Calibration of microscopic traffic model for simulating safety performance

David D.Q. Duong[1]*, Research Assistant and Ph.D Student
E-mail: dqduong@engmail.uwaterloo.ca

Frank F. Saccomanno[1], Professor, Ph.D, P.Eng
E-mail: saccoman@uwaterloo.ca

Bruce R. Hellinga[1], Associated Professor, Ph.D, P.Eng
E-mail: bhellinga@uwaterloo.ca

[1]Department of Civil and Environmental Engineering
University of Waterloo
200, University Street West
Waterloo, ON, Canada N2L 6P6
Tel: (519) 888-4567 x. 6596
Fax: (519) 888-5221

Text:		4178
Figures:	3 x 250 =	750
Tables:	7 x 250 =	1750
Total words:		6678

*Corresponding author

Compendium of papers of the 89th Annual TRB Conference held Jan. 10 – 14, 2010 in Washington, DC.

Revised November 15, 2009

ABSTRACT

A multi-criteria calibration procedure is proposed for parameter calibration of microscopic traffic simulation platforms. The impetus for this paper is provided by increased usage of microscopic simulation models in transportation safety studies. Before such models can be adopted in safety research it is important that we obtain parameter values that reflect real world traffic conditions. Current state-of-the-art Genetic Algorithm calibration procedures only allow for one measure of performance in parameter calibration. In safety studies, the fitting function used in calibration has been safety performance. Therefore, the underlying traffic-related factors, such as speed, volumes and density have not been directly considered in calibration. Since these models are based on simulating traffic and calibration needs to be based observed traffic attributes. The proposed multi-criteria procedure would allow for the direct calibration of these traffic attributes while at same time producing accurate estimates of safety performance. The multi-criteria procedure is applied to a sample of vehicle tracking data and the results are compared to parameter values suggested by a single-criteria approach and platform defaults. The application of a multi-criteria approach using traffic attributes yielded correspondingly good estimates of safety performance for the NGSIM dataset.

INTRODUCTION

In order to better appreciate the link between time-dependent traffic and corresponding high risk driving behaviour, researchers have turned to microscopic traffic simulation models. These models consider a range of individual driving situations that potentially could impact safety, and hence assist in the development and evaluation of cost-effective countermeasures.

A major challenge inhibiting the application of microscopic traffic simulation in safety studies is to bridge the gap between simulated and real-world driving experience. Researchers who have been critical of using these types of models in safety studies commonly cite three basic arguments [(1), (2)]: 1) traffic simulation is based on fundamental rules of crash avoidance, and as such, it fails to provide a full explanation for high risk driving behavior leading to crashes, 2) the results from simulation model applications are only as good as the accuracy and reliability of their relevant input parameters and the model's ability to explain "real world" traffic conditions, and 3) safety performance is a conceptual yardstick that is only relevant within the context of verifiable crash experience. These arguments present a number of technical challenges that have not been adequately addressed in the current safety literature.

The focus of this paper is to introduce a procedure for calibrating microscopic traffic simulation models, for the purpose of estimating safety performance. Best estimates of input parameters are obtained using a multi-criteria fitting function based on observed real-time traffic attributes. The calibration framework introduced in this paper is applied to a sample of vehicle tracking data obtained from the FHWA NGSIM (3) program as observed for freeway conditions.

The purposes of the application exercise are to:

a) Investigate the balance between separate traffic-based goodness-of-fit or error functions and to the overall goodness-of-fit for the model.
b) Compare parameter values obtained from three different calibration approaches: multi-criteria, single safety performance based criterion and default values.

REVIEW OF PREVIOUS CALIBRATION LITERATURE

As discussed by Hellinga (4), the use of microscopic traffic simulation models (e.g. VISSIM, INTEGRATION, PARAMICS, SIMTRAFFIC) requires the calibration of parameters that govern underlying driving behaviour, such as gap-acceptance, lane-change, and car-following. In calibrating these parameters the analyst must first select suitable measures of performance (MOP) to fit model outputs to "observational" traffic data. The main purpose of the calibration is to obtain parameter values that minimize error between observed and simulated MOP, normally expressed in terms of root mean squared error (RMSE) for volume or speed.

The selection of MOP is dictated by the type of study that is undertaken. For example, if the focus is safety performance, then MOP should provide an indication of unsafe driver behavior or crash potential at a given location (5). Alternatively if the focus of interest is traffic flow, then MOP should be based on traffic flow indicators, such as, temporal progression of vehicles, their average speed or volume [(6); (7)]. The approach adopted in this paper however, is that traffic is an input into safety performance, and as such, the minimization of error in traffic attributes should lead to a minimization of error in safety performance. The two objectives are then essentially linked.

The use of traffic simulation models in safety studies frequently requires the adoption of surrogate safety performance (SP) measures, since crashes cannot be predicted directly from simulation (8). There are a wide range of SP measures documented in the literature, including "time to collision" (TTC), "post encroachment time" (PET), "maximum deceleration rate required to avoid a crash (DRAC), among others [(5),(8), (9)]. Many of these measures are themselves functions of individual vehicle-pair interactions, such as, differences in speed, spacing and deceleration rates. Bonsal et.al (10) noted that with regard to modeling SP, it is important to obtain traffic model input parameters that reflect real-world driving behavior in terms of time dependent speed, spacing and deceleration at the vehicle specific level.

Previous research has focused on the use of either a practical heurist method (11) or the use of Genetic Algorithms (GA) based on a single SP fitting function [(5), (6), (12)]. In the approach proposed by Hourdakis et. al. (11), error was first minimized with respect to volume, and once this was achieved minimized with respect to speed. However, this approach is rather ad hoc in nature since the search technique for parameter input values is not objective and a minimum error for volume does not necessarily result in a minimum error for speed or for that matter for SP.

Cunto and Saccomanno (5) proposed the application of GA to obtain best estimate input parameters in microscopic traffic models. GA is a search technique that mimics processes found in nature, and involves a number of sequential steps (13):

1. Start with a population with random 'chromosomes' (initial population of parameter values)
2. Calculated the fitness function for each chromosome (e.g. RMSE of speed)
3. Randomly decide to either 'mutate' or 'crossover'
4. If mutation is selected, then we choose a 'chromosome' randomly from our population and change the parameter values (govern by random process) to create an 'offspring' chromosome
5. If crossover is selected, we choose two 'chromosomes' and randomly swap information (parameter values) to create an offspring chromosome
6. Drop the chromosome (parameter set) that yields the worse fitness results (e.g. higher RMSE)
7. Return to step 3, and repeat the process.
8. The process is terminated when either the maximum number of iterations is achieved or there are no changes in parameter sets between iterations

Table 1 summarizes the results of several recent calibration exercises involving single fitness functions applied to several microscopic traffic simulation platforms. Park and Qi (12) calibrated VISSIM parameters using travel time as the MOP at an intersection between US15 and US250 in Virginia. Kim et. al. (6) calibrated VISSIM parameters for an arterial road in Houston, Texas using travel time as the MOP. Ma and Abdulhai (7) demonstrated a GA based calibration procedure for PARAMICS using vehicle flow as the MOP. In a recent study of surrogate safety measures by Cunto and Saccomanno (5) used GA to calibrate various VISSIM input parameters based on a Crash Potential Index (CPI) as the SP fitting function. Observed estimates of the MOP were obtained from NGSIM vehicle tracking data (3) for an intersection on Lankershim Boulevard in Los Angeles.

Despite differences in the type of optimization and the nature of the underlying fitting functions, all these studies share a common limitation in that best estimates of model parameters are obtained using a 'single' fitting function (volume, travel time or SP). As such the minimization of error fails to reflect a more complete balance of model performance that encompasses both traffic attributes as well as safety. As noted previously, input parameters that satisfy the minimization of error for one fitting function (say volume) does not ensure that they minimize the error for other functions, say, vehicle speed or spacing or safety performance.

TABLE 1 Summary of Parameter Calibration Research Literature Results

Study	Type of optimization	Model	Network Type	Measure of Performance	results of best parameter estimate	Notes
Hourdarkis et. al (2003)	heuristic search	AIMSUM	Freeway	volume	8.84 % (RSPE)	Root mean square percentage error
Park and Qi (2005)	genetic algorithm	VISSIM	Freeway interchange	travel time	12.6 % (RSPE)	Root mean square percentage error
Kim et. al (2005)	genetic algorithm	VISSIM	Freeway network	travel time	1 % (MAER)	Mean absolute error ratio
Ma and Abdulhai (2002)	genetic algorithm	PARAMICS	Arterial network	flows	46.09 % (GRE)	Global relative error
Cunto and Saccomanno (2008)	genetic algorithm	VISSIM	Intersection	CPI (Crash Potential Index)	0.026 % (RSPE)	Root mean square percentage error

The multi-criteria calibration approach proposed in this paper permits model calibration based on **n** indicators of traffic and/or safety performance (e.g. speed, volume, travel time, CPI, etc.). The two most common multi-criteria methods cited in the literature are (14):

1) Weighting Method which can be expressed mathematically as:

$$\min z(\boldsymbol{x}) = w_1\, z_1(\boldsymbol{x}) + w_2 z_2(\boldsymbol{x}) + \cdots + w_i z_i(\boldsymbol{x}) \qquad (1)$$

$$\text{s.t. } \mathbf{x} \in \mathbf{X}$$

where, w_i is the weight for the measure of performance i
z_i is the measure of performance i (e.g. RMSE of speed)
$\mathbf{x}$ is a vector of parameters
$\mathbf{X}$ is the feasible region of the parameters

2) ε-Constrained Method which can be expressed as:

$$\min z_l(\boldsymbol{x}) \quad (2)$$
$$\text{s.t.} \quad \mathbf{x} \in \mathbf{X}$$
$$z_i(x) \leq \varepsilon_i \text{ for all i=1,2,…,l-1, l+1,…,p}$$

One of the MOP functions in the ε-constrained method is arbitrarily chosen for minimization and the other functions are treated as constraints.

In other types of applications, Madsen (15) used the MOP of low flow RMSE and peak flow RMSE to calibrate the MIKE 11/NAM rainfall-runoff model parameters for a Danish catchment basin. He demonstrated that using a two criteria (MOP) approach produced better estimates of runoff when compared to a single MOP fitting function. Grierson (16) demonstrated the use of multi-criteria parameter calibration for a building design (media centre) project in France, that considered several measures, such as, optimal lighting, thermal heat loss, aesthetics, and building costs. In the transportation field, Abdelghany and Mahmassani (17) used a multi-criteria approach to solve a dynamic trip assignment problem for an intermodal network, based on MOP values of average total travel time, average passenger time and average vehicle time for buses and cars.

In general, the solution for the above Equations 1 or 2 will comprise a Pareto set of solutions rather than a single unique solution, since there will be trade-offs between the different fitting function errors or measures of performance (MOP). Grierson (16) has noted that the set of Pareto solutions represents a front that satisfies the property of 'non-dominance', such that x_i is a Pareto front if two conditions are met:

1. For all non-members x_k there exists at least one member of the front i where $z_i(x_i) < z_i(x_k)$ for all i = 1,…,l.
2. It is not possible to find a x_k within the Pareto front such that $z_i(x_k) < z_i(x_i)$ for all i = 1,…,l.

All solutions along the Pareto front are considered 'equally' good, since any point on this front is not dominated by any other point. The above properties means that any member of the Pareto set will be better than other members for some criteria but could be worse for other criteria. However, there are methodologies that can be utilized to objectively select a unique Pareto-compromised set of solutions.

Grierson (16) demonstrated that the Pareto front can be normalized to eliminate the scaling problem inherent in the fitness functions for different measures of performance. The normalization procedure permits a geometric transformation of the fitting functions in order to find the Pareto member closest to the origin (the unique optimal solution). The application discussed in this paper is concerned primarily with obtaining the Pareto front and demonstrating the multi-criteria parameter calibration procedure as applied to a microscopic traffic simulation platform. The methodology of finding a unique Pareto-compromised solution will be demonstrated in future research

FORMULATION OF ERROR ANALYSIS

For this type of problem, “best estimate” parameter values cannot be obtained directly using conventional statistical fitting. The simulation model produces a range of traffic attributes for individual vehicles over time, including operating speed, spatial progression and acceleration. Microscopic measures of safety performance, such as “time to collision” (TTC), “deceleration rate to avoid the crash” (DRAC), “crash potential index” (CPI) are “composite” measures of individual vehicle traffic attributes expressed over time and space. Hence, for the simulation of safety performance a sound calibration process must reflect a duality of purpose: a) accurate measures of traffic flow, and b) accurate measures of safety performance. Best estimate parameter values for traffic simulation must satisfy both of these requirements.

The parameter input calibration can be quite complex, since depending on the underlying simulation platform, “best estimate” values need to be specified for a rather extensive array of significant model parameters affecting traffic. The two most widely distributed microscopic traffic models VISSIM and PARAMICS, for example, each require specification of a large number of input parameters that affect driving behavior and network performance. In this paper, we will focus on the VISSIM Version 4.3 platform, which requires specification 33 parameters.

The calibration process can be quite complex given this large number of parameters in need of specification. Consider a model consisting Z_k parameters, such that:

$$Z = [Z_1, Z_2, ...Z_k]$$

The first step in the process involves a screening procedure with the aim of reducing the original set of Z_k parameters to a smaller more manageable number for further analysis. Cunto and Saccomanno (5) applied factorial design methods to reduce the 33 parameters in VISSIM to 6 statistically significant parameters. The best estimate values for these parameters needs to be determined in order to obtain accurate measures of traffic attributes, and subsequently safety performance.

In general, the calibration of a traffic simulation platform, such as VISSIM, can be formulated as a Pareto optimization problem governed by ***two*** objective traffic-based fitting functions: e.g. root-mean squared percentage error (*RSPE) of speed* $\{f_1(\mathbf{z})\}$ and RSPE of volume $\{f_2(\mathbf{z})\}$ both of which can be solved using a GA to obtain best-fit estimates for parameters Z_1 to Z_k, such that:

$\mathbf{z} = [z_1, z_2, \ldots, z_k]^T$ in the feasible domain $\boldsymbol{\Omega}$:

$$\text{Minimize } F = \{ f_1(\mathbf{z}), f_2(\mathbf{z}) \} \qquad (3a)$$

$$\text{Subject to } \mathbf{z} \in \boldsymbol{\Omega} \qquad (3b)$$

The two objective fitting functions can be expressed as:

$$f_1(\mathbf{z}) = \Sigma [O^s - E^s(\mathbf{z})]_i^2 \qquad (i = 1,\ldots,N) \qquad (4a)$$

$$f_1(\mathbf{z}) = \Sigma [O^v - E^v(\mathbf{z})]_i^2 \qquad (i = 1,\ldots,N) \qquad (4b)$$

The feasible domain $\boldsymbol{\Omega}$ is defined by the following constraints for the Z_k parameters being calibrated:

$$\begin{aligned} 50 &\le \textit{Maximum look ahead distance} \le 300 \\ 0.5 &\le CC0 \le 3 \\ -15 &\le CC1 \le -4 \\ 0.1 &\le CC3 \le 2 \\ 2 &\le CC5 \le 20 \\ -3.5 &\le \textit{Accepted deceleration of trailing vehicle} \le -0.5 \\ 0.2 &\le \textit{Safety distance factor} \le 0.8 \end{aligned} \tag{3}$$

Where:

O^s = Observed aggregated traffic speed *(known data)*
O^v = Observed aggregated traffic volume *(known data)*
$E^s(\mathbf{z})$ = Estimated aggregated traffic speed *(found from traffic simulator)*
$E^v(\mathbf{z})$ = Estimated aggregated traffic volume *(found from traffic simulator)*
i = index of a time interval over which traffic speed and volume are aggregated
N = number of time intervals for which traffic speed and volume are aggregated
$\mathbf{z} = [z_1, z_2, \ldots, z_7]^T$ variable parameters used by the traffic simulator to estimate traffic speed and volume *(found by the Genetic Algorithm)*

The solution for this problem is a set of calibration parameters $\mathbf{z}^*$, for which each of the corresponding error terms $f_1(\mathbf{z^*})$ and $f_2(\mathbf{z^*})$ reflect a Pareto-optimal solution, such that no other feasible set of parameters $\mathbf{z}$ can be obtained where $f_1(\mathbf{z}) < f_1(\mathbf{z}^*)$ *and* $f_2(\mathbf{z}) < f_2(\mathbf{z}^*)$. The method also yields a Pareto-compromise set of parameter values $\mathbf{z}^o$ for which $f_1(\mathbf{z}^o)$ and $f_2(\mathbf{z}^o)$ represent an equally-balanced trade-off in the error estimates between speed and volume.

CALIBRATION FRAMEWORK

Cunto and Saccomanno (5) proposed a five step screen process to reduce the initial number of inputparameters in VISSIM to the most significant factors. Their process illustrated in Figure 1 consists of the following steps:

1. Selection of initial parameters based on engineering judgment and literature review.
2. Initial screening of parameters using a Plackett-Burnman factorial design for each measure of performance.
3. Establish linear expressions relating significant parameters to the measures of performance using fractional factorial experimentation.
4. Obtain parameter sets using genetic algorithm based on either the weighted method or ε-constrained method.
5. Choose the compromise parameter set solution from the resultant Pareto front

The above-mentioned framework will be demonstrated through a case study application.

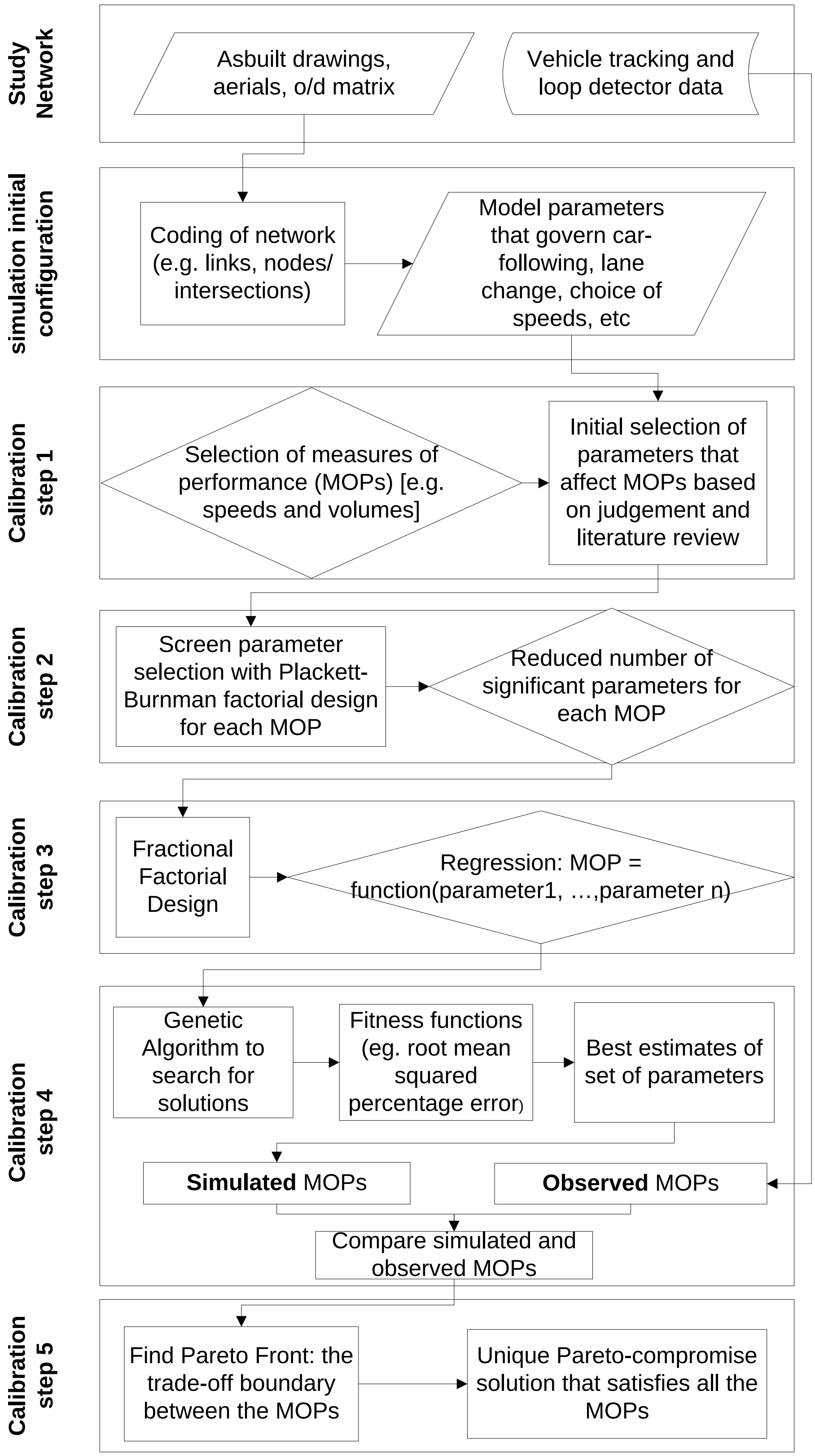

FIGURE 1 Calibration framework.

CASE STUDY

As noted previously, observational traffic data for calibration was obtained from the Next Generation SIMulation (FHWA, 2007). As illustrated in Figure 2, data were extracted from a freeway segment of Highway 101 in California.

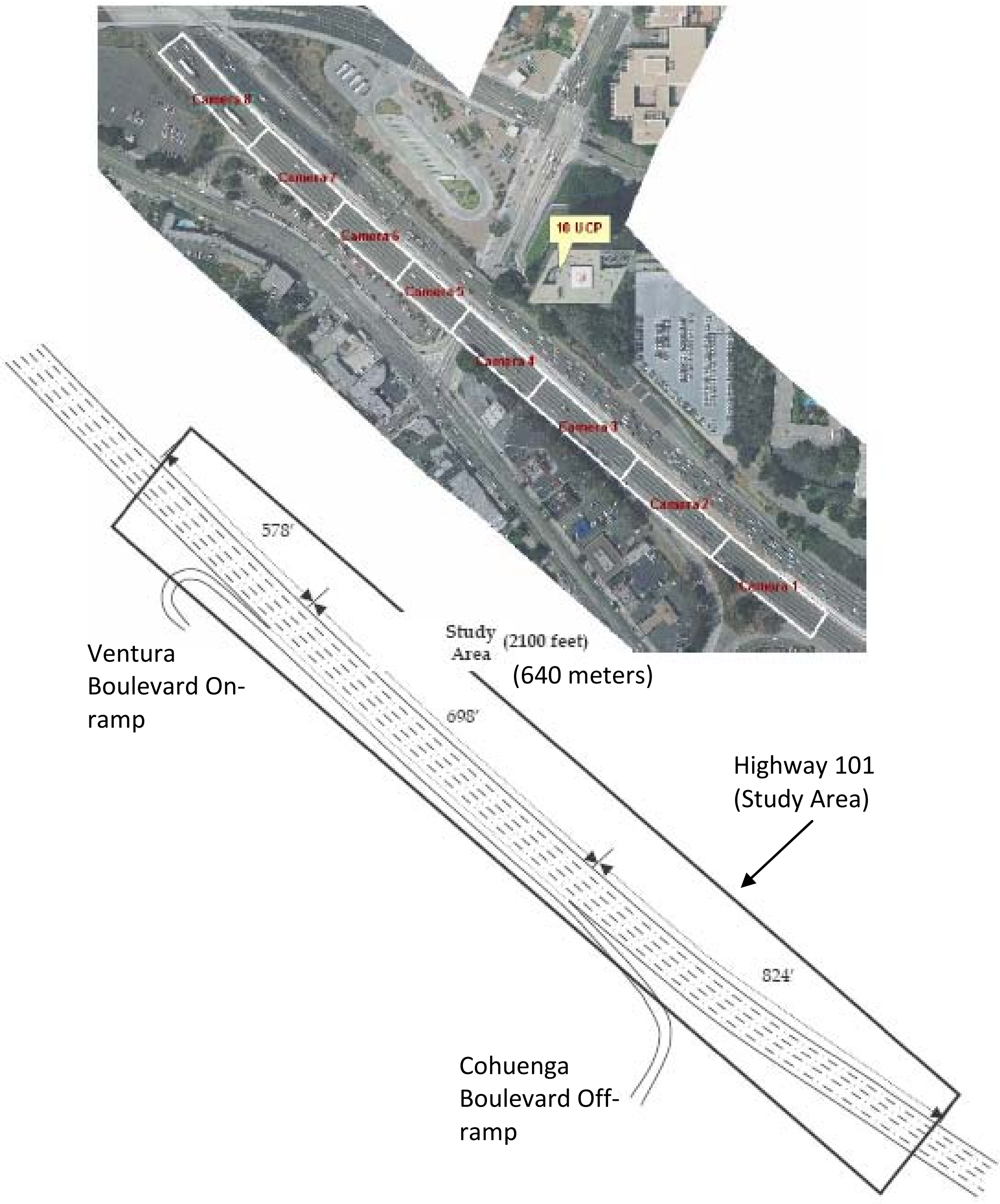

FIGURE 2 Study area of US Highway 101 (NGSIM 2007).

A sensitivity analysis was undertaken to find parameters with statistically significant effects on the various measures of performance. Cunto and Saccomanno (5) introduced a CPI measure of SP based on a comparison between maximum deceleration requirements and braking capability estimated in time increments on a vehicle specific basis. Other measures of MOP include speed and volume. The CPI SP measure is defined in terms of the probability that a given following or response vehicle deceleration rate needed to avoid a crash (DRAC) with a lead or stimulus vehicle exceeds its maximum available deceleration rate (MADR). Since MADR is vehicle and scenario-specific, separate values of MADR need to be specified for each vehicle in the traffic stream; such that

$$CPI_i = \frac{\sum_{t=ti_i}^{tf_i} P\left(DRAC_{i,t} > MADR_{i,t}\right) * \Delta t * b}{T_i} \qquad (2.1.3)$$

where,

CPI_i = crash potential index for vehicle i
ti_i = initial time interval for vehicle i
tf_i = final time interval for vehicle i
Δt = observation time interval (seconds)
T_i = total simulated time for vehicle i (seconds)

A Plackett-Burnman factorial design was applied to the original set of VISSIM parameters using ANOVA. Table 2 describes these parameters in VISSIM and it remains to obtain their best estimate values. The results shown in Tables 3, 4, and 5 produced seven significant parameters for further analysis. For this freeway application, the screening procedure yielded seven statistically significant parameters that affect traffic and SP attributes in VISSIM.

TABLE 2 VISSIM Parameters that Affect MOPs of Speed, Volume and CPI

	VISSIM Parameter	Description
A	Maximum look ahead distance	Defines the distance that vehicles can see forward to react to other vehicles in front or beside it on the same link
B	CC0:	Standstill distance (m), defines the desired distance between stopped vehicles
C	CC1	Is the headway time in seconds that a driver wants to keep
D	CC3	Threshold for entering Following, controls the start of the deceleration process
E	CC5	Following thresholds control the speed differences during the following state. Smaller values result in a more sensitive reaction of drivers to accelerations or decelerations of the preceding car
F	Acceptable deceleration of trailing vehicle	Affects lane change behaviour
G	Safety distance factor	Takes effect for; a) the safety distance of the trailing vehicle in the new lane for the decision whether to change lanes or not, b) the own safety distance during a lane change and c) the distance to the leading (slower) lane changing vehicle

TABLE 3 ANOVA Table for Main and Two-Factor Effects on Volume

Model		Coefficients Beta	t	Sig.
	(Constant)		497.396	.000
	A	.326	5.636	.000
	B	-.136	-2.349	.023
	C	-.406	-7.008	.000
	D	.229	3.954	.000
	E	.035	.599	.551
	F	.002	.034	.973
	G	-.427	-7.377	.000
	CD	.253	4.362	.000
	DE	-.108	-1.858	.069
	DF	-.130	-2.243	.029
	DG	.263	4.535	.000
	EF	-.366	-6.322	.000
	EG	.002	.034	.973
	FG	.038	.660	.512

TABLE 4 ANOVA Table for Main and Two-Factor Effects on Speed

Model		Coefficients Beta	t	Sig.
	(Constant)		425.789	.000
	A	-.303	-10.446	.000
	B	-.172	-5.927	.000
	C	-.598	-20.596	.000
	D	-.266	-9.168	.000
	E	-.112	-3.861	.000
	F	-.127	-4.390	.000
	G	-.298	-10.253	.000
	CD	-.311	-10.717	.000
	DE	-.127	-4.364	.000
	DF	-.141	-4.868	.000
	DG	-.247	-8.509	.000
	EF	-.260	-8.961	.000
	EG	-.115	-3.977	.000
	FG	-.117	-4.042	.000

TABLE 5 ANOVA Table for Main and Two-Factor Effects on CPI/Vehicle

Model	Coefficients		
	Beta	t	Sig.
(Constant)		8.690	.000
A	-.175	-2.503	.015
B	.000	.001	.999
C	.321	4.588	.000
D	-.266	-3.811	.000
E	-.020	-.285	.776
F	-.049	-.696	.490
G	.459	6.562	.000
CD	-.256	-3.654	.001
DE	.001	.018	.986
DF	.004	.062	.950
DG	-.338	-4.831	.000
EF	.374	5.342	.000
EG	-.019	-.277	.783
FG	-.062	-.882	.382

In this paper two fitting procedures were investigated: Procedure 1 involving a multi-criteria MOP where root mean square error is minimized with respect to the two traffic attributes: speed and volume. Procedure 2 involving a single SP based criterion where the root mean square error is minimized in terms of CPI/vehicle (Cunto and Saccomanno, 2008). In both procedures, a GA (MATLAB version R2007) was used to objectively search for the best fit parameter set. The population of parameter values in the GA was kept the same for both procedures so as to reduce the variability in error caused by randomness in the initial parameter selection. Tables 6 and 7 summarize the VISSIM parameter values and corresponding RSPE for different GA searches.

TABLE 6 Approach 1- Multi-Criteria Approach (Calibrating using both RSPE Speed and Volume; Corresponding RSPE CPI Shown)

Paramter Set Number	(max) Look ahead Distance (m)	CC0	CC1	CC3	CC5	Accepted deceleration of trailing vehicle for lance change	Safety distance reduction factor	Speed (km/h)	Volume (veh)	CPI/veh (10^{-10})	RSPE Speed (%)	RSPE Volume (%)	Relative Error in CPI/veh (%)	Sum of Errors (%)
1	294.62	2.87	1.49	-4.05	1.96	-0.29	0.79	73	1975	973,645	25.6	8.3	10.0	33.9
2	299.78	2.58	1.50	-4.01	1.99	-0.26	0.73	59	1926	866,478	1.9	10.5	2.1	12.4
3	297.55	2.99	1.50	-4.47	2.00	-0.25	0.80	70	1945	1,292,068	20.4	9.7	45.9	30.1
4	293.59	2.94	1.47	-4.24	1.99	-0.26	0.78	82	1989	854,907	41.2	7.6	3.4	48.9
5	282.86	2.74	1.37	-4.02	1.98	-0.28	0.79	80	1988	1,199,004	37.5	7.7	35.4	45.2
6	299.08	2.86	1.50	-4.01	1.98	-0.27	0.77	70	1940	1,508,226	20.1	9.9	70.3	29.9
7	297.26	3.00	1.50	-4.03	2.00	-0.25	0.80	71	1957	1,488,249	21.3	9.1	68.1	30.4
8	279.22	2.95	1.44	-4.16	1.93	-0.33	0.79	74	1997	1,332,134	27.7	7.3	50.5	34.9
9	277.90	2.94	1.48	-4.25	2.00	-0.25	0.80	74	1968	1,426,636	26.1	8.6	61.1	34.7
10	297.78	2.92	1.49	-4.04	1.97	-0.28	0.79	62	1947	943,160	6.2	9.6	6.5	15.8
11	294.80	2.98	1.48	-4.56	1.97	-0.28	0.80	63	1986	952,702	7.2	7.7	7.6	15.0
VISSIM Defaults	250	1.5	0.9	-8	0.35	-0.5	0.6	104	1992	539,547	78.7	7.5	39.1	86.2
Observed								**58**	**2153**	**885,402**				

TABLE 7 Approach 2 – Single Criterion using RSPE CPI (Corresponding RSPE Speed and Volumes are shown)

Parameter Set Number	(max) Look ahead Distance (m)	CC0	CC1	CC3	CC5	Accepted deceleration of trailing vehicle for lance change	Safety distance reduction factor	Speed (km/h)	Volume (veh)	CPI/veh (10^{-10})	RSPE Speed (%)	RSPE Volume (%)	Relative Error in CPI/veh (%)
1	176.18	0.50	1.00	-8.85	1.66	-0.65	0.50	103	2057	875,155	77.0	4.5	1.2
2	201.31	2.67	1.44	-4.96	1.11	-1.30	0.78	71	1958	1,432,328	21.8	9.1	61.8
3	279.16	1.89	1.47	-5.64	1.30	-1.08	0.74	87	1978	667,775	49.2	8.1	24.6
4	272.78	2.92	1.12	-11.21	0.75	-1.74	0.54	102	2025	394,974	75.2	5.9	55.4
5	208.35	1.40	1.39	-6.03	1.10	-1.32	0.60	95	2017	1,066,616	62.8	6.3	20.5
6	180.89	1.27	1.03	-7.60	1.25	-1.14	0.55	102	1774	769,106	75.8	17.6	13.1
VISSIM Defaults	250	1.5	0.9	-8	0.35	-0.5	0.6	104	1992	539,547	78.7	7.5	39.1
Observed								**58**	**2153**	**885,402**			

Figure 3 illustrates a comparison between the multi-criteria set of solutions and the single-criterion results. In addition, this figure provides an indication of the fitting errors associated with the VISSIM default parameter values.

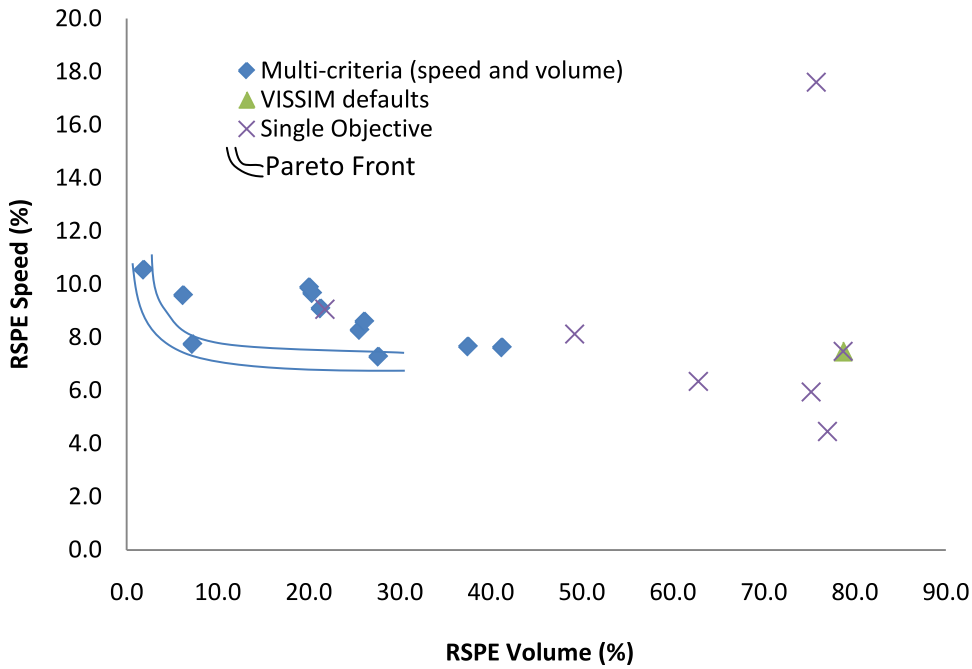

FIGURE 3 RSPE volume versus RSPE speed.

Both the multi-criteria and the single SP-based criterion yielded reasonable root mean squared percentage errors in comparison to NGSIM observations. However, as shown in Figure 3, the single criteria procedure provides good estimates of speed with respect to the multi-criteria values, but poor results for volume. This potentially presents a problem of faith in the relevance of the simulation approach applied to safety. An absence of precision in traffic attributes cannot yield precision in safety performance if these attributes are themselves inputs in the SP measure. The proposed multi-criteria procedure on the other hand, results in good estimates for both speed and volume (with respect to NGSIM) and also reasonable measures of safety performance. Particularly problematic is the arbitrary adoption of default values in safety studies. In Figure 3, default values resulted in the highest RSPE errors for volume.

We note that from Table 6 for the multi-criteria procedure the lowest error in speed does not match the lowest error in volume, or the lowest error in CPI/vehicle. This presents a major calibration challenge, and any calibration exercise will need to address the issue of balance between traffic attributes and SP, as well as overall model fitness considering all the traffic attributes. Unfortunately this issue is not within the scope of this paper, but is currently being investigated.

CONCLUSIONS

A multi-criteria calibration procedure has been presented for obtaining best estimate parameter values for a microscopic traffic simulation platform (VISSIM). The multi-criteria procedure based on traffic attributes is preferable to using a more abstract measure of safety performance that is not as easily verified in the observed traffic data. A comparison of current state-of-the-art calibration based on safety performance and the proposed multi-criteria procedure based on speed and volume indicates that the multi-criteria procedure is able to yield better estimates of traffic attributes, in addition to comparable estimates of safety performance.

When properly calibrated, simulation models can provide useful information on individual driver responses to changing traffic and geometric road conditions. Since traffic attributes are an integral part of safety performance, a calibration based on traffic attributes provides a more thorough basis for investigating safety at a given location.

ACKNOWLEDGEMENT

We are grateful to Professor Don Grierson for his valuable comments and suggestions on this research.

REFERENCES

1. Tarko, A.P., Songchitruksa, P. 2005.Estimating the Frequency of Crashes as Extreme traffic Events. Presented at 84th Annual Meeting of the Transportation Research Board, Washington D.C.

2. Saunier, N., Sayed, T. 2007. Automated Road Safety Analysis Using Video Data. *Transportation Research Record. Vol. 2019.* pp 57-64.

3. Federal Highway Administration (FHWA), *US Highway 101 Dataset*, Next Generation SIMulation Fact Sheet, FHWA-HRT-07-030, 2007, http://www.tfhrc.gov/about/07030.htm, Website accessed July 2007.

4. Hellinga, B. R. *Requirements for the calibration of traffic simulation models.* Proceedings of the Canadian Society for Civil Engineering, 1998.

5. Cunto, F., Saccomanno, F.F. Calibration and validation of simulated vehicle safety performance at signalized intersections. *Accident Analysis and Prevention*. Vol. 40, pp 1171-1179, 2008.

6. Kim, SJ, Kim W, and Rilett, LR. Calibration of micro-simulation models using non-parametric statistical techniques. *In Transportation Research Record: Journal of the Transportation Research Board, No.1935,* Transportation Research Board of the National Academies, Washington, D.C., pp. 111–119, 2005.

7. Ma, T, and Abdulhai, B. Genetic algorithm-based optimization approach and generic tool for calibrating traffic microscopic simulation parameters. *In Transportation Research Record: Journal of the Transportation Research Board, No.1800,* Transportation Research Board of the National Academies, Washington, D.C., pp. 6–15, 2002.

8. Gettman, D., and Head, L., Surrogate safety measures from traffic simulation models, Final Report, Federal Highway Administration, Publication No. FHWA-RD-03-050, 2003.

9. Archer, J. Methods for the assessment and prediction of traffic safety at urban intersection and their application in micro-simulation modelling, PhD Thesis, Department of Infrastructure, Royal Institute of Technology, Stockholm, Sweden, 2005.

10. Bonsall, P., Liu, R., and Young, W. Modelling safety-related driving behaviour – impact of parameter values. In Transportation Research Part A, Vol 39, pp. 425 – 444, 2005.

11. Hourdakis, J, Michalopoulos, P, and Kottommannil, J. A practical procedure for calibrating microscopic traffic simulation models. Presented at the 82nd Annual Meeting of the Transportation Research Board, Washington, DC., January, 2003.

12. Park B and Qi H. Development and evaluation of a procedure for the calibration of simulation models. *In Transportation Research Record: Journal of the Transportation Research Board, No.1934*, Transportation Research Board of the National Academies, Washington, D.C., pp. 208–217, 2005.

13. Elbetagi, E., Hegazy, T., and Grierson, D. Comparison among five evolutionary-based optimization algorithms. *Advance Engineering Informatics, Vol 19,* pp 43-53, 2005.

14. Goicoechea, A., Hansen, D.R., and Ducksten, L. *Multiobjective decision analysis with engineering and business applications*. Toronto: John Wiley & Sons Inc, 1982.

15. Madsen, H. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology,* Vol 235, 2000, pp. 276-288.

16. Grierson, D. Pareto multi-criteria decision making. *Journal of Advance Engineering Informatics, Vol 22,* 2008, pp 371-384.

17. Abdelghany, K.F. and Mahmassani, H.S. Dynamic trip assignment-simulation model for intermodal transportation networks. . *In Transportation Research Record: Journal of the Transportation Research Board, No.1771,* Transportation Research Board of the National Academies, Washington, D.C., pp. 52–60, 2001.