TECHNICAL TRANSACTIONS 12/2017
CZASOPISMO TECHNICZNE 12/2017
CIVIL ENGINEERING
DOI: 10.4467/2353737XCT.17.214.7757

Remigiusz Wojtal (rwojtal@pk.edu.pl)
Institute of Road & Railroad Engineering and Transport, Faculty of Civil Engineering, Cracow University of Technology

Laurence R. Rilett
Keith W. Klaasmeyer Chair in Engineering and Distinguished Professor, Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE, USA

Calibration of microsimulation models of advance detection and warning systems at signalized intersections

Kalibracja modeli mikrosymulacyjnych systemów wczesnej detekcji i ostrzegania na skrzyżowaniach z sygnalizacją świetlną

Abstract

This paper presents a calibration and validation procedure for microsimulation models, which used metrics (mean, variance, and mean absolute percentage error) with statistical tests (t-test, Shapiro-Wilk, Kolmogorov-Smirnov and Wilcoxon), to compare empirical and simulated data. A genetic algorithm was used to identify calibration parameters set. The paper justifies the approach using VISSIM microsimulations to analyze two safety countermeasures: Advance Detection System and Advance Warning System, which can be applied at signalized intersection. The end result was a calibrated and validated model, which could be used to compare different safety countermeasures at rural signalized intersections in the state of Nebraska (USA). The proposed approach could be utilized in similar studies.

Keywords: signalized intersection, safety countermeasures, calibration, validation, simulation model

Streszczenie

W artykule przedstawiono procedurę kalibracji i walidacji modeli mikrosymulacji z użyciem mierników (średnia, wariancja i średni bezwzględny błąd procentowy), oraz testów statystycznych (studenta, Shapiro-Wilka, Kolmogorow-Smirnova oraz Wilcoxona) dla porównania danych empirycznych i z symulacji. Algorytm genetyczny wykorzystano do doboru parametrów kalibracyjnych. Metodologia została sprawdzona w programie VISSIM do analizy dwóch systemów poprawy BRD: Systemu Wczesnej Detekcji i Wczesnego Ostrzegania, które stosuje się na skrzyżowaniach z sygnalizacją świetlną. W wyniku uzyskano skalibrowany i zweryfikowany model, który można użyć do porównania środków poprawy BRD na zamiejskich skrzyżowaniach z sygnalizacją w stanie Nebraska (USA). To podejście można wykorzystać w podobnych analizach.

Słowa kluczowe: skrzyżowanie z sygnalizacją, środki poprawy brd, kalibracja, walidacja, model symulacyjny

1. Introduction

The main advantage of using microsimulation tools to examine potential traffic safety improvements is their ability to handle the dynamic and stochastic nature of traffic. However, it is critical that the user calibrates and validates the models for the particular site and safety countermeasure being studied. Traffic micro-simulation models are very useful tools for simulating the effects of proposed network improvements, and have the potential to model a subset of safety countermeasures. Typically, these countermeasures are operation-based, e.g. an Advance Detection System (ADS) and an Advance Warning System (AWS).

The key features of VISSIM include driver and vehicle behavior management, a comprehensive toolbox for signal control that allows the user to define signal control logic, and flexibility in collecting disaggregated data. The basic VISSIM model contains a set of default parameter values related to the driver, the vehicle, and the system. However, it is crucial that these default values are adjusted for the specific problem being studied to ensure that the simulation output reflects reality. As these parameters directly affect modeled vehicle interactions, failure to properly calibrate them can result in erroneous conclusions. The calibration process for safety-related studies should include car-following, lane-changing, and signal control parameters.

This paper is focused on the calibration of a VISSIM microsimulation model (PTV, version 5.30, 2011) that will be utilized to examine ADS and AWS safety countermeasures. The model was calibrated for four high-speed, isolated, rural intersections in Nebraska (USA). Due to spatial constraints of the current paper, the process will be described with respect to a selected single location, the intersection of US-77 and Pioneers Boulevard outside of Lincoln, Nebraska. The calibration procedure utilized a genetic algorithm (GA). The GA procedure was selected because, by definition, it examined the entire solution space, and was less likely than other optimization techniques to identify a local minimum. Due to the stochastic nature of the model, nonparametric statistical techniques were also incorporated within the GA to ensure that empirical and simulated data were statistically identical. Because the study was safety-related, it was imperative that the resulting model adequately replicates the distribution of speeds, rather than the average speed only. Therefore, speed distribution was used as the primary calibration measure of effectiveness. Subsequently, the calibrated model was validated to ensure that the simulated speed distributions at key points matched the empirical distributions, that the simulated queues matched the observed queues, and that the simulated delays matched the observed delays. These comparisons were made using appropriate statistical tests. The end result was a calibrated and validated model for Nebraska that replicated driver behavior and could be utilized to compare various safety countermeasures at isolated, rural, high-speed signalized intersections.

2. Methodology

In recent years, a number of studies have been conducted in relation to the development of systematic calibration methodologies. Many of these studies were based on statistical comparisons and utilized GAs to identify potential parameter sets [3, 7, 8, 9, 11]. In addition, many made use of ITS data, which are becoming more widely available.

Mathew et al. [9] proposed heuristics and the GA based optimization for model calibration. The calibration parameters were identified through sensitivity analysis. The optimum values of these parameters were obtained by minimizing the error between the simulation and field delays using the GA technique. The authors found that the time and effort required to accurately tune a large number of potential simulation parameters could be reduced through the use of optimization methods. Park et al. [11] proposed an innovative calibration and validation procedure, successfully applying the approach to several case studies. The first step of the procedure required the identification of key calibration parameters and their acceptable ranges. The generation of a reasonable number of parameter sets using a statistical experimental design was then performed. Each parameter set was run five times to test the statistical feasibility of each set. The GA optimization program obtained an optimal calibration parameter set from the potential parameter ranges that were accepted during the feasibility test step. Since VISSIM is a microscopic and stochastic simulation model, a small number of runs was conducted for each feasible parameter set to reduce variability. An objective function of the GA was obtained through the comparison of field data to simulation output. A recent study indicated that an automatic calibration procedure could be more effective [7] than a manual approach. The authors used the GA procedure to determine ideal parameters. Five steps were proposed for the approach, and the procedure was iterative. The first step was the initialization of the GA. Next, the microscopic simulation model was run with the input file (generated parameters were translated into the appropriate VISSIM format). The model output and selection of the potential parameter set was then evaluated. This was the most important component in the calibration procedure. The model output was evaluated using statistical tests (Moses's test, the Wilcoxon test, and the Kolmogorov-Smirnov test). Finally, two descriptive statistics (median and dispersion) and the maximum difference in the cumulative function were tested using nonparametric testing methods. Cunto et al. [3] utilized a calibration procedure consisting of four steps. Their procedure included a heuristic selection of the initial model inputs; statistical screening using a Plackett–Burnman with factorial design; the development of a linear expression relating significant model inputs to safety performance (fractional factorial analysis); and the GA procedure to obtain best estimate model parameters. The next attempt to determine a formal calibration procedure was conducted by Park et al. [10], who implemented an experimental design. They argued that this was appropriate because the number of controllable feasible parameter combinations was so large that the set of possible scenarios could not be evaluated in a reasonable amount of time. This problem was compounded if multiple simulation runs were required to reduce stochastic variability. The authors also used statistical tests, including the t-test and the Kolmogorov–Smirnov test, to determine how well their calibration procedure performed.

Calibration is a process used to determine an appropriate set of model parameters that replicate observed information, such as local driver behavior. In general, the user adjusts the select behavioral parameters until the simulated output matches the empirical data at some predefined statistical level. Validation is then employed to ensure that the calibrated model is appropriate by determining whether the output of the model accurately reproduces the specified behavior. Similar to calibration, validation compares empirical data with simulated data. However, this empirical data need to be different than that used in the calibration.

The objective of the current paper was to demonstrate the calibration and validation of a VISSIM microsimulation model for signalized intersections in Nebraska. The model will be utilized to study ADS and AWS systems, which are engineering countermeasures intended to improve road safety. As such, it was critical that the model accurately reflects driver behavior, as measured by vehicle speeds, when these systems were active. A microsimulation model was developed for four Nebraska test sites. Key inputs included traffic volume, turning movements, and heavy vehicle percentages. The models were then calibrated based on speed distributions at particular points at each location. The calibrated model was then validated by examining speed distributions at other locations and waiting times on the minor approaches.

Fig. 1 displays the calibration procedure adopted in the current study, which is described more thoroughly in subsequent sections.

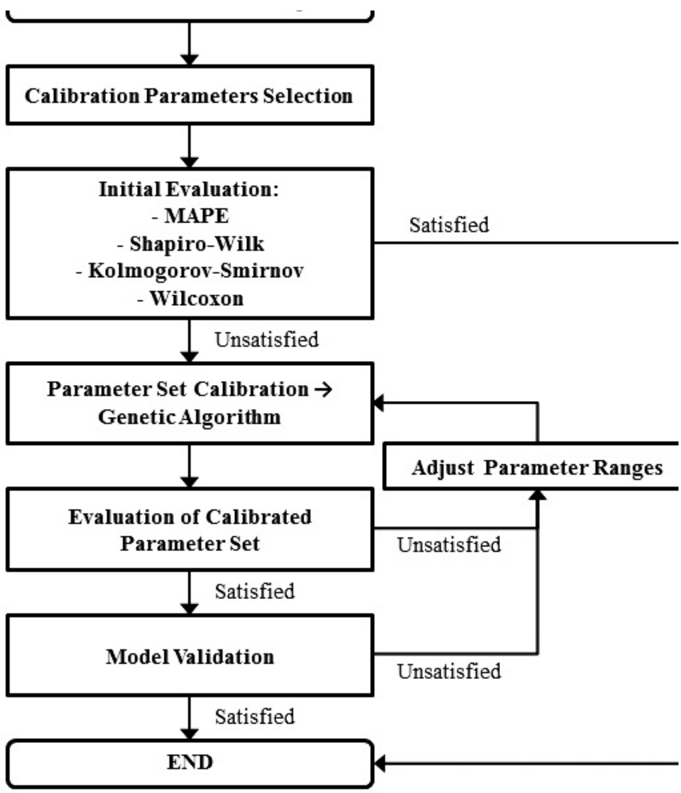

Fig. 1. Procedure for model calibration

3. Calibration procedure

The procedure of the calibration, presented in the current paper, was applied to four intersections in Nebraska: (1) US Highway 77 and Saltillo Rd in Lincoln, (2) US Highway 77 and Pioneers Boulevard in Lincoln, (3) Highway N-133 and Highway N-36 in Omaha, and (4) US Highway 75 and Platteview Road in Bellevue. However, based on space limitations of the current document, the US-77 and Pioneers Blvd test site was treated as the primary focus of the current paper.

3.1. Simulation model setup

The intersection of the US-77 and Pioneers Blvd was located in a rural area approximately five miles south of Lincoln. US Highway 77 is a four-lane divided expressway, as pictured in Fig. 2.

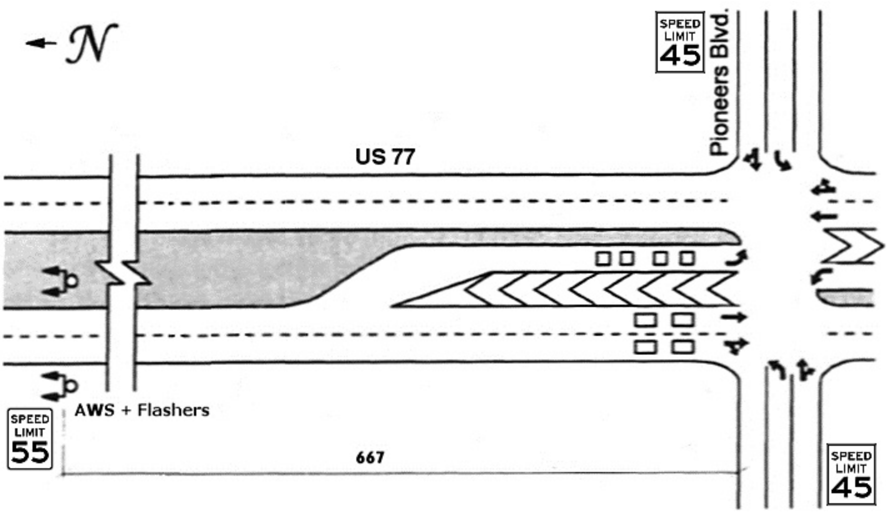

Fig. 2. Study site – US-77 & Pioneers Blvd in Lincoln, NE

The speed limit on US-77 before the intersection was 55 mph, and on Pioneers Blvd was 45 mph in both directions. The SB approach of US-77 was examined because it had been outfitted with an AWS by the Nebraska Department of Roads (NDOR) [14]. This approach had an exclusive left-turn lane, one through lane, and one shared through and right-turn lane. All lanes were 12 ft wide, while the left-turn lane was 180 ft long. The lane configuration for this approach, with detector locations, is pictured in Fig. 2. Additionally, the EB and WB approaches on Pioneers Boulevard had an exclusive left-turn lane and shared through with the right-turn lane. The advance warning sign, which operates in pulse mode, was located 667 ft before the intersection, as shown in Fig. 2. Stop line detection was provided for all approach lanes, and the detectors operated in presence mode. The SB approach on US-77 had one phase, which included all movements. It had a 15.0 sec minimum green, a 2.0 sec passage time, a 50.0 sec maximum green, a 4.5 sec yellow, and a 0.5 sec all-red interval.

This intersection was coded in VISSIM software and the plan and overall geometry of the intersection were based on a scaled map obtained from Google [5]. Lane width and length were confirmed on site. Additional data pertaining to the intersection, including signal timing and the location of the safety countermeasures (AWS and detectors), was obtained from materials provided by NDOR. Empirical traffic data, including volumes, turning movement counts, and heavy vehicles percentages, were collected by the authors. All VISSIM parameters were initially set to the default values. The signal control logic was coded using vehicle actuated programming (VAP) to enable the phase-based, traffic-actuated signal control logic to be implemented in the microsimulation [13].

Due to the nature of the traffic safety countermeasures being analyzed, it was critical that the speed distribution of drivers at critical locations be adequately modeled. Therefore, the empirical data collection effort focused on speed distribution as a function of space and time. The speed and location of every northbound vehicle were collected using wide area detectors (WAD) mounted on a mobile trailer. These devices were connected to programmable controllers. Three WADs were used to collect data: two Wavetronix SmartSensor Advance models and one Wavetronix SmartSensor HD. The two Wavetronix SmartSensor Advance models were used to track approaching vehicles upstream and downstream of the trailer location. The sensors recorded distance and speed at 0.1 sec intervals. The Wavetronix SmartSensor HD was used to record vehicle information equidistant with the pole location [1]. More detailed information on the specific capabilities of these devices can be obtained from the Wavetronix website [6]. The data collection scheme, including the location of each WAD and their coverage (up to 600 ft upstream and downstream), is illustrated in Fig. 3.

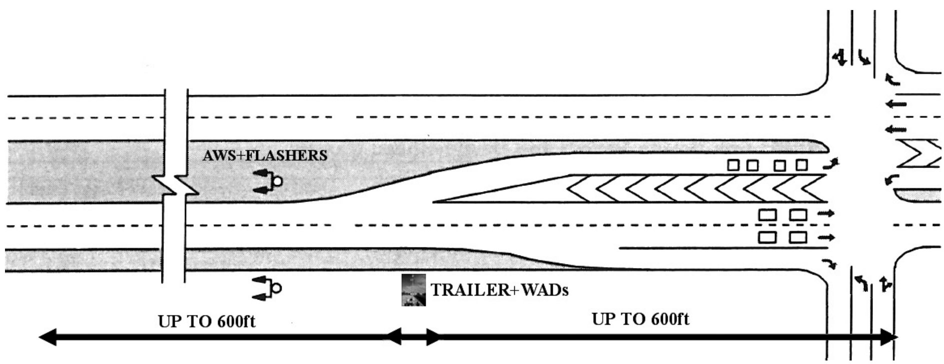

Fig. 3. WADs data coverage [14]

In calibrating a microsimulation model, it is critical that the end application is known so that key metrics are appropriately modeled. For example, for safety applications, the distribution of metrics, such as speed, is often required; however, if the model is calibrated to mean speed, it may fail to perform in the desired manner. The motivation behind the current paper was to analyze two safety countermeasures (ADS and AWS). A complete description of these systems can be found elsewhere [14].

To model the ADS countermeasure, two critical locations were identified, as illustrated in Fig. 4. The first location, denoted as DS1, was 2,000 ft upstream of the stop bar. This location was selected because the status of the traffic signal was deemed to have no effect upon observed speed. The second location, denoted as DS2, was 600 feet from the stop bar. This location was selected in order to model driver behavior in the vicinity of the intersection as a function of traffic signal status.

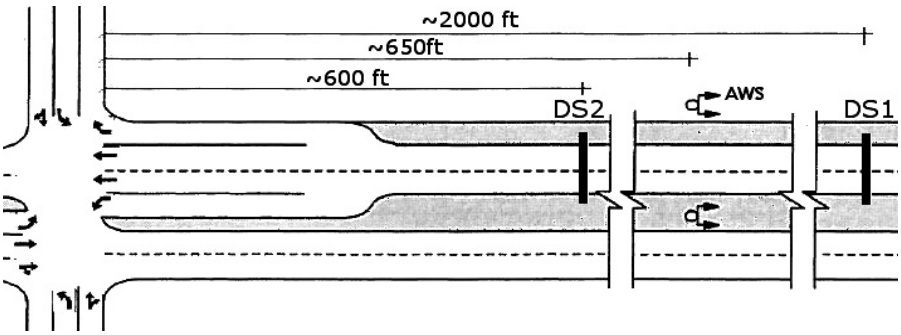

Fig. 4. Location of desired speed decision points [14]

Empirical speed distribution data were collected at both locations, and are presented in Fig. 5. Distribution DS1 corresponded to location DS1, and distribution DS2a corresponded to location DS2. In order to model speed change, a decision point was input in VISSIM. In essence, the desired speed distribution was added as part of the input parameter set. The simulated vehicles were then assigned a speed from this input distribution when they crossed the point. In this way, the behavior of drivers slowing down as they approached the intersection was modeled.

There were three key points to note from Fig. 5. First: because the ADS could not be detected by drivers in real life, it was not necessary to model driver reactions to that system in VISSIM. Second, users are able to add as many decision points as they wish to simulate driver deceleration upon approaching an intersection. While two decision points were adequate for the four test sites analyzed in the current project, other test sites may require additional decision points. Third, the distribution is defined by the user. In the current paper, five points were used to define the curve. However, users may choose as many points as desired in order to model the desired speed distribution.

In essence, the AWS simulation followed ADS logic. The primary difference was that the AWS utilized an AWS device (i.e., a sign with the flashers). Because drivers react to active AWS signs, it was necessary to model the reactions of simulated drivers to the AWS sign in VISSIM. In Nebraska, the recommend sign location was 650 ft from the intersection stop bar. The flasher activated (i.e., began to flash) a few seconds (typically 7 sec, but 8 sec in this case) before the traffic signal transitioned from green to yellow. Any driver upstream of the AWS sign was then in a position to slow and come to a gradual stop. To model this component of driver behavior in VISSIM, two speed distributions were utilized at location DS2: DS2a and DS2b. When AWS flashers were inactive (i.e., from the end of the red phase until 8 sec prior to the transition from

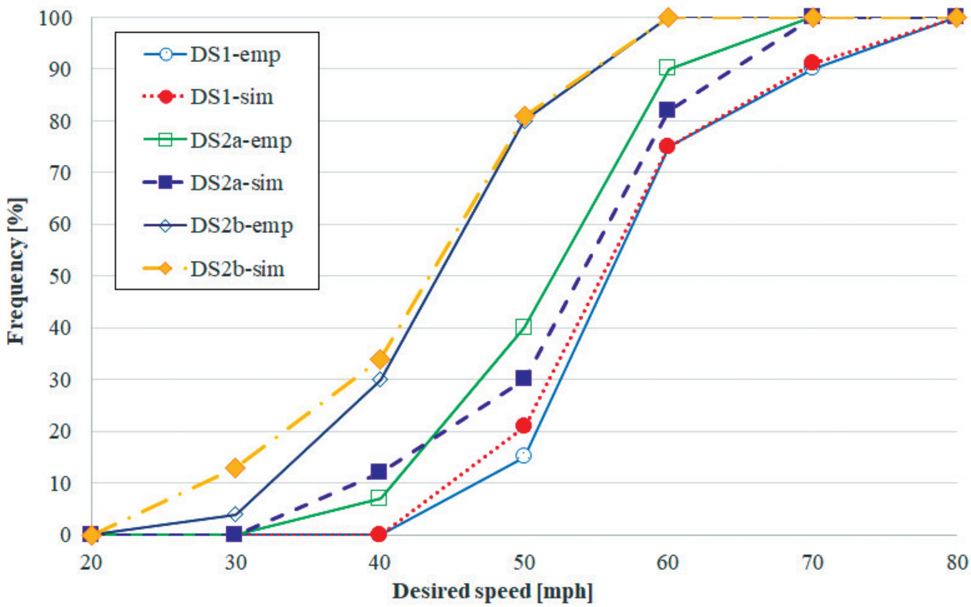

Fig. 5. Empirical and simulated speed distributions at the DS1 and DS2 locations

green to yellow), vehicles crossing DS2 followed the speed distribution DS2a, as detailed in Fig. 5. This was the same speed distribution followed by all vehicles in the ADS simulation. When the AWS flashers were active (i.e., eight 8 sec prior to yellow phase and then during the yellow and red phases) drivers followed the adjusted speed distribution DS2b, as shown in Fig. 5. In this latter situation, each vehicle that crossed location DS2 was assigned a new speed from the speed distribution DS2b, and was traveling, on average, at the slower rate [13]. Therefore, when the AWS was active, simulated drivers traveled, on average, at a slower rate compared to periods when the AWS was off (i.e., during the green signal and until 8 sec prior to the yellow signal). It was further assumed that a change in speed would occur at DS2. Note that the model could be easily adjusted to add additional decision points to reflect a more gradual change in driver speeds. Additionally, when the system was active, change in speed distribution implied that a percentage of vehicles potentially sped up to beat the yellow signal. Note that the percentage of drivers that exhibited this behavior was controlled by the user through the input speed distribution. However, red light running is not modeled in VISSIM, so therefore the number of vehicles entering the intersection during the red signal was precluded from being an MOE in the current analysis.

3.2. Selection of calibration parameters

Calibration was carried out for the parameters of car-following, lane change, desired speed distribution, and signal control [13]. The initial set of VISSIM parameters utilized in the calibration was identified and selected based upon engineering judgment and the review of the salient literature [3, 4, 9–12]. All 19 parameters and their acceptable ranges, identified in the literature review, are presented in Table 1.

Table 1. Set of VISSIM parameters for the initial evaluation and calibration

BEHAVIOR	PARAMETER		DEFAULT VALUE	ACCEPTABLE RANGE	UNIT
CAR FOLLOWING	Number of observed preceding vehicles	NUMB_PRECED	2	1 – 4	-
	Maximum look ahead distance	OBS_DISTANCE MAX	820.21	700 – 900	ft
	CC0 (Standstill Distance)	CC0	4.92	3.0 – 6.0	ft
	CC1 (Headway Time)	CC1	0.9	0.5 – 3.0	s
	CC2 ('Following' Variation)	CC2	13.12	1.0 – 30.0	ft
LANE CHANGE	Waiting time before diffusion	T_DISAPPEAR	60.0	30 – 90	s
	Minimum headway (front/rear)	MIN_LC_GAP	1.64	0.5 – 3.0	ft
DESIRED SPEED 1	Frequency at 50mph	DESIRED_SPEED 1	0.150	0.10 – 0.30	-
	Frequency at 60mph	DESIRED_SPEED 1	0.750	0.60 – 0.80	-
	Frequency at 70mph	DESIRED_SPEED 1	0.900	0.85 – 0.95	-
DESIRED SPEED 2a	Frequency at 40mph	DESIRED_SPEED 2	0.070	0.05 – 0.15	-
	Frequency at 50mph	DESIRED_SPEED 2	0.400	0.30 – 0.50	-
	Frequency at 60mph	DESIRED_SPEED 2	0.900	0.80 – 0.95	-
DESIRED SPEED 2b	Frequency at 30mph	DESIRED_SPEED 3	0.040	0.03 – 0.15	-
	Frequency at 40mph	DESIRED_SPEED 3	0.300	0.25 – 0.45	-
	Frequency at 50mph	DESIRED_SPEED 3	0.800	0.70 – 0.90	-
SIGNAL CONTROL	Reaction to amber signal: α	AMBER_ALPHA	1.59	1.0 – 15.0	-
	Reaction to amber signal: β1	AMBER_BETA1	-0.26	- 0.40 – -0.20	-
	Reaction to amber signal: β2	AMBER_BETA2	0.27	0.10 – 0.30	-

The default values of all parameters were utilized in the initial evaluation of the simulation model. The definitions of these parameters and their functions in VISSIM can be obtained from the VISSIM manual [13], as well as from a number of papers describing VISSIM model calibration [3, 4, 9–12].

3.3. Initial evaluation

The primary goal of the initial evaluation was to determine whether the default model (e.g. based on the default values of the parameters) was able to adequately model real traffic conditions at the test intersection. If the output did not match the empirical data, it became necessary to conduct additional steps in the calibration procedure.

A calibration procedure was performed to obtain a set of driving behavior parameters that would result in simulation results similar to the observed empirical values. As discussed previously, the procedure involved the use of a GA and a variety of metrics and associated statistical tests.

Speed–in particular, speed distribution–was selected as a criterion because it effectively characterized the nature of road traffic and could also be used to measure safety. Following each simulation run, the approach speeds of all vehicles were recorded at the cross-section 900 ft from the stop bar. This location was selected for the provision of speed data reflecting driver behavior under the influence of activated flashers. A speed distribution was created and parameters were calculated: mean, median, mode, standard deviation, and kurtosis. Mean absolute percentage error (MAPE) was used as a measure of accuracy to determine the difference between the empirical and simulated speed distributions, as demonstrated in Equation (1).

$$\mathrm{MAPE} = \frac{1}{5} \cdot \sum_{i=1}^{5} \frac{R_i - S_i}{R_i} \tag{1}$$

where: R_i = empirical speed mean, median, mode, standard deviation, and kurtosis;
S_i = simulated speed mean, median, mode, standard deviation, and kurtosis.

A MAPE value of less than 5% was targeted to indicate sufficient model fit and merit further analysis.

In addition, the Shapiro-Wilk (S-W) and Kolmogorov-Smirnov (K-S) tests were performed to check the normality of the approach speed. The K-S and Wilcoxon tests were also used to test the hypothesis that the empirical and simulated speed distributions were identical. Descriptions of all statistical tests utilized in the current study can be found elsewhere [14].

After 20 simulation runs, based on the default values of the select VISSIM parameters, the lowest noted MAPE value was 5.4% (see Table 2).

Table 2. Average MAPE in the initial model evaluation

Distribution parameter	Unit	Simulated speed	Empirical speed	MAPE [%]
Mean	[mph]	56.25	54.92	2.4
Median	[mph]	55	55	0.0
Mode	[mph]	58	58	0.0
Standard Deviation	-	7.76	6.30	23.2
Kurtosis	-	0.295	0.290	1.6
			avg MAPE =	**5.4**

The difference between the four distribution parameters was less than 5%, but the difference in standard deviation was higher than 20%. That is, in comparison to the empirical data, simulation speed was distributed more evenly. The MAPE results indicated that a simulation model using the default parameter set could be satisfactory for this type of analysis; however, as discussed previously, simply using the MAPE may not provide adequate results if the user is interested in the distribution of a particular metric, such as speed.

While the functions were similar, it can be seen in Fig. 6 that the greatest differences occurred within the 46-53 mph range, and at speeds above 58 mph. Only for low speeds (< 46 mph), and for speeds in the range of 53-58 mph, were the distributions close. Overall, the simulated speed distribution had higher tails than did the empirical speed distribution. Tests for normality were

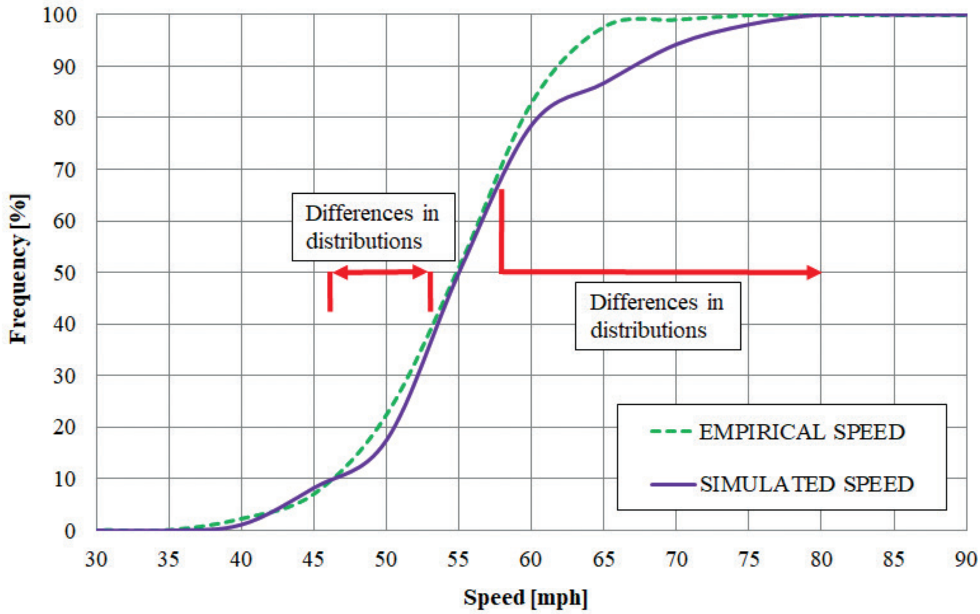

Fig. 6. Graphical comparison of speed distributions

performed to determine whether both speed distributions were normally distributed. The S-W test indicated that, at the 5% significance level, empirical speed (p = 0.0499) was normally distributed, but simulated speed (p < 0.0001) was not normally distributed. Additionally, the K-S test confirmed the non-normality of both speed distributions with p < 0.01. The K-S and Wilcoxon tests were also used to check the equality of the distributions, providing separate results. The K-S test rejected the hypothesis that both speed distributions resulted from the same continuous distribution (p = 0.0352 and < 0.05). At the same time, the Wilcoxon test accepted the hypothesis that the analyzed speed distributions were equal (p = 0.1919).

In summary, the microsimulation model resulted in a MAPE slightly higher than 5%, and failed the normality test. Additionally, the K-S and Wilcoxon tests provided different results as to the question of whether both speed distributions follow the same continuous distribution. Therefore, it was determined that the model may not have been acceptable and required calibration to facilitate further analysis. A full analysis can be found elsewhere [14].

3.4. Parameter set calibration

The calibration procedure was designed to identify the "best" parameter set for a given problem. In the current study, 19 driving behavior-related parameters were selected for testing. The desired parameter set depended on the MAPE and the results of the statistical tests. Following the calibration, the 10 parameter sets with the lowest MAPE value, having passed all statistical tests, were output.

As discussed previously, many researchers have identified the GA method as an appropriate tool for calibrating traffic microsimulation models. Based on these experiences, a GA was selected for the current research. The theory of the genetic algorithm is based on the Darwinian biological evolutionary processes occurring in nature; as such, it is ideal for solving complex problems. Put simply, a GA sets select parameters as "chromosomes" [8]. The GA for each task must include the following elements: the representation of potential solutions, a method to create an initial population, an evaluation function, a selection procedure, basic operators, and the values of parameters, such as population size, etc. The biggest advantage of using a GA is that, as a product of its approach, it considerably reduces the number of search steps and the amount of time required to determine a solution to a given problem. Fig. 7 illustrates the calibration procedure using GA.

The objective of the calibration process was to minimize MAPE values. First, agents that represented all select parameters needed to be defined. GA uses agent and gene terms, where a gene is represented by the binary digits 0 or 1. One agent is defined as a group of genes

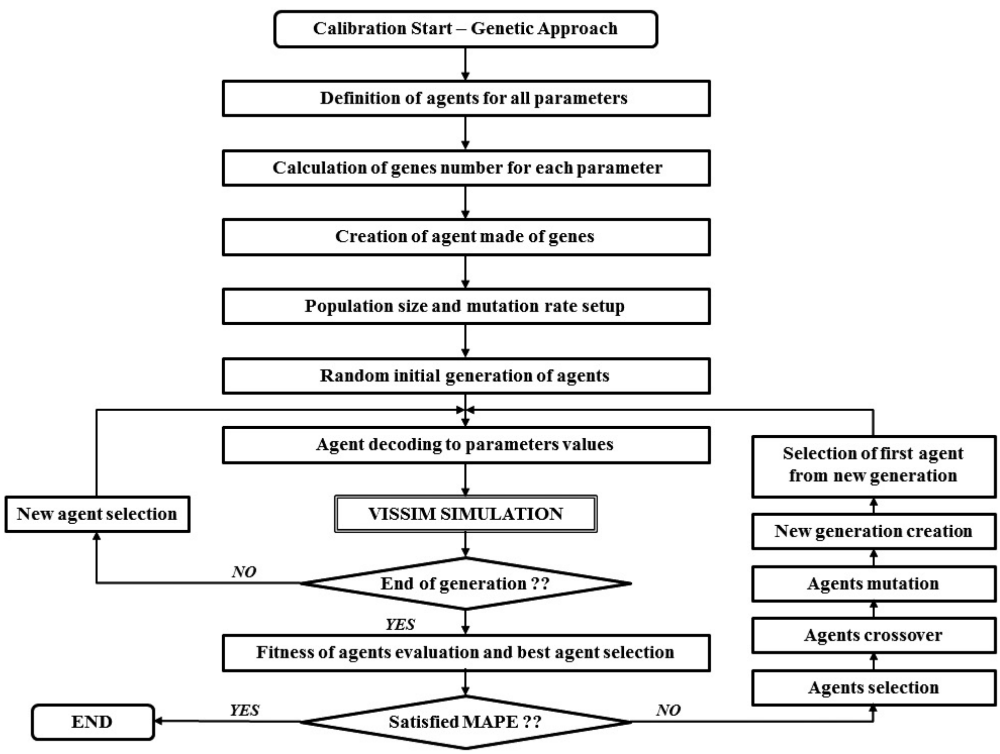

Fig. 7. Procedure based on GA for VISSIM model calibration [14]

used to represent a value of each parameter. One generation is then defined as the specified number of agents. The size of the population is defined as the number of agents included in one generation [15]. Second, for each calibrated parameter, the number of genes needed to be estimated, which depended on the characteristics of the parameter and the increment of its value. The number of genes (n_i) for each parameter was calculated using the following equation:

$$n_i = \log_2\left(\frac{\max(x_i) - \min(x_i)}{\alpha_i} + 1\right) \quad (2)$$

Where: x_i = value of i parameter;
$max(x_i)$ = maximum value of x_i;
$min(x_i)$ = minimum value of x_i;
a_i = increment value of $x_{i.}$

Prior to calculation, the acceptable range of values (i.e., $max(x_i)$ and $min(x_i)$ values) for every parameter needed to be identified. The results of the calculations for all 19 VISSIM parameters are presented in Table 3.

Table 3. Number of genes with increments for calibrated parameters

PARAMETER		Min (x_i)	Max (x_i)	Number of genes (n_i)	Increment (a_i)
Number of observed preceding vehicles	NUMB_PRECED	1	4	2	1
Maximum look ahead distance	OBS_DISTANCE MAX	700	900	5	7
CC0 (Standstill Distance)	CC0	3	6	3	0.5
CC1 (Headway Time)	CC1	0.5	3	5	0.1
CC2 ('Following' Variation)	CC2	1	30	5	1
Waiting time before diffusion	T_DISAPPEAR	30	90	6	1
Minimum headway (front/rear)	MIN_LC_GAP	0.5	3	5	0.1
Frequency at 50mph	DESIRED_SPEED 1	0.10	0.30	5	0.007
Frequency at 60mph	DESIRED_SPEED 1	0.60	0.80	5	0.007
Frequency at 70mph	DESIRED_SPEED 1	0.85	0.95	4	0.007
Frequency at 40mph	DESIRED_SPEED 2a	0.05	0.15	4	0.007
Frequency at 50mph	DESIRED_SPEED 2a	0.30	0.50	5	0.007
Frequency at 60mph	DESIRED_SPEED 2a	0.80	0.95	4	0.01
Frequency at 30mph	DESIRED_SPEED 2b	0.03	0.15	4	0.008
Frequency at 40mph	DESIRED_SPEED 2b	0.25	0.45	5	0.007
Frequency at 50mph	DESIRED_SPEED 2b	0.70	0.90	5	0.007
Reaction to amber signal: α	AMBER_ALPHA	1	15	6	0.23
Reaction to amber signal: β1	AMBER_BETA1	-0.4	-0.2	3	0.03
Reaction to amber signal: β2	AMBER_BETA2	0.1	0.3	3	0.03
				84	

In the next step of the GA approach, the initial generation, based on agents, was set. These agents were composed of genes. Table 3 illustrates that there were 84 genes representing the 19 calibrated parameters in the current study. A population size of 32 was defined as the number of agents in one generation. In the initial generation, each gene of the agent was randomly assigned a 0 or 1. To decode each agent to the value of the parameter value $x_{i,}$ the following approach was adopted:

$$x_i = \alpha_i \cdot A \cdot B + \beta \tag{3}$$

Where: $A = (a_1, a_2, a_3, \ldots . a_n)$

$$B = \begin{pmatrix} 2^{n_i - 1} \\ . \\ . \\ . \\ 2 \\ 1 \end{pmatrix}$$

$i = 1, 2, 3, \ldots , 19$
x_i = value of i parameter
a_i = increment value of x_i (0 or 1)
$b_i = \min(x_i)$ as shown in Table 3
A = vector representing agent
B = coefficient vector
n_i = number of genes of agent representing x_i

MAPE, as a measure of accuracy or fitness, was utilized to evaluate agent quality. The given parameter set was deemed to be more ideal if the fitness value was higher (e.g. lower MAPE). The procedure was automated, as shown in Fig. 8. Following each generation, the assessment of each agent was executed based on MAPE, and the best agent was selected.

When one generation was complete and agents were evaluated, steps for agent selection, crossover, and mutate were performed, resulting in the subsequent generation of agents. The selection was based on probability, and agents with lower MAPE values were the most likely to be selected. To crossover, two agents interchanged part of their genes to create two new agents. One agent was mutated to create a new agent by changing one of its genes from 1 to 0, or from 0 to 1 (15). After these steps were completed for the agents of the previous generation, additional agents were presented to create a new generation.

The GA described in the current paper was implemented using the MatLab software. The toolbox developed by the University of Sheffield provided all the necessary functions to implement the GA operators, i.e., selection, crossover, and mutation (2). The complete calibration procedure used in this study combined the MatLab software, Visual Basic, GA toolbox, and VISSIM.

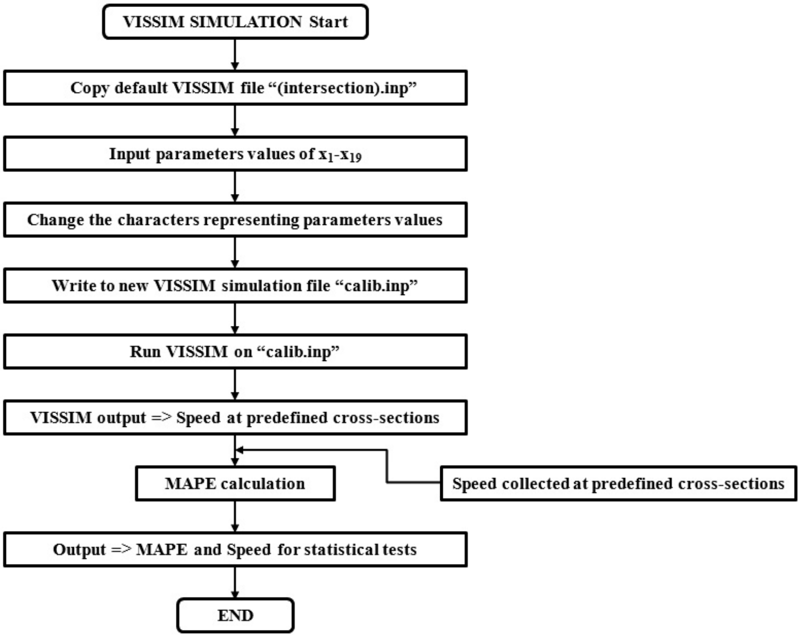

Fig. 8. Procedure for VISSIM model calibration [14]

3.5. Calibration results

After nine generations of the GA, the procedure was complete, and 10 sets of parameters with the lowest MAPE values were recorded, as shown in Table 4.

Table 4. Calibration results for the intersection of US-77 & Pioneers Blvd

		DEFAULT VALUE	1	2	3	4	5	6	7	8	9	10
PARAMETER	NUMB_ PRECED	2	4	4	3	3	3	3	3	4	3	4
	OBS_ DISTANCE MAX	820.21	889	889	791	791	791	791	791	847	700	889
	CC0	4.92	3.5	3.5	6.5	6.5	6.5	6.5	6.5	3	5.5	3.5
	CC1	0.9	1.9	1.3	0.8	1.6	2.9	0.8	0.8	3.6	3.1	1.9
	CC2	13.12	1	18	12	12	18	12	12	20	8	1
	T_DISAPPEAR	60.0	70.0	54.0	34.0	34.0	54.0	34.0	34.0	49.0	62.0	70.0
	MIN_LC_GAP	1.64	3	3.5	3.1	3.1	3.5	3.1	3.1	2.1	3.5	3
	DESIRED _SPEED 1	0.150	0.210	0.270	0.230	0.230	0.270	0.230	0.230	0.230	0.210	0.210
		0.750	0.750	0.810	0.810	0.810	0.810	0.810	0.810	0.750	0.740	0.750
		0.900	0.910	0.950	0.950	0.950	0.950	0.950	0.950	0.930	0.890	0.910
	DESIRED _SPEED 2a	0.070	0.120	0.110	0.110	0.110	0.110	0.110	0.110	0.060	0.100	0.120
		0.400	0.300	0.520	0.520	0.520	0.520	0.520	0.520	0.300	0.330	0.300
		0.900	0.820	0.900	0.940	0.940	0.900	0.940	0.940	0.840	0.880	0.820

PARAMETER	DESIRED _SPEED 2b	0.040	0.130	0.130	0.140	0.140	0.130	0.140	0.140	0.130	0.130	0.130
		0.300	0.340	0.340	0.290	0.290	0.340	0.290	0.290	0.380	0.380	0.340
		0.800	0.810	0.870	0.830	0.890	0.880	0.830	0.830	0.720	0.720	0.810
	AMBER_ ALPHA	1.59	12.5	3.53	7.9	7.9	3.53	7.9	7.9	1.23	1	12.04
	AMBER_BETA1	-0.26	-0.25	-0.22	-0.34	-0.34	-0.22	-0.34	-0.34	-0.37	-0.37	-0.37
	AMBER_BETA2	0.27	0.22	0.25	0.22	0.22	0.25	0.22	0.22	0.31	0.13	0.13
TESTS	MAPE [%]	5.45	4.96	5.34	5.8	5.88	6	6.13	6.23	6.23	6.35	6.39
	Kolmogorov-Smirnov (Passed/Not Passed)	N	P	N	P	N	N	P	P	P	P	N
	Wilcoxon (Passed/Not Passed)	P	P	N	P	N	N	P	P	P	P	P

Additionally, each parameter set was checked to verify that the simulated speed distribution and empirical speed distribution were equally distributed, and to determine whether they displayed the same continuous distribution. The latter check was performed using the K-S and Wilcoxon tests. The set with the lowest value of MAPE (4.96%) passed the non-parametric tests (K-S and Wilcoxon), which meant that the empirical approach speed and simulated approach speed distributions originated from the same continuous distribution. Consequently, this parameter set was utilized for further evaluation. While the MAPE values for the calibrated and default parameter sets were similar, the distributions from the latter

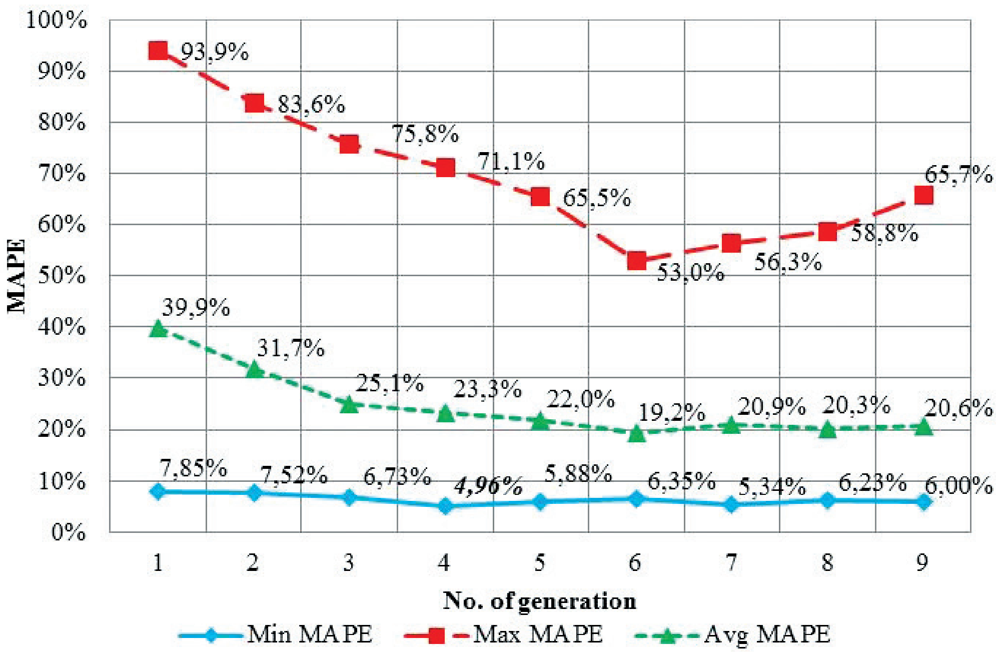

Fig. 9. Calibration results for the analyzed intersection

scenario did not differ statistically from the empirical distribution. Therefore, the calibrated parameter set was utilized in subsequent analyses.

Fig. 9 displays the results for each generation in the GA.

As can be seen, following nine generations, MAPE values were nearly steady; consequently, the algorithm was terminated.

3.6. Evaluation of the calibrated parameter set

Table 5 displays a comparison between the simulated results from the default parameter set and the calibrated parameter set.

It can be seen that the calibrated results had lower MAPE values, as calculated using the five measures of speed distribution. In addition, the calibrated model passed statistical testing, indicating that the simulated and empirical speed distributions were statistically equivalent. The calibrated VISSIM parameter set is displayed in Table 6.

Table 5. Evaluation of calibrated parameter sets

	Distribution Parameter	MICROSIMULATION MODEL	
		Default	Calibrated
MAPE [%]	Mean	2.4	1.8
	Median	0.0	0.0
	Mode	0.0	0.0
	Standard Deviation	23.2	22.3
	Kurtosis	1.6	0.8
TEST	Average MAPE [%]	5.40	4.96
	K-S test	Not Passed	Passed
	Wilcoxon test	Passed	Passed
	MAPE Improvement	8%	

Table 6. Calibrated parameter set for microsimulation model

BEHAVIOR	PARAMETER		DEFAULT VALUE	CALIBRATED MODEL
CAR FOLLOWING	Number of observed preceding vehicles	NUMB_PRECED	2	4
	Maximum look ahead distance	OBS_DISTANCE MAX	820.21	889
	CC0 (Standstill Distance)	CC0	4.92	3.5
	CC1 (Headway Time)	CC1	0.9	1.9
	CC2 ('Following' Variation)	CC2	13.12	1

LANE CHANGE	Waiting time before diffusion	T_DISAPPEAR	60.0	70.0
	Minimum headway (front/rear)	MIN_LC_GAP	1.64	3.0
DESIRED SPEED 1	% at 50mph	DESIRED_SPEED 1	0.150	0.210
	% at 60mph	DESIRED_SPEED 1	0.750	0.750
	% at 70mph	DESIRED_SPEED 1	0.900	0.910
DESIRED SPEED 2a	% at 40mph	DESIRED_SPEED 2a	0.070	0.120
	% at 50mph	DESIRED_SPEED 2a	0.400	0.300
	% at 60mph	DESIRED_SPEED 2a	0.900	0.820
DESIRED SPEED 2b	% at 30mph	DESIRED_SPEED 2b	0.040	0.130
	% at 40mph	DESIRED_SPEED 2b	0.300	0.340
	% at 50mph	DESIRED_SPEED 2b	0.800	0.810
SIGNAL CONTROL	Reaction to amber signal: α	AMBER_ALPHA	1.59	12.5
	Reaction to amber signal: β1	AMBER_BETA1	-0.26	-0.25
	Reaction to amber signal: β2	AMBER_BETA2	0.27	0.22

3.7. Model validation

As mentioned previously, validation was performed to determine whether the calibrated microsimulation model performed properly by comparing the output with data not utilized in the original calibration. The calibrated model was run 10 times to determine whether the model was capable of reflecting actual traffic conditions at the test intersection. Waiting time on the minor approaches was used as a validation parameter. Real values of waiting times were gathered during data collection. The mean value of waiting time for the EB and WB approaches, the output from 10 simulation runs, was compared to the values derived from the collected empirical data. The results are displayed in Table 7.

Table 7. Summary of calibrated model validation

Test intersection	Minor approach	Average waiting time [s]		t-test (p)	Result (Passed /Not passed)	Validation
		Empirical data	Simulated data			
US-77 & Pioneers Blvd	EB	21.6	22.2	0.397	Passed	**YES**
	WB	17.1	16.3	0.395	Passed	

Two statistical tests, the F-test, and the t-test, were performed to determine whether statistically significant differences existed between the analyzed values at the 5% level of confidence. It can be observed that the statistical tests were passed ($p = 0{,}397 > 0{,}05$ and $p = 0{,}395 > 0{,}05$), indicating that the microsimulation model could effectively mimic real traffic at the signalized intersection [14] by providing the values of the wait time on minor approaches that were proximate to the empirical values determined through data collection (see Table 7). Therefore, the calibrated model was acceptable, and was deemed eligible for further safety analysis.

4. Conclusions

This study resulted in a calibrated stochastic model of a signalized intersection located in Nebraska. The model was calibrated using a genetic algorithm with non-parametrical statistical tests. The GA approach provided a quick and effective method for finding the "best" set of VISSIM parameters, and seemed to be a very effective tool for the calibration of traffic and the development of a stochastic simulation model for the studied signalized intersection. An innovative approach described in this paper utilized speed distribution as the objective function in the calibration process. While other researchers have calibrated microsimulation models to measures of central tendency (e.g. mean), this paper proposed calibrating to approach speed distribution. Mean absolute percentage error, calculated for the speed distribution parameters, was used as a measure of fitness. Non-parametric statistical tests were then utilized to indicate whether the subsequent value of MAPE was acceptable at a 5% level of significance (see Table 2). The distributions of the observed and calibrated speeds were compared, and it was determined that no statistically significant differences existed at the 5% confidence level (see Table 5). The calibration process resulted in a microsimulation model (see Table 6), which was also validated to the wait time on the minor approaches (see Table 7).

The VISSIM model that was developed, calibrated, and validated in this paper could be used as the basis for a methodology to implement specific safety countermeasures at signalized intersections.

References

[1] Burnett N. P., *Effect of information on driver's risk at the onset of yellow*, Master Thesis at University of Nebraska-Lincoln, July, 2011.

[2] Cao Y., Wu Q., *Teaching Genetic Algorithm using Matlab*, International Journal of Electrical Engineering Education, Vol. 36/1999, 139–153.

[3] Cunto F., Saccomanno F., *Calibration and validation of simulated vehicle safety performance at signalized intersections*, Accident, Analysis and Prevention, 40/2008, 1171–1179.

[4] Duong D., Saccomanno F., Hellinga B., *Calibration of microscopic traffic model for simulating safety performance*, Compendium of papers of the 89th Annual TRB Conference held Jan. 10–14, 2010 in Washington, D.C., No. 10-0858/2010.

[5] http://maps.google.com (access: 01.05.2010).
[6] www.wavetronix.com/en/products (access: 01.05.2010).
[7] Kim S-J., Kim W., Rilett L.R., *Calibration of microsimulation models using nonparametric statistical techniques*, Transportation Research Record, No. 1935/2005, 111–119.
[8] Kim K., Rilett L.R., *Genetic-algorithm-based approach for calibrating microscopic simulation models*, 2001 IEEE Intelligent Transportation Systems Conference Proceedings, 2001, 298-–704.
[9] Mathew T.V., Radhakrishnan P., *Calibration of microsimulation models for nonlane-based heterogeneous traffic at signalized intersections*, ASCE Journal of Urban Planning and Development, Vol. 136/2010, Issue 1, 59–66.
[10] Park B., Schneeberger J. D., *Microscopic simulation model calibration and validation: Case study of VISSIM simulation model for a coordinated actuated signal system*, Transportation Research Record, No. 1856/2003, 185–192.
[11] Park B., Qi H., *Development and evaluation of a procedure for the calibration of simulation models*, Transportation Research Record, No. 1934/2005, 208–217.
[12] Santhanam S., Park B., *Development of VISSIM Base Model for Northern Virginia (NOVA) Freeway System*, Report No. UVACTS-13-0-124, June 2008.
[13] VISSIM 5.30-05 User Manual, PTV Planung Transport Verkehr AG, Innovative Transportation Concepts, 2011.
[14] Wojtal R., *Development of a methodology for analyzing safety treatments at isolated signalized intersections*, Ph.D. Dissertation, University of Nebraska-Lincoln, Lincoln 2012.
[15] Yu L., Lei Y., Xumei C., Tao W., Jifu G., *Calibration of VISSIM for bus rapid transit systems in Beijing using GPS data*, Journal of Public Transportation, No. 3/2006, 239–257.
[16] Zalzala A., Fleming P., *Genetic algorithms in engineering systems*, The Institution of Engineering and Technology, February, 1999.