

# CALIBRATION OF MICRO-SIMULATION MODELS TO ACCOUNT FOR SAFETY AND OPERATION FACTORS FOR TRAFFIC CONFLICT RISK ANALYSIS

**Hong Yang**

Graduate Research Assistant,  
Rutgers Intelligent Transportation Systems (RITS) Laboratory,  
Department of Civil and Environmental Engineering,  
Rutgers, The State University of New Jersey, Piscataway, NJ, USA,  
E-mail: yanghong@eden.rutgers.edu

**Kaan Ozbay**

Professor, Rutgers Intelligent Transportation Systems (RITS) Laboratory,  
Department of Civil and Environmental Engineering,  
Rutgers, The State University of New Jersey, Piscataway, NJ, USA,  
E-mail: kaan@rci.rutgers.edu

*Submitted to the 3<sup>rd</sup> International Conference on Road Safety and Simulation,  
September 14-16, 2011, Indianapolis, USA*

## ABSTRACT

Road safety is a critical issue for transportation systems. The use of crash data-based methodologies to analyze traffic safety problems has been problematic due to shortcomings such as unavailability and low quality of historical crash data. Other than crash data-based analysis, development of micro-simulation models in conjunction with surrogate safety measures has been shown to complement traditional safety analysis. However, for the adopted simulation model to achieve high fidelity, it is important to calibrate and validate it before use. This paper proposes a numerical optimization approach to calibrate a traffic simulation model for rear-end traffic conflict risk analysis on highways. The proposed calibration approach is developed based on the stochastic gradient approximation algorithm to find optimal parameters of stochastic traffic simulation models. The calibration methodology accounts for multiple calibration criteria and is implemented on a selected traffic simulation platform to test its performance. Simulated operational measurements and traffic conflict risk in terms of surrogate safety measures are quantified and compared with observations derived from real-world vehicle trajectory data from the Next Generation Simulation (NGSIM) program. The calibrated traffic model has been validated by using independent vehicle trajectory data saved as a hold-out sample. The results show that the fine-tuning of parameters using the proposed calibration approach can significantly improve the performance of the simulation model to describe actual traffic conflict risk and operational performances.

**Keywords:** simulation model, safety, traffic conflict risk, calibration, surrogate safety measures

## INTRODUCTION

Micro-simulation models have been developed primarily for analysis, evaluation, and optimization of traffic operations. The concept of using micro-simulation models for traffic safety evaluation is still a challenging and relatively unexplored topic. Nevertheless, a number of researchers have recently proposed the potential of using a micro-simulation approach for safety evaluation. Since the initial recognition by Darzentas et al. (1980), this approach has gained increasing attention. Several studies (e.g., Young et al., 1989; Algiers et al., 1997; Archer & Kosonen, 2000) have revealed that micro-simulation models can capture driver behavior and individual interactions of vehicles to be studied, and many parameters used in the models have some implications on the safety issues of vehicle-to-vehicle interactions.

Using micro-simulation models for safety assessments can be thought of as employing traditional traffic conflict techniques (TCT) in conjunction with the same simulation models used for operational performance analysis of traffic. Traditional TCTs are extended to use simulated vehicle trajectory information to automate conflict analysis based on traffic simulation data. The overall level of safety under certain operational conditions can then be determined by using surrogate safety measures. The simulation-based approach using surrogate safety measures provides a proactive safety evaluation technique to diagnose traffic safety problems to select and quantify appropriate remedial measures. It requires less human involvement to extract conflict information and therefore avoids one of the main sources of error encountered when using traditional TCT. The simulated surrogate safety measures and driver attributes have potential to be associated with the likelihood of a conflict or collision (Sayed et al., 1994; Muchuruza, 2006). Among many simulation-based safety studies conducted in recent years are intersection safety evaluation studies by Sayed and Zein (1999) and Saccomanno et al. (2008), freeway safety analysis by Fazio et al. (1993), safety of truck-lane restrictions by Liu and Garber (2007), and safety of ramp merging sections by Barceló et al. (2003).

Although great effort has been made toward using microscopic simulation models for safety assessment, questions related to the use of this approach still exist. One major concern is the calibration of the micro-simulation models. In a microscopic simulation model, driver behaviors are captured via sub-models representing car-following, gap-acceptance, and lane-changing behavior of each driver in the simulation. These models are in turn dependent on input parameters deemed to represent relevant aspects of driver behavior. Thus one of the major steps in using simulation models is to ensure that input parameters are determined based on observational data, so that the models replicate safety performance of a given facility under a given operational condition, and can be verified from field observations (as described by Cunto & Saccomanno, 2008). In fact, it has recently been recognized that input parameters can have a direct effect on the resulting safety measures from simulations (Klunder et al., 2006). These realistic input parameters can be determined through a robust calibration process that incorporates safety-related aspects of traffic through the use of relevant observed safety data, such as observed conflicts and accidents. Unfortunately, there is limited simulation calibration experience with an emphasis on safety evaluation. In fact, the calibration process can even be more demanding, expensive, and time-consuming, when the objective of safety evaluation is combined with the traditional operational performance-related objectives.

The main objective of this paper is to develop a novel procedure for calibrating traffic model input parameters based on both operational and safety performance measures, to analyze highway rear-end conflict risks. Multicriteria optimization and stochastic approximation approach are used to estimate optimal parameters, and NGSIM (FHWA, 2005) vehicle trajectory data are used as major sources of observed data. The calibration approach is further validated by comparing simulated results with the actual observations using additional trajectory data.

## METHODOLOGY

### Simulation Calibration Problem

Calibration of a micro-simulation model can be defined as finding a set of parameters to optimize the difference between simulation output and corresponding observations. The simulation output can be a set of measurements such as flow, speed, and travel time, etc. The difference between each simulated measurement and its observation can be used to quantify a performance criterion. Essentially, calibration is a multi-objective optimization problem, though a single performance criterion is frequently used as calibration target in practice. Mathematically, the calibration problem can be described as finding the parameter set  $\theta^* = [\theta_1^*, \theta_2^*, \dots, \theta_n^*]^T$  that can optimize the objective function set  $z(\theta)$ :

$$z(\theta) = [z_1(\theta), z_2(\theta), \dots, z_m(\theta)]^T \quad (1)$$

Here, “optimize” means that the optimal parameter set  $\theta^*$  can yield the solution of each performance criterion  $z_i(\theta)$  at a defined acceptable level. For example, one criterion is to find  $\theta^*$  to minimize the difference between simulated and observed traffic counts, and another is to reduce the difference of simulated and observed travel times to less than 5 percent. Specifically, the objective function defined as  $z(\theta)$  in this paper consists of both safety and operational performance criteria. It should be noted that these performance criteria may be in conflict with each other. In other words, it is rarely the case that a single set  $\theta^*$  can simultaneously optimize all the performance criteria. Therefore, it is necessary to search for the “trade-offs” among the safety and operational performance criteria. In this paper, we use the idea of multicriteria optimization approach to achieve the goal. Specifically, the point estimate weighted-sums method (Steuer, 1986) is used to simplify the typical calibration problem in terms of minimizing differences between simulated output and observations.

The point estimate weighted-sums method can be described as follows:

$$\begin{aligned} \min z(\theta) &= \omega_1 z_1(\theta) + \omega_2 z_2(\theta) + \dots + \omega_m z_m(\theta) \\ \text{s.t. } \theta &\in \Theta \end{aligned} \quad (2)$$

where  $\omega_i$  is a user defined non-negative scalar weight of the  $i^{th}$  performance criterion  $z_i(\theta)$ .  $z(\theta)$  becomes the aggregated objective function.  $\Theta$  is the possible domain of parameters to be calibrated. If there is no interest to calibrate  $z_k(\theta)$ , one can set  $\omega_k = 0$ . Otherwise, a positive scalar weight has to be assigned to each performance criterion. Without loss of generality, we can assume that  $\omega_1 + \omega_2 + \dots + \omega_m = 1$ . This method is straightforward to aggregate safety and operational performance measures. By using the scalar weights, the multi-objective of the optimization problem is converted into a single criterion problem that is easier to analyze.

Moreover, each  $\omega_i$  can be modified to reflect the priority corresponding to this individual criterion by decision makers. In this method, the optimal solution of the target function is controlled by the selected weighting vector. Depending on the information available, some subjective and objective approaches have been proposed to determine the weights (e.g., Saaty, 1980; Hwang & Yoon, 1981; Gass, 1987). The selecting of optimal weighting vector, however, is usually difficult given the different units, scales and numbers of observations that various objectives may have (Steuer, 1986). If there is no such difficulty, a solution obtained with equal weights may offer least objective conflict among all criteria.

### Parameter Estimation Algorithm

To efficiently solve the problem defined in equation (2), it is necessary to resort to numerical optimization methods. In this paper, we use the simultaneous perturbation stochastic approximation (SPSA) algorithm because of the highly stochastic nature of the underlying microscopic traffic simulation models. SPSA algorithm was first introduced by Spall (1987, 1988 & 1992) and expanded in subsequent work (e.g. Fu and Hill, 1997; Sadegh, 1997; Bhatnagar and Borker, 2003). The algorithm, developed based on the iterative form of generic stochastic approximation (SA), is shown in equation (3):

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \mathbf{g}_k(\hat{\theta}_k) \quad (3)$$

where  $\mathbf{g}_k(\hat{\theta}_k)$  is the gradient of objective function when parameter vector  $\theta = \hat{\theta}_k$ , and  $a_k$  is a positive gain sequence of step sizes.

SA approach attempts to mimic the gradient search method used in deterministic optimization. Based on equation (3), the recursive procedure must obtain the gradient of the objective function in order to update the parameters in the  $k^{\text{th}}$  iteration. The Robbins-Monro algorithm (Robbins and Monro, 1951) can be used to perform parameter updates when the gradient of the objective function is available. However, our simulation model does not allow the computation of  $\mathbf{g}(\theta)$  because there is no clear mathematical expression of an objective function as a response of simulation parameters. Thus, it is necessary to approximate  $\mathbf{g}(\theta)$ . When a finite-difference (FD) method (Dennis and Schnabel, 1989) is used to approximate the gradient, the well-known form of SA called Kiefer-Wolfowitz algorithm (Kiefer and Wolfowitz, 1952) is obtained. The general form of the algorithm takes the following iterative form:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{\mathbf{g}}_k(\hat{\theta}_k) \quad (4)$$

where  $a_k$  is a positive gain sequence of step sizes, and  $\hat{\mathbf{g}}_k(\hat{\theta}_k)$  is the approximation of  $\mathbf{g}(\theta)$  at each iteration.

According to Spall (1998), SPSA uses the following formula to obtain the approximation of  $\hat{\mathbf{g}}_k(\hat{\theta}_k)$ :

$$\hat{\mathbf{g}}_k(\hat{\boldsymbol{\theta}}_k) = \begin{bmatrix} \frac{z(\hat{\boldsymbol{\theta}}_k + c_k \Delta_k) - z(\hat{\boldsymbol{\theta}}_k - c_k \Delta_k)}{2c_k \Delta_{k1}} \\ \dots \\ \frac{z(\hat{\boldsymbol{\theta}}_k + c_k \Delta_k) - z(\hat{\boldsymbol{\theta}}_k - c_k \Delta_k)}{2c_k \Delta_{kp}} \end{bmatrix} = \frac{z(\hat{\boldsymbol{\theta}}_k + c_k \Delta_k) - z(\hat{\boldsymbol{\theta}}_k - c_k \Delta_k)}{2c_k} \begin{bmatrix} \Delta_{k1}^{-1} \\ \Delta_{k2}^{-1} \\ \dots \\ \Delta_{kp}^{-1} \end{bmatrix} \quad (5)$$

where  $c_k$  is gain sequence at  $k^{\text{th}}$  step.  $\Delta_k$  is perturbation vector at  $k^{\text{th}}$  step.  $p$  is the number of parameters.

At iteration step  $k$  we take a random perturbation vector:

$$\Delta_k = [\Delta_{k1}, \Delta_{k2}, \dots, \Delta_{kp}]^T \quad (6)$$

Common gain sequence of  $a_k$  and  $c_k$  are positive with the form of power functions shown in equations (7) and (8):

$$a_k = \frac{a}{(1 + A + k)^\alpha} \quad (7)$$

$$c_k = \frac{c}{(1 + k)^\gamma} \quad (8)$$

where  $\Delta_{ki}$  denotes a sequence of independent identically distributed, symmetrically distributed, bounded random variables satisfying certain conditions (Spall, 1992). A standard perturbation can be a sequence  $\Delta_{ki}$  of Bernoulli  $\pm 1$  distribution with probability  $P(\Delta_{ki} = +1) = 1/2$  and  $P(\Delta_{ki} = -1) = 1/2$ .  $a, c, \alpha, \gamma$ , and  $A$  are the coefficients.  $a$  and  $c$  control the noise setting.  $A$  is a constant introduced to stabilize the optimization process. The exponents  $\alpha$  and  $\gamma$  control the speed of the convergence, and Li et al. (2006) presented the typical constraints for  $\alpha$  and  $\gamma$ . For more practical use, these coefficients can be determined based on some guidelines provided by Spall (1998). For instance,  $\alpha = 0.602$  and  $\gamma = 0.101$  are shown to yield good results in several cases.

As the numerator in equation (5) is the same for each component of  $\hat{\mathbf{g}}_k(\hat{\boldsymbol{\theta}}_k)$ , the number of function evaluations needed to estimate the gradient in SPSA is only two. This property of SPSA provides the potential for a large improvement of the overall efficiency of the optimization analysis. The convergence of the gradient approximation has been proved in many cases (e.g., Spall, 1992, Kushner & Yin, 1997).

Step-by-step implementation of the proposed calibration algorithm which makes use of the above SPSA technique can be summarized as follows:

### Step 1: Initialization and Parameter Selection

Set the iteration index  $k = 0$ . Pick initial guess  $\hat{\boldsymbol{\theta}}_0$  for equation (4). In our simulation model, we can use the default configuration values as an alternative. Select the nonnegative algorithmic coefficients  $a, c, \alpha, \gamma$ , and  $A$ .

**Step 2: Generation of Simultaneous Perturbation Vector**

Generate a  $p$ -dimensional random perturbation vector  $\Delta_k$ , where each of the  $p$  components of  $\Delta_k$  is independently generated from a Bernoulli  $\pm 1$  distribution with probability of 0.5 for each  $\pm 1$  outcome.

**Step 3: Loss Function Evaluations**

Run the traffic simulation model with perturbed parameters  $\hat{\theta}_k \pm c_k \Delta_k$  based on  $c_k$  and  $\Delta_k$  from Step 1 and Step 2. Obtain two measurements of the loss function:  $z(\hat{\theta}_k + c_k \Delta_k)$  and  $z(\hat{\theta}_k - c_k \Delta_k)$ .

**Step 4: Gradient Approximations**

Compute the simultaneous perturbation approximation to the unknown gradient  $\hat{g}_k(\hat{\theta}_k)$  according to equation (5).

**Step 5: Update Parameter Estimation  $\hat{\theta}_k$** 

Use the recursive equation (4) to update  $\hat{\theta}_k$  to a new value  $\hat{\theta}_{k+1}$ . Check for constraint violation and modify the updated  $\theta$  if necessary.

**Step 6: Check Convergence**

Check whether the maximum number of iterations has been reached or the predefined convergence criteria are satisfied. If not, return to Step 2 with iteration  $k+1$ . Otherwise, terminate the algorithm and report optimal values of parameters  $\theta$ .

To make the SPSA algorithm more suitable for the analysis in this study, several enhancements have been made:

- (1) original simulation parameters have been normalized to 0 to 1.0 for perturbation in step 3 and inverse scaling of the perturbed parameters have been performed when running simulation
- (2) multiple simulation runs with different random seeds have been conducted to obtain the average gradient in step 4
- (3) multiple initial parameters for the simulation were tested and compared to obtain better parameters.

**CASE STUDY****Data Collection**

A field vehicle tracking dataset namely “I-101 Dataset” generated by the NGSIM program was obtained to demonstrate the implementation of the proposed methodology. The dataset is “*specifically collected to improve the quality and performance of simulation tools, promote the use of simulation for research and applications, and achieve wider acceptance of validated simulation results (FHWA, 2005).*” The data were collected at a segment of southbound direction of U.S. highway 101 in Universal City neighborhood of Los Angeles, California. The schematic illustration of the location is shown in Figure 1. The length of the segment used for data

collection was approximately 2100 feet and the length of the auxiliary lane for on-ramp vehicle merging and vehicle diverging is about 698 ft. About 6,000 vehicle trajectories were collected based on video data with a 0.1-second time increment. This amount of detailed trajectory data is unique compared with previous traffic studies and provides a better basis to objectively investigate real-world traffic conflicts. The dataset was also separated into three 15-minute periods representing transitional and congested flow conditions on the morning of June 15, 2005. Data between 08:05 AM and 08:20 AM are retrieved for model calibration and data between 08:20 AM and 08:35 AM are used for validation.

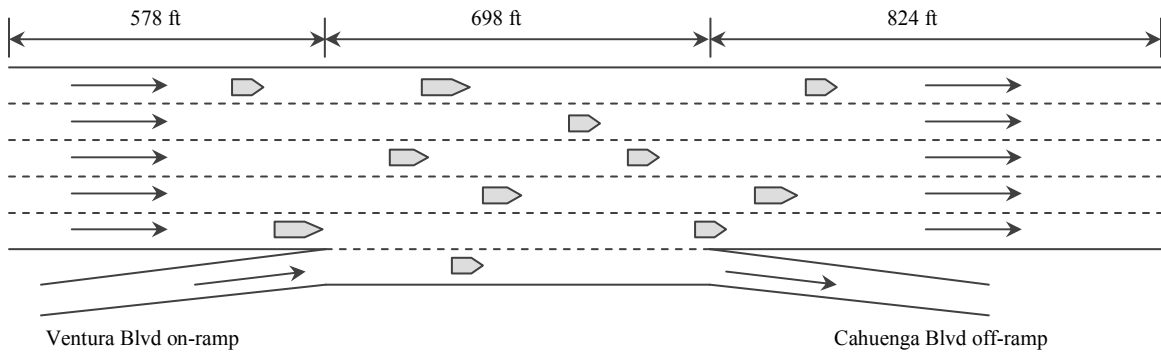


Figure 1 Study area schematic of I-101

### Measuring Traffic Conflict Risk

A number of surrogate safety measures have been developed to identify potential conflicts through the use of simulation models. Generally, these measures fall into four categories: time-based, distance-based, deceleration-based, and other composite measures. Some of the most frequently used time-based measures are the time-to-collision (TTC) and post-encroachment time (PET), developed in the 1970s (Hayward, 1972; Allen et al., 1978). TET (time exposed time-to-collision) and TIT (time integrated time-to-collision) extended from TTC are also introduced by Minderhoud and Bovy (2001). Ozbay et al. (2008) introduced a modification of the time-to-collision (MTTC). Deceleration rate to avoid the crash, (DRAC), has been recognized as a typical deceleration-based safety indicator (Gettman and Head, 2003; Archer, 2005). Crash potential, proposed by Saccomanno and Cunto (2006), can also be an indirect use of such a measure. One important example of a distance-based indicator could be the possibility index for collision with urgent deceleration (PICUD) proposed by Uno et al. (2003). Besides the aforementioned measures, several other studies have also proposed specific measures such as unsafe density (UD) and J-value in support of safety evaluation (Barceló et al., 2003; Pham et al., 2007). There is no unique measure to describe all types of traffic conflicts. For instance, PET is more appropriate for intersecting conflicts (Songchitruksa & Tarko, 2006) while TTC is for measuring rear-end conflicts (Gettman and Head, 2003).

In this paper, we use the concept of conflict probability (CP) introduced by Yang and Ozbay (2011) as the surrogate measure to describe the potential rear-end conflict risk. It is derived from the modification of the time-to-collision (MTTC) index introduced by Ozbay et al. (2008). Considering the fact that the shorter MTTC is the higher probability of conflict is, CP adopts an exponential decay function as an alternative of defining single threshold value of MTTC to identify potential risk of the conflict. The function of the potential conflict probability (CP)

associated with a subject vehicle is shown in equation (9). CP is a continuous monotone decreasing function of MTTC such that as  $MTTC \in [0, +\infty)$ ,  $CP \in [1, 0)$ . When MTTC is 0, two consecutive vehicles definitely conflict with each other. When MTTC is relatively large, conflict probability will be small. The same MTTCs may not indicate the same chance of conflict under different traffic conditions. So a parameter  $\lambda$  is used for adjusting the impact of MTTC at different cases, such as a freeway versus a local road.

$$CP = \Pr(\text{Conflict} | MTTC) = \text{Exp}\left(\frac{-MTTC}{\lambda}\right) \quad (9)$$

Depending on the objective, CP of subjective vehicles can be aggregated by time and space to describe the conflict risk of the target facility. In this study, the conflict risk  $C_m$  is represented by aggregating the simulated CPs over the  $m^{\text{th}}$  section (100-ft) and simulation time period (15-minute).

### Simulation Modeling

Previous studies (e.g., Gettman & Head, 2003; Archer, 2005) have given some insights into the strengths and weaknesses of various simulation software packages that can be used to support safety analysis. However, there are still no definitive conclusions as to which simulation package is a better tool to conduct safety analysis. Paramics simulation tool is selected as our test platform to model the study section mainly because of its relatively superior customization potential. The simulated trajectory information obtained through the Application Programming Interface (API) facility of Paramics provides vehicle's position, speed and acceleration for a user-defined small time resolution. The simulated data provide sufficient information needed to numerically compute the surrogate safety measure and other operational measures. Major global and local parameters of the simulation model have been summarized in Table 1. Though each parameter may affect operational and/or safety performances, calibration of the last three discrete parameters are beyond the scope of this study. The search space defined by factors ranging A to H consists of an eight-dimensional hyperplane. It is difficult to enumerate all possible parameter sets on the hyperplane and to run the simulation model with all the sets. The SPSA based approach is then used to search acceptable parameter sets in a faster manner.

Table 1 Major Parameters in Paramics

Factor	Parameters	Default Value	Feasible Range	Low Level (-1)	High Level (+1)
A	Mean Target Headway (s)	1.0	0.5 to 3.0	0.5	3.0
B	Mean Driver Reaction Time (s)	1.0	0.5 to 2.0	0.5	2.0
C	Minimum Gap (ft)	5.00	1.0 to 9.0	1.0	9.0
D	Queue Gap Distance (ft)	32.81	5.0 to 40.0	5	40
E	Queue Speed (mph)	4.47	1.0 to 8.0	1.0	8.0
F	Link Headway Factor	1.0	0.5 to 2.0	0.5	2.0
G	Link Reaction Factor	1.0	0.5 to 2.0	0.5	2.0
H	Signpost (ft)	696.2	1.0 to 1500.0	1.0	1500.0
I	Driver Aggressiveness	4	1 to 8	1	8
J	Driver Awareness	4	1 to 8	1	8
K	Speed Memory	5	1 to 20	1	20



To calibrate the simulation model, the following objective functions in terms of root mean square percentage error (RMSPE) namely,  $z_i(\theta)$ , have been defined:

$$\min z_1(\theta) = \sqrt{\frac{1}{M} \sum_{m=1}^M \left[ \frac{C_m^s - C_m^o}{C_m^o} \right]^2} \quad (10)$$

$$\min z_2(\theta) = \sqrt{\frac{1}{N} \sum_{n=1}^N \left[ \frac{L_n^s - L_n^o}{L_n^o} \right]^2} \quad (11)$$

$$\min z_3(\theta) = \sqrt{\frac{1}{I} \sum_{i=1}^I \left[ \frac{F_i^s - F_i^o}{F_i^o} \right]^2} \quad (12)$$

$$\min z_4(\theta) = \sqrt{\frac{1}{J} \sum_{j=1}^J \left[ \frac{V_j^s - V_j^o}{V_j^o} \right]^2} \quad (13)$$

Based on the point estimate weighted-sums method we defined the aggregated objective function  $z(\theta)$  as follows:

$$\begin{aligned} \min z(\theta) &= \omega_1 z_1(\theta) + \omega_2 z_2(\theta) + \omega_3 z_3(\theta) + \omega_4 z_4(\theta) \quad (14) \\ &= \omega_1 \sqrt{\frac{1}{M} \sum_{m=1}^M \left[ \frac{C_m^s - C_m^o}{C_m^o} \right]^2} + \omega_2 \sqrt{\frac{1}{N} \sum_{n=1}^N \left[ \frac{L_n^s - L_n^o}{L_n^o} \right]^2} + \omega_3 \sqrt{\frac{1}{I} \sum_{i=1}^I \left[ \frac{F_i^s - F_i^o}{F_i^o} \right]^2} + \omega_4 \sqrt{\frac{1}{J} \sum_{j=1}^J \left[ \frac{V_j^s - V_j^o}{V_j^o} \right]^2} \end{aligned}$$

where  $z(\theta)$  quantify the overall error of the simulation;  $z_1(\theta)$ ,  $z_2(\theta)$ ,  $z_3(\theta)$  and  $z_4(\theta)$  are criteria that quantify the performance of traffic conflict, lane change, traffic count and speed, respectively.  $C_m^o$ ,  $L_n^o$ ,  $F_i^o$  and  $V_j^o$  are an observation of traffic conflict risk, lane change, traffic count, and speed value, respectively.  $C_m^s$ ,  $L_n^s$ ,  $F_i^s$  and  $V_j^s$  are the corresponding simulated values.  $M$ ,  $N$ ,  $I$  and  $J$  are total numbers of observations. We assume that equal weights are assigned to individual performance criterion. Calibrations based on each single objective function and the aggregated objective function have been implemented.

## SIMULATION RESULTS & DISCUSSION

### Calibration Results

NGSIM provides summarized lane change and speed information for each sub-section (100-ft) of the entire segment (2100-ft) in their summary reports (FHWA, 2005). These values are used as baseline data when computing  $z_2(\theta)$  and  $z_4(\theta)$ , respectively. The simulated throughput and the reported 15-minute throughput are compared to obtain  $z_3(\theta)$ . To obtain  $z_1(\theta)$ , traffic conflict probability is calculated using the original trajectory data and the simulated trajectories. Only one percent of the car-following scenarios are sampled. This is equivalent to screen the status of an individual vehicle every 10 seconds. The conflict probability of each sample is then calculated using equation (9) and aggregated by sub-section.  $z(\theta)$  is computed based on equation (14).

To confirm the need for adequate model calibration, initial runs with default input parameters and 13 sets of guessed input parameters were also conducted. The simulation results based on these input parameter sets are presented in Table 2. No matter which performance measure is used, simulation results are not stable. For instance,  $z_1(\theta)$  ranges from 0.176 to 0.706 and  $z_2(\theta)$  varies from 0.261 to 1.820. Large variance of these performance measurements suggest that neither default values nor guessed parameter sets can definitely yield optimal results. It is thus necessary to calibrate the parameters.

Table 2 Simulation Results based on Different Parameter Guesses

Initial Inputs	Value of Parameters ( $\theta$ )								Simulation Results				
	A	B	C	D	E	F	G	H	$z_1(\theta)$	$z_2(\theta)$	$z_3(\theta)$	$z_4(\theta)$	$z(\theta)$
Default	1.00	1.00	5.00	32.81	4.47	1.00	1.00	696.20	0.353	0.261	0.054	0.140	0.202
Guess 1	0.90	1.60	6.00	10.00	2.80	1.50	0.70	700.00	0.364	0.309	0.184	0.230	0.272
Guess 2	1.45	0.75	7.30	27.00	7.50	1.20	1.10	800.00	0.176	0.416	0.179	0.144	0.229
Guess 3	0.90	0.90	6.20	15.00	7.00	1.70	1.70	1450.00	0.529	0.508	0.239	0.123	0.350
Guess 4	0.70	0.70	4.50	22.81	4.90	1.50	1.50	570.00	0.292	0.300	0.053	0.102	0.187
Guess 5	1.80	1.50	6.60	29.00	6.40	1.20	1.20	1350.00	0.571	1.029	0.397	0.168	0.541
Guess 6	1.68	1.45	3.30	24.00	6.50	1.20	1.20	570.00	0.595	1.111	0.349	0.154	0.552
Guess 7	1.60	1.30	7.70	40.00	5.00	1.00	1.00	800.00	0.302	0.264	0.256	0.212	0.259
Guess 8	1.70	1.30	4.20	36.00	5.20	1.35	1.85	1000.00	0.706	1.820	0.456	0.143	0.781
Guess 9	1.25	1.05	4.30	28.00	3.15	1.30	0.50	630.00	0.194	0.440	0.099	0.153	0.222
Guess 10	1.75	1.35	5.20	15.00	4.75	1.40	0.70	300.00	0.324	0.360	0.307	0.177	0.292
Guess 11	0.75	1.70	6.00	10.00	2.80	1.50	0.75	720.00	0.592	0.301	0.171	0.225	0.322
Guess 12	1.50	1.50	7.30	32.00	5.90	1.20	1.00	900.00	0.417	0.340	0.314	0.197	0.317
Guess 13	2.40	1.70	6.00	19.00	5.50	1.00	1.00	1200.00	0.555	1.187	0.457	0.283	0.621

The SPSA based calibration approach described in the previous section of this paper is employed to calibrate the input parameters. Initially, the simulation model is calibrated just using a single objective function in terms of the measurements of conflict risk, lane change, throughput or speed (shown in equations 10-13). After calibrating simulation model using a single performance criterion, the model is then calibrated based on the aggregated multicriteria objective function shown in equation (14). Figure 2a illustrates an example of the convergence diagram using the measurements of conflict risk as the calibration based objective function,  $z_1(\theta)$ . Despite using different initial guessed parameter sets, the fitted values of  $z_1(\theta)$  converged after a number of iterations using SPSA. Similarly, Figure 2b (right) shows the convergence diagram of the fitted value of  $z(\theta)$ . The converged value of  $z(\theta)$  is about 0.15.

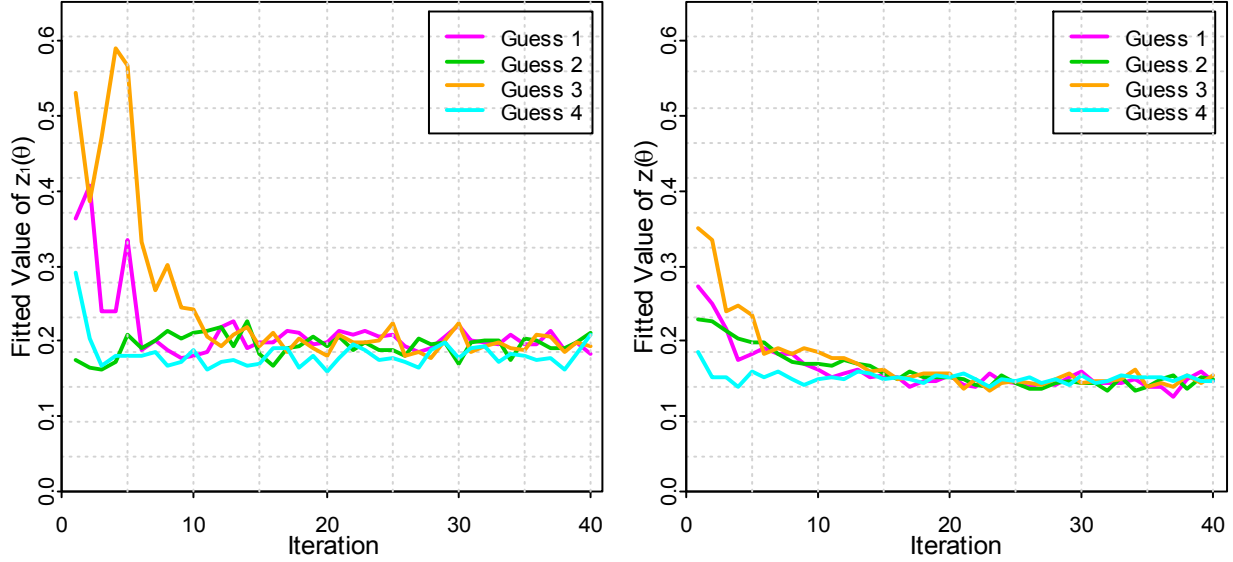


Figure 2 Calibration convergence diagram: a) single criterion (left), b) multicriteria (right)

Table 3 presents an example of the final calibrated results when running simulation with the initial input parameters of guess vector 4 (listed in Table 2). Similarly, Table 4 summarizes the results of guess vector 5. The parameter set is either calibrated by minimizing a single objective function,  $z_i(\theta)$  ( $i=1, 2, 3,$  and  $4$ ) or the multicriteria objective function,  $z(\theta)$ . When an objective function is minimized, the corresponding measurements of other four performance functions are also obtained. For instance, the first row in Table 4 shows that  $z_1(\theta)$  is the objective function and its minimized value is 0.137. In the meanwhile, the measured  $z_2(\theta)$ ,  $z_3(\theta)$ ,  $z_4(\theta)$ , and  $z(\theta)$  are calculated as 0.403, 0.117, 0.143, and 0.200, respectively, for this specific calibration scenario.

Table 3 Calibration Results Using Different Objective Functions (Initial Inputs: Guess vector 4)

Objective Function	Calibrated Value of Parameters ( $\theta$ )								Simulation Results				
	A	B	C	D	E	F	G	H	$z_1(\theta)$	$z_2(\theta)$	$z_3(\theta)$	$z_4(\theta)$	$z(\theta)$
$\min z_1(\theta)$	0.50	0.50	6.72	17.47	4.08	1.40	1.06	522.73	0.210	0.273	0.051	0.079	0.153
$\min z_2(\theta)$	0.72	0.52	5.83	15.10	5.26	0.69	1.27	934.07	0.216	0.269	0.051	0.088	0.156
$\min z_3(\theta)$	0.50	0.58	4.33	20.37	5.04	1.35	1.32	551.50	0.177	0.322	0.051	0.077	0.157
$\min z_4(\theta)$	0.50	0.53	4.27	23.85	5.37	1.40	1.47	599.39	0.204	0.288	0.052	0.077	0.155
$\min z(\theta)$	0.50	0.50	3.38	18.36	3.86	1.19	1.22	769.03	0.174	0.290	0.051	0.076	0.148

Table 4 Calibration Results Using Different Objective Functions (Initial Inputs: Guess vector 5)

Objective Function	Calibrated Value of Parameters ( $\theta$ )								Simulation Results				
	A	B	C	D	E	F	G	H	$z_1(\theta)$	$z_2(\theta)$	$z_3(\theta)$	$z_4(\theta)$	$z(\theta)$
$\min z_1(\theta)$	1.28	0.63	8.27	13.81	6.40	1.21	0.75	1452.26	0.137	0.403	0.117	0.143	0.200
$\min z_2(\theta)$	1.96	0.80	3.39	35.97	5.98	1.75	0.57	536.59	0.426	0.330	0.332	0.152	0.310
$\min z_3(\theta)$	0.58	0.92	7.15	27.13	6.37	0.87	1.03	1036.37	0.196	0.299	0.051	0.087	0.158
$\min z_4(\theta)$	1.51	1.13	7.59	31.94	5.82	1.54	1.40	1239.80	0.520	0.331	0.376	0.125	0.338
$\min z(\theta)$	0.50	0.99	8.50	35.80	5.98	1.61	0.71	1137.27	0.191	0.271	0.052	0.091	0.151

The difference between the final calibrated parameters across the tables suggests that there are multiple solutions for simulation models and SPSA is still a local optimization algorithm that can find one of the many possible solutions. To obtain optimal results, use of multiple initial input parameter sets is recommended. The calibrated results in both tables also suggest that minimization of safety performance function cannot guarantee the minimization of the operational performance functions, and vice versa. For instance, when safety criterion  $z_1(\theta)$  is the selected objective function, the corresponding value of lane-change criterion  $z_2(\theta)$  is 0.403 in Table 4. However, lower values of 0.330 can be obtained if  $z_2(\theta)$  was selected as the objective function. By comparing two tables we found that when  $z(\theta)$  is the objective function, the measurement of each performance function is more stable than that of using a single objective function  $z_i(\theta)$ . The minimized values of  $z(\theta)$  are about 0.15 in both tables. The corresponding  $z_1(\theta)$  is about 0.19,  $z_2(\theta)$  is about 0.29,  $z_3(\theta)$  is about 0.05, and  $z_4(\theta)$  is about 0.09. Though not all the criteria are simultaneously minimized,  $z(\theta)$  avoids the cases of minimizing a single criterion by deteriorating the performance of other criteria. Therefore,  $z(\theta)$  is preferred because when minimizing  $z(\theta)$  all the four performance functions can be equally considered.

## Validation

To evaluate the performance of the calibration model, optimized input parameters are tested using NGSIM trajectory data collected between 08:20 AM and 08:35 AM. The simulated conflict risk, lane changes, speed and throughput are then compared with the computed results using the actual trajectory data. Figure 3 illustrates validation results when using calibrated input parameters. Simulated results can accurately capture observed conflict risk, lane-change, and speed along the 2100-ft segment. Both simulated results using the calibrated parameter set and actual results show that: (a) conflict risks are higher for the upstream of the weaving section; (b) the majority of lane changes occurred at the weaving section; and (c) speed when approaching the weaving section is lower, whereas the speed of the downstream section is higher. Table 5 compares simulation results based on the calibrated parameter set and its original guessed set. RMPSE of traffic conflict risk is reduced from 0.566 to 0.123. Similarly, RMSPE of other measures are also greatly reduced. Observed traffic throughput is 1915 and simulated throughput using calibrated parameter set is 1847. These findings suggest that the calibrated model shows generally good performance in comparison with actual observations.

Table 5 Validation Results Using Calibrated Parameter Set

Parameter Set	Parameters ( $\theta$ )								RMSPE				
	A	B	C	D	E	F	G	H	$z_1(\theta)$	$z_2(\theta)$	$z_3(\theta)$	$z_4(\theta)$	$z(\theta)$
Guess	0.90	1.60	6.00	10.00	2.80	1.50	0.70	700.00	0.566	4.165	0.350	0.094	1.294
Calibrated	0.50	0.96	5.80	14.56	2.18	1.49	0.51	730.67	0.123	0.582	0.035	0.054	0.199
	Difference (%)								78.27	86.03	90.00	42.55	84.62

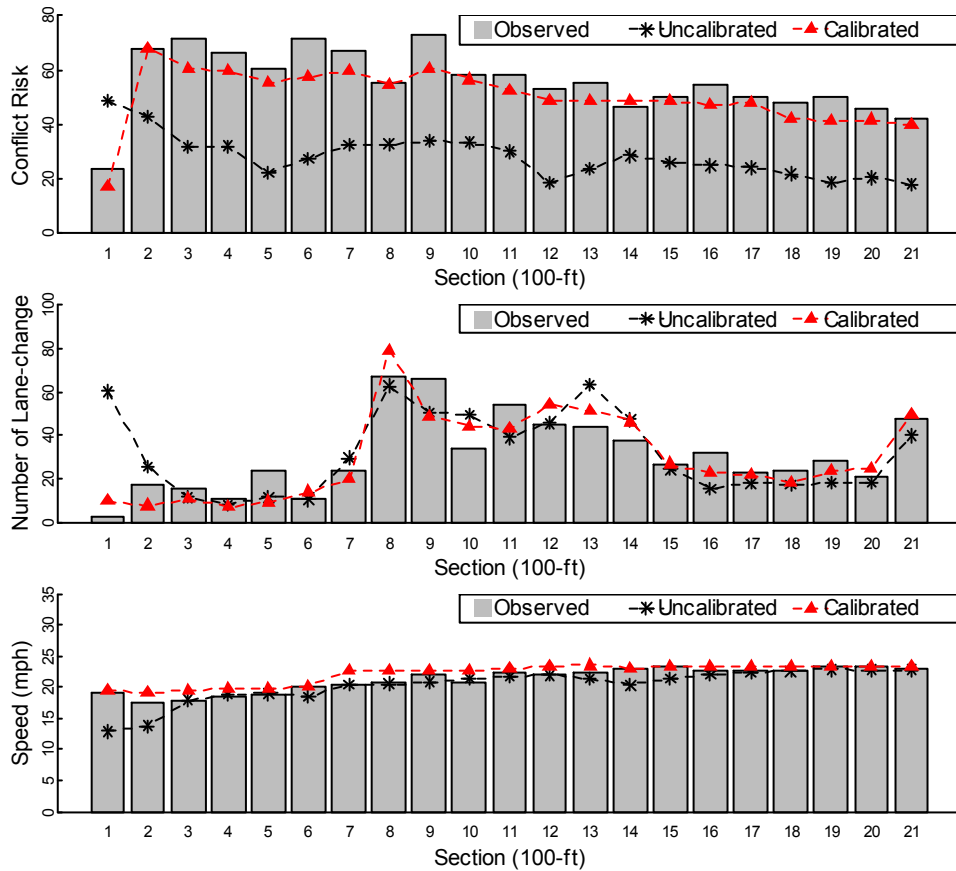


Figure 3 Simulation results using uncalibrated parameters and calibrated parameters versus observations: conflict risk (top), lane-change (middle), speed (bottom)

## CONCLUSIONS

This study develops a numerical approach to calibrate stochastic micro-simulation models for traffic conflict analysis. It has been found that neither default parameters in the simulation model nor randomly guessed parameters can guarantee the accuracy of the simulation model. Moreover, calibration of simulation models using a single criterion may cause deterioration of other important criteria. When simulating traffic conflict risk, the accuracy of safety performance is of great interest. However, calibration solely based on safety criteria can be in conflict with other operational performance criteria such as speed and traffic counts. Instead of solely calibrating to optimize the models estimates in terms of safety performance and ignoring operational criteria, this study adopted the concept of multicriteria optimization by simultaneously considering all of these criteria together. The weighted-sums method is then used to simplify the calibration problem by developing an aggregated objective function. Since there are many input parameters that need to be calibrated, it is impossible to enumerate all possible combinations and run the simulation model for all these combinations. To efficiently find the optimal parameters of the highly stochastic simulation model, SPSA approach is used to perturbate and update all parameters in the searching process. The approach has been shown to be able to find the acceptable parameter set in a relatively fast manner. The proposed SPSA-based multicriteria calibration approach is implemented using Paramics traffic simulation platform for a study network, for which vehicle trajectory data are available through NGSIM program (FHWA,

2005). In the case study, this stochastic and gradient-based calibration approach is shown to be able to identify input parameters that make the aggregate objective function - quickly converge to a stable, almost optimal, value. The consistency of the calibrated parameters has been further validated by using additional vehicle trajectory data not used for calibration. The results show that the fine-tuning of parameters can greatly improve the performance of simulation models to describe traffic conflict risk, as well as the operational measures quantified using the field data.

## REFERENCES

- Algers, S., Bernauer, E., Boero, M., Breheret, L., Di Taranto, C., Dougherty, M., Fox, K., and Gabard, F. F. (1997). Review of Micro-Simulation Models. SMARTTEST Program Report, Institute for Transport Studies, University of Leeds, Leeds, U.K., 1997.
- Allen, B. L., Shin, B. T., and Cooper, P. J. (1978). Analysis of Traffic Conflicts and Collisions. Transportation Research Record: Journal of the Transportation Research Board, No.667, Transportation Research Board of the National Academies, Washington, D.C., pp. 67-74.
- Archer, J. (2005). Indicators for Traffic Safety Assessment and Prediction and Their Application in Micro-Simulation Modelling: A Study of Urban and Suburban Intersections. Ph.D. Thesis, KTH Royal Institute of Technology, Stockholm.
- Archer, J., and Kosonen, I. (2000). The Potential of Micro-Simulation Modeling in Relation to Traffic Safety Assessment. 12th European Simulation Symposium. Hamburg, Germany, September 28-30, 2000.
- Barceló, J., Dumont, A. G., Montero, L., Perarnau, J., and Torday, A. (2003). Safety Indicators for Microsimulation-Based Assessments. In TRB 82nd Annual Meeting Compendium of Papers CD-ROM, Transportation Research Board 82nd Annual Meeting Washington, D.C., January 2003.
- Bhatnagar, S., and Borkar, V. S. (2003). Multiscale Chaotic SPSA and Smoothed Functional Algorithms for Simulation Optimization. Simulation, Vol. 79, No. 10, pp. 568-580.
- Cunto, F., and Saccomanno, F. F. (2008). Calibration and Validation of Simulated Vehicle Safety Performance at Signalized Intersections. Accident Analysis and Prevention, Vol.40, Issue 3, pp.1171-1179.
- Darzentas, J., Cooper, D. F., Storr, P. A., and McDowell, M. R. C. (1980). Simulation of Road Traffic Conflicts at T-junctions. Simulation, pp. 155-164.
- Dennis, J. E., and Schnabel, D. J. (1989). A view of Unconstrained Optimization. Optimization (Handbooks in Operation Research and Management Science, Vol. 1, G. L. Nemhauser et al., eds.), Elsevier, New York, pp. 1-72.
- Fazio, J., Holden, J., and Roupail, N.M. (1993). Use of Freeway Conflict Rates as an Alternative to Crash Rates in Weaving Section Safety Analysis. In Transportation Research Record 1401, TRB, National Research Council, Washington, D.C., pp. 61-69.

FHWA (2005). Next Generation Simulation (NGSIM). Dataset available through website: <http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm>.

Fu, M. C., and Hill, S. D. (1997). Optimization of Discrete Event Systems via Simultaneous Perturbation Stochastic Approximation. *IIE Transactions*, Vol. 29, No. 3, pp. 233-243.

Gass, S. I. (1987), The Setting of Weights in Linear Goal-Programming Problems. *Computers and Operations Research*, Vol. 14, Issue 3, pp. 227-229.

Gettman, D., and Head, L. (2003). Surrogate Safety Measures from Traffic Simulation Models, Final Report. Publication No.: FHWA-RD-03-050, Federal Highway Administration, U.S. Department of Transportation.

Hayward, J. C. (1972). Near-miss Determination through Use of a Scale of Danger. *Highway Research Record*, No.384, pp. 24-34.

Hwang, C. L., Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications*. Springer, Berlin.

Kiefer, J., and Wolfowitz, J. (1952). Stochastic Estimation of a Regression Function. *Annals of Mathematical Statistics*, Vol. 23, pp. 462-466.

Klunder, G., Abdoelbasier, A., and Immers, B. (2006). Development of a Model of Driver Behaviour on Intersections. Presented at 13th ITS World Congress and Exhibition, London, October 8-12, 2006.

Kushner, H. J., and Yin, G. G. (1997). *Stochastic Approximation Algorithms and Applications*, Springer-Verlag, New York.

Li, Q., Sato, I., and Murakami, Y. (2006). Simultaneous Perturbation Stochastic Approximation Algorithm for Automated Image Registration Optimization. *Proceedings of IEEE Conference on Geoscience and Remote Sensing Symposium*, pp. 184-187.

Liu, Q., and Garber, N. (2007). Identifying the Impact of Truck –Lane Restriction Strategies on Traffic Flow and Safety Using Simulation. Report No.: UVACTS-14-5-103, Center for Transportation Studies, The University of Virginia.

Minderhoud, M. M., and Bovy, P. H. L. (2001). Extended Time to Collision Measures for Road Traffic Safety Assessment. *Accident Analysis and Prevention*, Vol. 33, Issue 1, pp. 89-97.

Muchuruza, V. (2006). *Simulation and Modeling of Traffic Crashes Using Cell-Based Micro-Simulation*. PhD Dissertation, The Florida State University.

Ozbay, K., Yang, H., Bartin, B., and Mudigonda, S. (2008). Derivation and validation of new simulation-based surrogate safety measure. *Transportation Research Record: Journal of the*

Transportation Research Board No.2083, Transportation Research Board of the National Academies, Washington, D.C., pp. 105-113.

Pham, M. H., Mouzon, O., Chung, E., and Dumont, A. G. (2007). Applicability of Road Safety Indicators to Assess Driving Risks in Swiss Road Conditions. Presented in the 7th Swiss Transport Research Conference, Monte Verità / Ascona, September 12-14, 2007.

Robbins, H., and Monro, S. (1951). A Stochastic Approximation Method. *Annals of Mathematical Statistics*, vol. 22, pp. 400-407.

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill, New York.

Saccomanno, F. F., and Cunto, F. (2006). Evaluation of Safety Countermeasures at Intersections Using Microscopic Simulation. In *TRB 85th Annual Meeting Compendium of Papers CD-ROM*, Transportation Research Board 85th Annual Meeting Washington, D.C., January 2006.

Saccomanno, F. F., Cunto, F., Guido, G., and Vitale, A. (2008). Comparing Safety at Signalized Intersections and Roundabouts Using Simulated Rear-End Conflict. *Transportation Research Record: Journal of the Transportation Research Board*, No.2078, Transportation Research Board of the National Academies, Washington, D.C., pp. 90-95.

Sadegh, P. (1997). Constrained Optimization via Stochastic Approximation with a Simultaneous Perturbation Gradient Approximation, *Automatica*, Vol. 33, Issue 5, pp. 889-892.

Sayed, T., and Zein, S. (1999). Traffic Conflict Standards for Intersections. *Transportation Planning and Technology*, Vol.22, pp. 309-323.

Sayed, T., Brown, G., and Navin, F. (1994). Simulation of Traffic Conflicts at Unsignalized Intersections with TSC-Sim. *Accident Analysis & Prevention*. Vol.26, No.5, pp. 593-607.

Songchitruksa, P., and Tarko, A. P. (2006). Practical Method for Estimating Frequency of Right-Angle Collisions at Traffic Signals. *Transportation Research Record: Journal of the Transportation Research Board* No.1953, Transportation Research Board of the National Academies, Washington, D.C., pp. 89-97.

Spall, J. C. (1987). A Stochastic Approximation Technique for Generating Maximum Likelihood Parameter Estimates, *Proceedings of the American Control Conference*, pp. 1161-1167.

Spall, J. C. (1988). A Stochastic Approximation Algorithm for Large-Dimensional Systems in the Keifer-Wolfowitz Setting. *Proceedings of the IEEE Conference on Decision and Control*, pp. 1544-1548.

Spall, J. C. (1998). Implementation of the Simultaneous Perturbation Algorithm for Stochastic Optimization. *IEEE Trans. Aerospace and Electronic Systems*, Vol. 34, pp. 817-823.

Spall, J. C. (1992). Multivariate Stochastic Approximation Using a Simultaneous Perturbation



Gradient Approximation. IEEE Transactions on Automatic Control, Vol. 37, pp. 332-341.

Steuer, R. E. (1986). Multiple Criteria Optimization: Theory, Computation, and Application. John Wiley & Sons, New York.

Uno, N., Iida, Y., Yasuhara, S., and Masumi, S. (2003). Objective Analysis of Traffic Conflict and Modeling of Vehicular Speed Adjustment at Weaving. Journal of Infrastructure Planning, Vol. 4, No.4, pp. 989-996.

Yang, H., and Ozbay, K. (2011). Estimating Traffic Conflict Risk Associated with Merging Vehicles on a Highway Merge Section. In Transportation Research Record: Journal of Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C. (In Press).

Young, W., Taylor, M. A. P. and Gipps, P. G. (1989). Microcomputers in Traffic Engineering. Research Studies Press, Taunton.