to crash rates, but have been found to be correlated with safety rules of thumb such as, "Higher delays or longer queues indicate a higher probability of crashes" (Davis et al., 2008; Gettman and Head, 2003a; Gettman and Head, 2003b).

In recent years the use of microsimulation models for safety analyses has become more widespread as the use of the surrogate measures, discussed above, has increased (Caliendo and Guida, 2012; Cunto and Saccomanno, 2008; Fazio and Rouphail, 1990; Kosonen and Ree, 2001; Ozbay et al., 2008; Sayed et al., 1994; Shahdah et al., 2015). In recent years the Surrogate Safety Assessment Model (SSAM) has been developed which measures relative safety using surrogate safety measures from microsimulation output (Gettman et al., 2008; Pu and Joshi, 2008). This software, in conjunction with a microsimulation model, has been used in a number of safety studies (Gettman et al., 2008; So et al., 2015). Note that most of the SSAM models are focused on obtaining surrogate safety measures based on the individual vehicle trajectories, and their associated conflicts, output form the model. Because the model developers never developed their models to have an accurate conflict resolution process some authors have questioned the validity of this approach (So et al., 2015). In addition, because the safety counter-measures in this paper seek to eliminate, or at least reduce, the most serious conflicts it was decided to 1) focus on an approach that explicitly models driver behavior with respect to the traffic signs and 2) uses MOEs that are directly related to the countermeasures that are analyzed.

The measures of effectiveness (MOE) selected for this paper could be divided into two groups. The first type were operational in nature and have previously been related to safety (Gettman and Head, 2003a; Gettman and Head, 2003b). These metrics included average total delay per vehicle (on all approaches), average approach speed along the southbound major approach, and maximum queue length (on all approaches). The second type were MOEs specifically related to the safety treatments that were studied (Machiani and Abbas, 2015). In particular, the ADS and AWS treatments were specifically designed to reduce the number of vehicles in the dilemma zone. In addition, the AWS treatment was designed so that vehicles would reduce their speeds upon approaching an intersection. In the current paper, the average number of vehicles in the dilemma zone at the southbound approach was used as the safety MOE. It is important to note that this MOE is not a standard output for microsimulation models. Consequently, this MOE had to be calculated using detailed information output from the model.

It is important to note that the change in vehicle speeds once the flashing warning light became active in the AWS scenario was not selected as an MOE. This was because the microsimulation model was calibrated so that the simulated drivers reacted in a similar way to actual drivers measured in the field. In other words, while the authors could "measure" the change in simulated vehicle when the advance warning sign was active, this behavior was not an emergent property of the model; rather, it was "hard-wired" by the authors based on empirical data. In summary, using changes in vehicle speed in reaction the warning sign would be a false MOE because the authors directly input this behavior into the model. Blindly using the resulting microsimulation output directly, or in a post-processor such as SSAM, could result in false conclusions relating to the efficacy of a countermeasure.

The first MOE chosen was average total delay per vehicle on all approaches. This value was computed for every vehicle traversing a distance from 500 to 180 ft upstream of the stop-bar on all four approaches. It was calculated by subtracting the theoretical travel time, which was the time that it would take for a vehicle to traverse the distance if there were no other vehicles and no traffic controls, from actual travel time (PTV, 2011). This MOE was a standard metric included in the VISSIM output.

The next MOE to be utilized was average approach speed along the southbound major approach. This was the average instantaneous speed of every vehicle at a designated cross-section located upstream of the stop-bar at the SB approach. For the study site, nine locations along the SB approach were selected for analysis. The first location was located 200 ft from the stop-bar, and each subsequent location was 100 ft further back. The maximum distance was 1,000 ft. This MOE was a standard metric that included in VISSIM output.

The maximum queue length was the third MOE to be considered. This was the average maximum queue length in feet, counted from the queue counter (located at the stop-bar) to the final vehicle that in the queue condition over the course of the simulation (PTV, 2011). This MOE was a standard metric included in VISSIM output.

The final MOE was the average number of vehicles in the dilemma zone on the southbound approach. For the purposes of this research, the definition of the dilemma zone was adapted from the literature review (Liu et al., 2006; McCoy and Pesti, 2002; Urbanik and Koonce, 2007). The "Type I Dilemma Zone" definition was chosen because it is commonly used in practice. The Type I Dilemma Zone was defined as a range in which a vehicle approaching the intersection during the yellow phase could neither safely clear the intersection nor stop comfortably at the stop-bar (Liu et al., 2006). For the test-site, the dilemma zone ranged from 500 ft upstream of the stop-bar to 250 ft upstream of the stop-bar. It was felt that drivers in this location would have a choice on whether to stop or proceed and that removing drivers from this area by 1) extending the green time, 2) warning them ahead of time so they would start decelerating, or 3) both would lead to greater safety by reducing the number of potential conflicts at the intersection.

Note that VISSIM does not give the user the option of outputting the number of vehicles in a dilemma zone directly. Therefore, the authors output detailed disaggregate data on signal status and every vehicle's location at 0.1 second intervals. A MatLab program was coded to read the VISSIM output files and to extract the number of vehicles in the dilemma zone for each traffic signal transition (i.e., green to yellow). The average and standard deviation of this parameter over the one hour simulation time was used as the MOE. It is also important to note that any definition for the dilemma zone could be examined using the methodology presented in this paper, as long as the selected microsimulation model provides output for individual vehicle time and location data.

2.4. Statistical tests

One of the biggest advantages—and challenges—associated with microscopic simulation models is the fact that the simulation output differs for different random number seeds. To minimize the effect of obtaining an unrepresentative result for a single run, each treatment was run 10 times, each time with a different random seed number. The decision was made to use 10 runs, as the literature review and preliminary analysis indicated that this provided an appropriate trade-off between computation time and result accuracy (Park and Schneeberger, 2003; Spiegelman et al., 2010; Wiegand and Yang, 2011). It is also important to note that the simulation runs for each treatment used the same 10 randomly generated seed numbers. This allowed the paired t-test, which is a stronger statistical test than the regular t-test, to be used to measure statistically significant differences between the mean results of the safety treatments (Spiegelman et al., 2010). Each simulation run lasted one hour. This value was chosen since it allowed approximately 70 cycles to be completed and further enabled the model to operate in a steady state of conditions for a majority of the simulation run time. Once the MOEs for each scenario were obtained, it was possible to statistically compare the mean results between treatments. Two approaches were adopted, as shown in Fig. 3. In the first approach, 95% confidence intervals (CI) were utilized. If the CI of a given metric between two MOE's overlapped the mean values for compared treatments, the null/alternative hypothesis that the means did not differ at the 5% significance level was rejected. The confidence intervals were calculated for all MOEs for all three treatments. The mean values and associated CIs were placed in graphs, allowing statistically significant differences to be visually identified.

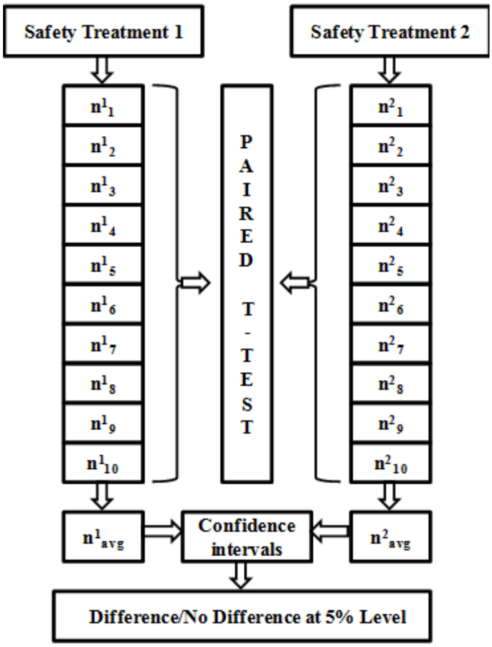

Fig. 3. Statistical tests in the comparison of safety treatments

Because the same 10 random number seeds were used to test each treatment, a paired t-test was used to test for statistically significant differences between treatments. When comparing two different treatments, the paired t-test indicates whether there exists a statistically significant difference at the 5% level between the means of the measures of effectiveness. This statistical test, which is an example of the repeated measures design, is more powerful than the regular t-test or CI approach; as such, it should be expected to identify a greater (or at least equivalent) number of statistically significant differences between the two treatments. Because each scenario is used as its control, individual differences are not present, and can be ruled out of the random error term, minimizing its effect. Therefore, the test results in a smaller error and, consequently, a larger t-value.
Both the regular and paired t-tests could be used to identify statistically significant differences between the analyzed treatments. As discussed above, the paired t-test is a more powerful test. The results of both the regular t-test and paired t-test are presented in the current paper in order to demonstrate both techniques. In practice, a user would pick one approach and the authors recommend the paired t-test.

3. Analysis

All three scenarios, including the two safety treatments and the NST, were analyzed using the methodology described above. Input volumes and turning movements were based on the empirical data collected from the test site. The approach volumes on the major and minor approaches were 600 veh/h and 150 veh/h, respectively. For the major directions (i.e., northbound and southbound), the share of the through movement was 90% while the left and right turning movements were 10% each. For the minor directions, (eastbound and westbound) the share of through movement was 20% while the left and right turning movements were 40% each. The heavy vehicle percentage for the entire intersection was set at 10% for all approaches and movements; this figure was based on empirical measurements at the test site.
The selected metrics for the different scenarios are briefly provided in the following sections in order to demonstrate the practical use of the developed methodology. A complete description, and all results of the analysis, may be found in the literature (Wojtal, 2012).

3.1. Average total delay per vehicle

Fig. 4 displays the average total delay per vehicle for the southbound approach for the three scenarios: NST, AWS and ADS (i.e., as previously described) It can be seen that the AWS scenario displayed the lowest average total delay, which was approximately 16% less than the NST delay. In contrast, the ADS scenario resulted in an approximate 5% delay reduction in comparison to the NST. The AWS and ADS treatments resulted in delays that significantly differed from those of the NST case at the 5% significance level, evidenced by the fact that their confidence intervals did not overlap. Similar conclusions applied for the major NB approach.
While the AWS logic was designed to increase safety at the intersection by avoiding side crashes, the use of the green extension on the major direction

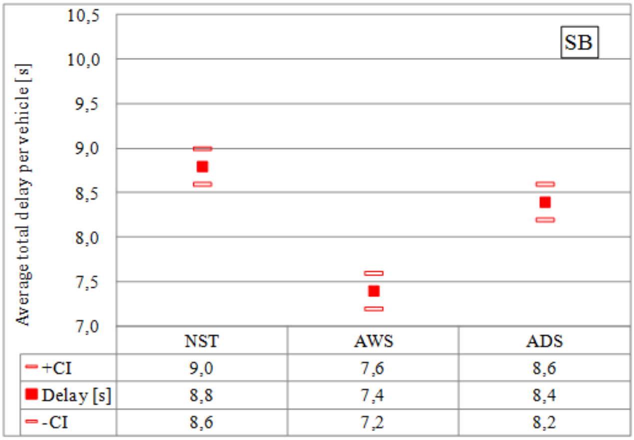

Fig. 4. Average total delay at southbound approach at the test site

can lead to a lower level of service on the minor approaches. As would be expected, delays at the minor approach did increase in the AWS scenario, by approximately 30% (Wojtal, 2012). The ADS treatment resulted in a decrease in delay of 4%. However, these delays did not statistically differ from the NST delay, as evidenced by the confidence intervals. Similar conclusions applied to the westbound minor approach.

The results of the paired two-sample t-test between the NST and the two safety treatments supported the conclusions drawn from the confidence intervals. There was a statistically significant difference at the 5% significance level in delay between the NST and AWS treatments for all approaches. The ADS treatment decreased delays in comparison to the NST case; the difference was found to be statistically significant for both major approaches. However, there was no statistically significant difference in average total delay for both minor approaches (Wojtal, 2012).

For this metric, the conclusions of the regular t-test (e.g., confidence intervals) and paired t-test were the same.

3.2. Average approach speed

Fig. 5 illustrates the difference in speed between the NST case (represented in the graph by the value “0.0”) and all other safety treatments as a function of distance from the stop-bar. A positive value implied that the approach speed under the safety treatment was higher than that of the NST scenario, while a negative value implied a lower approach speed for the safety treatment. Confidence intervals for the NST scenario are shown in Fig. 5 as dotted lines.

In the AWS treatment, the vehicle approach speed decreased in comparison to the NST treatment when vehicles were closer to the intersection (300-400 ft). Furthermore, when the distance to the stop-bar was greater (500-1,000 ft), the approach speed of the AWS treatment was higher than that of the NST treatment.

The ADS treatment resulted in a lesser speed increase than the NST for the cross-sections located close to the intersection (200-500 ft). The confidence intervals indicated a statistically significant difference in approach speed at the 5% significance level between the NST scenario and the AWS treatment for the cross-sections located between 500 and 800 ft of the stop-bar. For the ADS treatment, change in speed was not statistically significant.

The paired two-sample t-test was also performed to check for statistically significant differences at the 5% significance level between the approach speed from the NST scenario and all other safety treatments as shown in Table 2. The AWS treatment changed the approach speed significantly at a 5% level of significance for the cross-sections located between 400 and 1,000 ft of the stop-bar, in comparison to the NST scenario. The ADS scenario increased the approach speed for the cross-sections located between 200-300 ft from the intersection. This change was statistically significant (Wojtal, 2012).

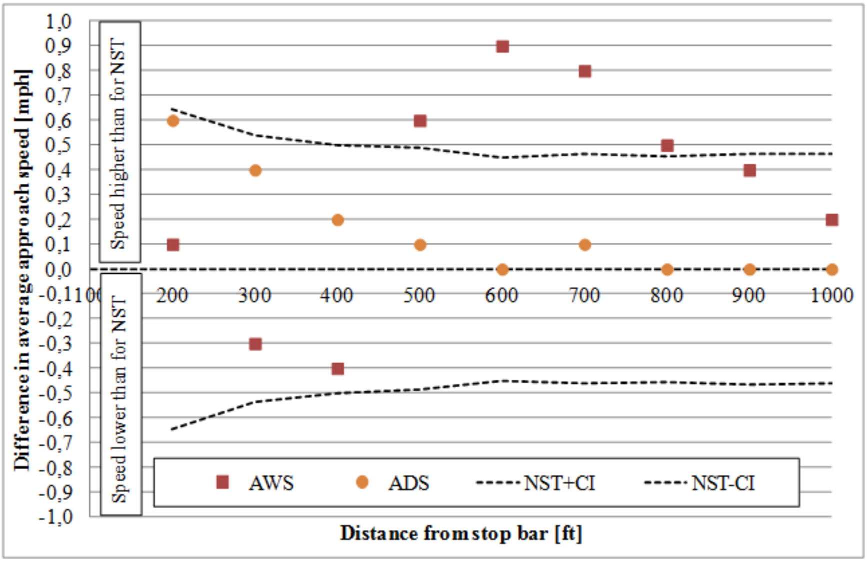

Fig. 5. Difference in speed between NST and two safety treatments

Table 2. Results of paired t-test of average approach speed at the SB approach

Distance [ft]	Treatment 1	Treatment 2	Difference in means [mph]	Result
200	NST	AD	-0.6	Difference
300	NST	AD	-0.4	Difference
400	NST	AWS	+0.3	Difference
500	NST	AWS	+0.2	Difference
600	NST	AWS	-0.9	Difference
700	NST	AWS	-0.7	Difference
800	NST	AWS	-0.5	Difference
900	NST	AWS	-0.4	Difference
1000	NST	AWS	-0.2	Difference

The conclusions derived from the regular t-test (i.e., confidence intervals) and paired t-test differed for this metric. In particular, the paired t-test revealed statistically significant differences at 400, 900, and 1,000 ft for the AWS treatment, and at 200 and 300 ft for the ADS treatment, compared to the regular t-test.

3.3. Maximum queue length

Fig. 6 shows the maximum queue length for the major southbound approach for each of the three treatments. It can be seen that the AWS treatment experienced a higher maximum queue length (approximately 13%) compared to the NST scenario. The ADS treatment resulted in a 4% approximate decrease in comparison to the NST scenario. The confidence intervals indicated that there was no statistically significant difference in maximum queue length at the 5% significance level between the NST scenario and the treatments. Similar results were obtained for the NB approach (Wojtal, 2012). On the minor eastbound approach, the AWS and ADS treatments exhibited a reduction of 22% and 14%, respectively, in maximum queue length, as compared to the NST scenario. The confidence intervals indicated that these results were statistically significant at the 5% significance level. Very similar results were obtained for the WB approach (Wojtal, 2012).

The results of the paired two-sample t-test between the NST scenario and the safety treatments revealed that for the major NB and SB approaches, there was a statistically significant difference at the 5% level of significance in maximum queue length between the NST and AWS scenarios There was not a statistically significant difference observed between the NST and ADS treatments. It was concluded that the AWS treatment increased queue lengths, from a statistical point of view, along the major approaches. The minor approaches indicated rather divergent results; therefore, no clear conclusions regarding the treatment and queue length could be drawn (Wojtal, 2012).

For the maximum queue length metric, the conclusions obtained from the regular t-test (e.g., confidence intervals) and paired t-test differed. In particular, the paired t-test revealed statistically significant differences between the NST and AWS scenarios for the SB and NB approaches, in comparison to the regular t-test. Similar conclusions were drawn for the minor EB, but not the WB, approach.

3.4. Number of vehicles in dilemma zone

Fig. 7 displays the numbers of vehicles in the dilemma zone as a function of the treatment for the SB major approach. The AWS and ADS treatments are specifically designed to reduce and hopefully eliminate vehicles becoming caught in the dilemma zone. As was expected, the use of the AWS and ADS treatments decreased the number of vehicles in the dilemma zone (by 39% and 27%, respectively) in comparison to the NST scenario. Both results were statistically significant at the 5% level, as shown in Fig. 7

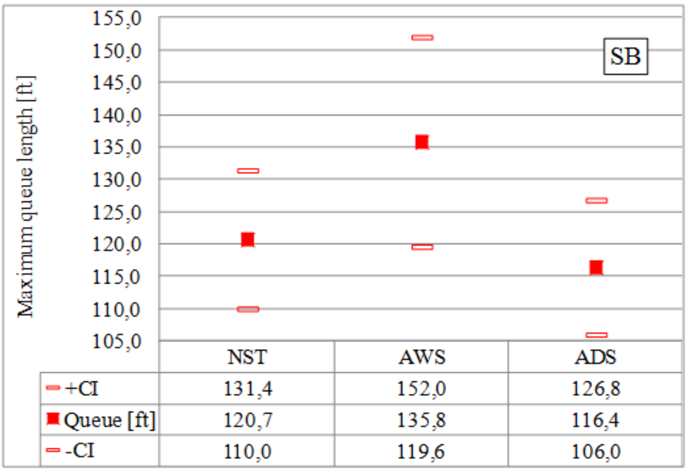

Fig. 6. Maximum queue length at southbound approach along the test site

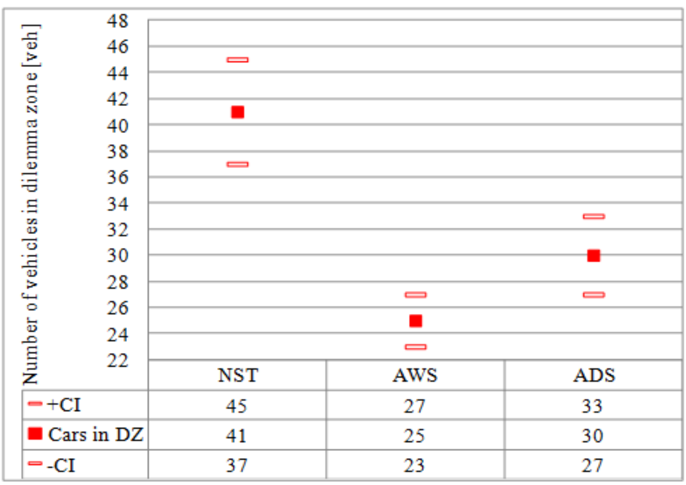

Fig. 7. Number of vehicles in the dilemma zone at test site southbound approach

The paired two-sample t-test was also performed to check for statistically significant differences between the safety treatments; it was determined that there existed a statistically significant difference at the 5% significance level between the NST scenario and the AWS and ADS treatments. Additionally, the AWS was more effective in reducing the number of vehicles in the dilemma zone than was the ADS as evidenced by the statistically significant difference between these treatments at the 5% significance level. This result was to be expected, since the AWS was designed not only to extend the green time, but also to encourage drivers to reduce their speeds when the signal is about to change from green to yellow. As discussed earlier, the drivers' behavior with respect to the sign status was based on empirical data (Wojtal, 2012). If the models were not adjusted for this behavior the AWS and ADS treatments would have given similar results which illustrate the danger of using uncalibrated microsimulation models directly for safety analyses.

In terms of the number of vehicles in the dilemma zone, the conclusions drawn from the regular t-test (i.e., confidence intervals) and paired t-test differed. In particular, the paired t-test revealed statistically significant differences between the AWS and ADS treatments that were not identified by the regular t-test.

As discussed previously it would be impossible to confirm these types of analyses using a designed experiment. For example, no transportation agency would allow an experimental design in which a system would be installed and then turned on and off to analyze the resulting safety impacts in terms of crash rates. However, it has been demonstrated that these systems do reduce crashes in comparison to unequipped intersections; this reduction has been attributed to a reduced number of vehicles caught in the dilemma zone (Appiah et al., 2011). In addition, empirical measurements of similar intersections with and without these systems present have demonstrated that the treatments reduce driver speeds at the onset of flashing warning lights. The important point to consider is that microsimulation models, when correctly calibrated to key metrics—in this case, the distribution of vehicle speed and driver behavior at the onset of warning lights—can be utilized to estimate the effectiveness of these types of safety countermeasures.

4. Summary

This paper demonstrated a general methodology for the analysis of safety treatments at signalized intersections. The proposed approach was used to analyze two operations-based safety treatments at a high-speed rural intersection in Lincoln, Nebraska. The two treatments examined were an Advance Detection System and an Advance Warning System; these treatments were compared to the "do-nothing," or, no safety treatment scenario. The conclusions can be broken down into two categories, i.e., those related to the proposed methodology and those strictly related to the analysis of the test site.

Conclusions of the proposed methodology

- The methodology described in this paper was an effective tool for the analysis of engineering safety treatments because it takes into account the stochastic nature of traffic and allows for the testing of various measures of effectiveness (e.g., number of vehicles in the dilemma zone) that would be difficult, if not impossible, to conduct using standard analytical models. Moreover, the model can be utilized for sensitivity analysis of safety metrics as a function of key traffic parameters such as volume or heavy vehicle percentage.
- The VISSIM model can be used to model safety treatments at signalized intersections through adjusting driving behavior parameters (e.g., speed distribution). The calibration and validation methodology could be utilized in any microsimulation model. The only limitation is the feasibility of the microsimulation in terms of accurately modeling the proposed safety countermeasure. It was demonstrated that, at present, microsimulation models can only be used for a small subset of the total number of potential safety treatments at signalized intersections.
- As stated previously, the microsimulation model appears to be a very useful and accurate tool for safety analysis, but it must be emphasized that only properly calibrated and validated models can provide accurate results. Additionally, it is critical to select the calibration procedure to include parameters the affect driver behavior, such as the distribution of approach speeds.
- A genetic algorithm with non-parametric tests is a very effective tool for the calibration of traffic and the stochastic simulation models of signalized intersections. While previous researchers have

calibrated microsimulation models to measures of central tendency (e.g., mean), the current study utilized approach speed distribution, since the distribution of vehicle speeds is directly related to the number of vehicles in the dilemma zone—a key measure of effectiveness of the examined safety treatments. The model was calibrated in an appropriate timeframe, and all results were statistically accurate.

- The number of vehicles in the dilemma zone is a highly effective measure of safety because it directly gauges what the modeled safety countermeasures are attempting to improve. However, this metric is typically not output as part of microsimulation models. In this paper, an automatic technique was used to calculate this measure.
- While the Type I Dilemma Zone definition was utilized in this paper, it is important to note that any definition could be utilized so long as the chosen microsimulation model produces individual vehicle location and time data.
- It was revealed that the paired t-test was a more powerful tool than the regular t-test for identifying differences among the treatments. For example, using the same number of simulation runs, the former test revealed more statistically significant differences among the treatments. The authors recommend that the paired t-test be used.

<u>Rural, high speed, isolated, signalized intersection test site</u>

- Using the proposed methodology, it was concluded that both safety treatments effectively improved safety, based on the four measures of effectiveness utilized in the current study: average total delay, average approach speed along the major approach, maximum queue length, and the number of vehicles in the dilemma zone. It was shown that both treatments improved traffic safety at the test intersection by reducing the number of vehicles in the dilemma zone by 39% and 27%, respectively. At the same time, the average total delay at the major approaches was reduced, with a simultaneous increase in delays on the minor approaches. Therefore, for the base case, the recommended safety treatment, without taking cost into account, could be an AWS, which had the largest statistically significant decrease in the number of vehicles in the dilemma zone. The AWS treatment also resulted in a statistically significant decrease in total delay along the major approach.
- Based on the results of the analysis, it is clear that some trade-off between safety and traffic operations is necessary. Intuitively, it is very unlikely that a safety treatment could simultaneously enhance both safety and traffic operations for all approaches. It should also be noted that the proposed methodology was not developed to directly select the best treatment. Rather it was developed to provide information regarding safety and efficiency, which could be used by the appropriate decision-makers to select the best safety treatment or treatments at a signalized intersection.

It was demonstrated that the proposed methodology could be utilized to analyze a subset of safety countermeasures at signalized intersections. Future work will depend on the creation of new, and/or the adjustment of existing, micro-simulation models that can better model safety-related impacts.

This paper focused on differences in measures of central tendency among the MOE's. However, users are often interested in the distribution as well as the average, and this should be examined. In addition, this paper examined only four safety-related measures of effectiveness. However, there is much room for improvement in the selection of MOE's that better reflect changes in safety. For example, it would be ideal if models could accurately predict the number of red light runners for a given safety treatment.

References

[1] APPIAH, J., RILETT, L. R., & WU, Z., 2011. Evaluation of NDOR's Actuated Advance Warning System. *Report TSPT1-03*, Nebraska Transportation Center, University of Nebraska-Lincoln, NE.

[2] CALIENDO, C., & GUIDA, M., 2012. Microsimulation Approach for Predicting Crashes at Unsignalized Intersections Using Traffic Conflicts. *Journal of Transportation Engineering*, 10.1061/(ASCE)TE.1943-5436.0000473, pp. 1453-1467.

[3] CUNTO, F., & SACCOMANNO, F. F., 2008. Calibration and validation of simulated vehicle safety performance at signalized intersections. *Accident Analysis & Prevention*, 40(3), pp. 1171-1179.

[4] DAVIS, G. A., HOURDOS, J., & XIONG, H., 2008. Outline of causal theory of traffic conflicts and collisions. In Transportation Research Board 87th Annual Meeting, No. 08-2431.

[5] ESSA, M., & SAYED, T., 2015. Transferability of calibrated microsimulation model parameters for safety assessment using simulated conflicts. *Accident Analysis & Prevention*, Volume 84, pp. 41-53.

[6] FAZIO, J., & ROUPHAIL, N. M., 1990. Conflict simulation in INTRAS: Application to weaving area capacity analysis. *Transportation Research Record*, Vol. 1287, pp. 96-107.

[7] GETTMAN, D, & HEAD, L., 2003. Surrogate Safety Measures from Traffic Simulation Models. *Final Report, FHWA Report FHWA-RD-03-050*, p 126.

[8] GETTMAN, D., & HEAD, L., 2003. Surrogate safety measures from traffic simulation models. *Transportation Research Record*, Vol. 1840, pp. 104–115.

[9] GETTMAN, D., PU, L., SAYED, T., & SHELBY, S. G., 2008. Surrogate safety assessment model and validation: *Final report. Report No. FHWA-HRT-08-051*.

[10] HUANG, X. P., & PANT, P. D., 1994. Simulation Neural-Network Model for Evaluating Dilemma Zone Problems at High-Speed Signalized Intersections. *Transportation Research Record*, Vol. 1456, pp. 34-42.

[11] Institute of Traffic Engineers, 2003. Making Intersections Safer: A Toolbox of Engineering Countermeasures to Reduce Red-Light Running. IR-115 500/STP/CAA/PMR/0503, Washington, DC.

[12] KOSONEN, I., & REE, S., 2001. The potential of microscopic simulation in traffic safety and conflict studies. Proceedings of the Conference on Road Safety on Three Continents, Pretoria, South Africa, VTI Konferens 15A, No. 15A, pp. 786-794.

[13] LIU, Y., CHANG, G., TAO, R., HICKS, T., & TABACEK, E., 2006. Empirical Observations of Dynamic Dilemma Zones at Signalized Intersections. *Transportation Research Record*, Vol. 2035, pp. 122–123.

[14] MACHIANI, S. G., & ABBAS, M., 2016. Safety Surrogate Histograms (SSH): A Novel Real-Time Safety Assessment of Dilemma Zone Related Conflicts at Signalized Intersections. *Accident Analysis & Prevention*, Vol. 96, pp. 361-370.

[15] McCOY, P. T., & PESTI, G., 2002. Advance Detection on High-Speed Signalized Intersection Approaches. *Report No. SPR-PL-1(035) P525*, Department of Civil Engineering, University of Nebraska-Lincoln, NE.

[16] OZBAY, K., YANG, H., BARTIN, B., & MUDIGONDA, S., 2008. Derivation and validation of new simulation-based surrogate safety measure. *Transportation Research Record*, Vol. 2083, pp. 105–113.

[17] PARK, B., & SCHNEEBERGER, J. D., 2003. Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System. *Transportation Research Record*, Vol. 1856, pp. 185-192.

[18] PARK, S. Y., XU, L., & CHANG, G., 2015. Design and Evaluation of an Advanced Dilemma Zone Protection System: Advanced Warning Sign and All-Red Extension. Compendium of Papers from the 94th Annual Meeting of the Transportation Research Board, Washington, D.C.

[19] PERKINS, D., & BOWMAN, B., 1986. Effectiveness Evaluation by Using Non-accident Measures of Effectiveness. *Transportation Research Record*, Vol. 905, pp. 138-142.

[20] PTV Plannung Transport Verkehr AG, 2011. VISSIM 5.30-05 User Manual. Innovative Transportation Concepts.

[21] PU, L., & JOSHI, R., 2008. Surrogate Safety Assessment Model (SSAM): Software User Manual. *Publication No. FHWA-HRT-08-050*.

[22] SAYED, T., BROWN, G., & NAVIN, F., 1994. Simulation of traffic conflicts at unsignalized intersections with TSC-Sim. *Accident Analysis & Prevention*, Vol. 26(5), pp. 593-607.

[23] SHAHDAH, U., SACCOMANNO, F., & PERSAUD, B., 2015. Application of traffic microsimulation for evaluating safety performance of urban signalized intersections. *Transportation Research Part C: Emerging Technologies 60*, pp. 96-104.

[24] SO, J., PARK, B. B. WOLFE, S. M., & DEDES, G., 2015. Development and Validation of a Vehicle Dynamics Integrated Traffic

Simulation Environment Assessing Surrogate Safety. *ASCE Journal of Computing in Civil Engineering 29*, 04014080. Online publication date: 2015.
[25] SPIEGELMAN, C., PARK, E.S., & RILETT, L.R., 2010. *Transportation Statistics and Microsimulation*. Chapman and Hall/CRC; 1 edition.
[26] URBANIK, T., & KOONCE, P., 2007. The Dilemma with Dilemma Zones. Proc., ITE District 6.
[27] WIEGAND, J., & YANG, D., 2011. Traffic Simulation Runs: How Many Needed?. *Public Roads Volume: 74*, Issue Number: 4, 2011-1., pp 30-35.
[28] WOJTAL, R., 2012. *Development of a Methodology for Analyzing Safety Treatments at Isolated Signalized Intersections*. Ph.D. Dissertation, University of Nebraska-Lincoln, NE.