Holistic Calibration of Microscopic Traffic Flow Models: Methodology and Real World Application Studies

Alexander Paz, Victor Molano and Javier Sanchez-Medina

Abstract This study proposes and applies a methodology to calibrate microscopic traffic flow simulation models. The proposed methodology has the capability to calibrate simultaneously all the calibration parameters as well as demand patterns for any type of network. Parameters considered include global and local as well as driver behaviour and vehicle performance parameters. Demand patterns, in terms of turning volumes, are included in the calibration framework. Multiple performance measures involving link counts and speeds are used to formulate and solve the proposed calibration problem. In addition, multiple time periods were considered. A Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm is used to search for the vector of the model's parameters that minimizes the difference between actual and simulated network states. (Punzo V, Ciuffo B, Montanino M Transp Res Rec J Transp Res Board 2315(1):11–24 2012, Punzo et al. [1]) commented on the uncertainties present in many calibration methodologies. The motivation to consider simultaneously all model parameters is to reduce that uncertainties to a minimum, by leaving to the experience of the engineers as little parameter tuning as possible. The effects of changing the values of the parameters are taken into consideration to adjust them slightly and simultaneously. This results in a small number of evaluations of the objective function. Three networks were calibrated with excellent results. The first network was an arterial network with link counts and speeds used as performance measurements for calibration. The second network included a combination of freeway ramps and arterials, with link counts used as performance measurements.

A. Paz · V. Molano
Department of Civil and Environmental Engineering, University of Nevada, Las Vegas, USA
e-mail: apaz@unlv.edu

V. Molano
e-mail: victor.hugo.molano@gmail.com

J. Sanchez-Medina (✉)
CICEI, Department of Computer Science, University of Las Palmas de Gran Canaria (ULPGC), Las Palmas, Spain
e-mail: javier.sanchez@ulpgc.es; javier.sanchez.medina@ieee.org

N.D. Lagaros and M. Papadrakakis (eds.), *Engineering and Applied Sciences Optimization*, Computational Methods in Applied Sciences 38,
DOI 10.1007/978-3-319-18320-6_3

The third network was an arterial network, with time-dependent link counts and speed used as performance measurements. The experimental results illustrate the effectiveness and validity of this proposed methodology. The same set of calibration parameters was used in all experiments.

1 Introduction

Micro-simulation models provide tremendous capabilities to model, at a high level of resolution, complex systems in a broad range of fields, including economy, sociology, physics, chemistry, and engineering [2].

In the context of vehicular traffic systems, microscopic traffic flow models enable the modelling of many aspects of the actual system, including the manoeuvres of individual vehicles and their interactions, the various types and characteristics of facilities, and the vast number of control settings. These capabilities are associated with a large number of modelling parameters that typically need to be tailored for each vehicular system. For example, driver behaviour includes parameters associated with car following, lane-changing manoeuvres, and gap acceptance.

In, Punzo et al. reflect on the uncertainties present in many of the current car-following based traffic flow simulation calibration methodologies. It is a fact that the accuracy of a model and the validity of its results are highly dependent on the correctness of the chosen parameters [3–9].

Punzo et al. [1] discussed uncertainties present in many of the existing methodologies for the calibration of car-following-based traffic flow simulation models. It is clear that the accuracy of a model and the validity of its results are highly dependent on the correctness of the chosen parameters [3–9].

Hence, it is important to consider all these model parameters simultaneously with the aim to capture their intricate interactions, thereby seeking convergence and stability of the solutions.

In [10] we drafted a method for the simultaneous calibration of all of the parameters of a CORSIM model. In the present work we have sharpen, extended and applied that methodology to three different big test cases with excellent results: (i) Pyramid Highway, in Reno, Nevada, USA; (ii) Interstate-75 in Miami, Florida, USA; and (iii) a Network of McTrans Sample Data Sets.

This study proposes a methodology to calibrate simultaneously all model parameters and demand patterns based on link counts and speeds. In addition, multiple performance measures were used, demand patterns were not pre-calibrated, and multiple time periods were explicitly considered with target performance values for each period. That is, the proposed methodology implements a Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm to determine an adequate set for all model parameters and turning volumes for multiple time periods using multiple performance measures. Even though there is a significant body of literature around the proposed problem context, to the best of the authors knowledge, no study has considered simultaneously all the aspects listed in this paragraph and included in our implementation and experimental framework. The state-of-the-art is summarized in the following subsection.

The SPSA was chosen based on its computationally efficiency and ability to handle large numbers of parameters [11–18]. Only two traffic flow simulation evaluations per iteration of the SPSA are required to update all model parameters. Running a low number of traffic flow simulations represents important savings in terms of time and other resources. However, the SPSA algorithm performs better when the initial model parameters relatively close to the optimal solution.

Comparative studies between SPSA and other algorithms could be found in the literature [11, 12, 18]. In addition, the SPSA algorithm has been used to calibrate and optimize various transportation applications [13, 19, 20].

The rest of this paper is organized as follows: We do a brief literature review in the next subsection. We expose the proposed methodology in Sect. 2. Then we share the experiments performed alongside with the corresponding results in Sect. 3. Finally we put together some concluding remarks in section .

1.1 State of the Art

A broad number of optimization algorithms, ranging from genetic algorithms to finite difference stochastic approximation, have been used to determine an adequate set of model parameters for a particular traffic system [3, 4, 6, 21, 22].

For example, the sequential simplex algorithm was used to calibrate parameters for car-following, acceleration/deceleration, and lane-changing behaviour [6]. However, only a subset of parameters was considered, maybe because of the lack of enough computing power in 2002. Moreover, parameters associated with infrastructure and vehicle performance were not considered. The algorithm provided adequate results under congested conditions. However, under low-congestion conditions, manual calibration provided better results [6].

In [23] they calibrate the VISSIM model of the NGSIM corridor, using a quite limited optimization technique, exploring only the limits. They calculate a number of restrictions for some parameters and accept values only if they satisfy all the restrictions. Additionally, they are only tuned to a specific period of the day.

In a recent study, [24], Markov Chain Monte Carlo (MCMC) method using Bayesian estimation theory. Only five parameters of a linear car following model [25] are calibrated.

Genetic Algorithms (GA) has been extensively used to calibrate traffic simulation parameters. In [26] the use a simple GA to calibrate the parameters of a CORSIM [27] based simulation of a 5.8 km expressway in Singapore. In [28], a freeway segment in California was used as a test example to attempt the optimization of two PARAMICS calibration parameters.

In both cases, the results proved limited success reducing discrepancies between real word and simulations.

Genetic Algorithms were used for the calibration of global and local capacity and occupancy parameters [20, 29]. A sequential approach was used to update global and local parameters.

In [30] a Genetic Algorithm was used to calibrate a small subset of all the PARAMICS [31] parameters.

In [32] a Multiobjective version of the Non-dominated Sorting Genetic Algorithm (NGSA-II, [33]) was applied to solve the multi-objective optimization task of parameter calibration. Results are modest and they were optimizing or calibrating a very few, only five of VISSIM's [34].

In [35] five PARAMICS [31] parameters were optimized for a larger model of down town Toronto, Canada. They tested three different GA approaches but they finally did not obtained significant improvements in the accuracy of the model.

In [8] they tuned 11 CORSIM [36] parameters of a 22.4 km segment of Interstate 10 in Houston, Texas. The authors used a Genetic Algorithm to perform an automated calibration of these parameters. Their results were remarkable, including a sensitivity analysis. As happens for every GA approach to traffic simulation calibration, there were a few set-up parameters in the Genetic Algorithm that must be carefully selected, because the quality of results is very dependent on them. There was no computing performance information provided for such work, which should be a very interesting element for comparison with SPSA-based approaches, likely to be faster, more suited to real world on-line applications.

In [37] yet another GA based PARAMICS parameter calibration was proposed. The authors only calibrate 5 parameters that needed to be initialized at "default values". In addition, there were eight additional configuration parameters that need to be tuned for the Genetic Algorithm to obtain better performance. This parameter adjustment required significant trial-and-error and experience by the researcher.

Regarding specifically SPSA algorithms we have selected a few interesting and related studies. In [13] Lee used SPSA algorithms to calibrate model parameters using distributions to generate input for various stages. The calibration capabilities of GA and SPSA algorithms were shown to be similar in [20]; however, SPSA algorithms were less computationally expensive.

In [38], the authors proposed a SPSA algorithm for the calibration of a simulation model of the Massachusetts Bay Transportation Authority (MBTA) Red Line. The authors used a generic simulator, SimMETRO. The effort involved a multiple objective function and simultaneous parameter calibration. It is important to notive that the simulation of one Metro line involves less parameters compared to a vehicular traffic system. This makes the problem more computationally affordable and less complex.

In they proposed a rail simulation SPSA based parameters calibration for the test case of the Massachusetts Bay Transportation Authority (MBTA) Red Line, using a generic simulator they called SimMETRO. Even when it is not exactly the same problem to solve than in our case, this is a remarkable application of multiple objective simultaneous parameter calibration. It is also true, though, that a one Metro line simulation has not as many calibration parameters as a vehicular simulation like

CORSIM may include, making the problem more computationally affordable and also less complex.

Another very interesting application of a SPSA algorithm to Intelligent Transportation Systems was published in [39]. A dynamical emission model was optimized to estimate aggregate emission patterns for traffic fleets so as to predict local air conditions.

SPSA and Finite Difference Stochastic Approximation algorithms have been proposed for the calibration of time depending Origin-Destination matrices. For example, in [11] driver behaviour parameters where pre-calibrated considering various time intervals. Other important performance measures, such as speed, were not considered.

In [40], a SPSA algorithm is used for the simultaneous adjustment of a dynamic traffic O-D matrix using traffic counts and speeds. However, the author states that some parameters must be tuned by hand to get close to the desired solutions. Hence, the proposed approach is infeasible for a large amount of calibration parameters as it requires significant user involvement and experience.

$$Min.\,NRMS = \frac{1}{\sqrt{N}} \times \sum_{t=1}^{T} \left(W \times \sqrt{\sum_{i=1}^{N} \left(\frac{V_i - \tilde{V}(\theta)_i}{V_i}\right)^2} + (1 - W) \times \sqrt{\sum_{i=1}^{N} \left(\frac{S_i - \tilde{S}(\theta)_i}{S_i}\right)^2} \right)$$

Subject to:
Lower bound $\leqq \theta \leqq$ Upper bound

(1)

Ben-Akiva et al. worked on the calibration of a dynamic traffic O-D matrix [41] for a large network in Beijing. The SPSA algorithm was used given its capability to address noise. The significant work conducted using the SPSA algorithm to perform related research motivated its use in the proposed study.

are

2 Methodology

2.1 Formulation of the Calibration Problem

The calibration problem for all model parameters, θ, is formulated using a mathematical programming approach. The analysis period is divided into a number T of discrete time periods. The objective function, normalized root mean square (NRMS), as denoted by Eq. 1, is the sum over all calibration time-periods of the average of the sum over all links I of the root square of the square of the normalized differences between actual and simulated link counts and speeds. The normalization enables the consideration of multiple performance measures, in this case, link counts and speeds. In our experimental set-up, the initial parameters for a model are selected as the default values used in CORSIM models. The calibration problem is formulated as shown in Eq. 1, where:

- V_i = actual link counts for link i
- $\tilde{V}(\theta)_i$ = simulated link counts for link i
- S_i = actual speeds for link i
- $\tilde{S}(\theta)_i$ = simulated speeds for link i
- N = total number of links in the model
- T = total number of time periods t
- W = weight used to assign more or less value to counts or speeds

$$g_k\theta_k = \frac{y(\theta_k + c_k\Delta_k) - y(\theta_k - c_k\Delta_k)}{2c_k}[\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \Delta_{k3}^{-1}, \ldots, \Delta_{kp}^{-1}]^T \tag{2}$$

2.2 Calibration Criteria

The calibration criteria for this study were based on guidelines from the Federal Highway Administration. The difference between actual and simulated link counts should be less than 5 % for all links; and, the GEH statistic, in Eq. 3, should be less than 5 for at least 85 % of the links [27].

$$\begin{aligned} GEH &= \sqrt{\frac{2(V_i - \tilde{V}(\theta)_i)^2}{V_i + \tilde{V}(\theta)_i}} \\ V_i &= \text{actual link counts at the link i.} \\ \tilde{V}(\theta)_i &= \text{simulated link counts at the link i.} \end{aligned} \tag{3}$$

2.3 Simultaneous Perturbation Stochastic Approximation Algorithm

The SPSA algorithm is an iterative approach that uses gradient estimations of the objective function to determine an optimal solution. Details of its implementation are provided by Spall [15–18]. In each iteration of SPSA, the vector of model parameters is updated using Eq. 4; where:

$$\theta_{k+1} = \theta_k - a_k g_k \theta_k \tag{4}$$

- θ_{k+1}= vector of updated parameters at iteration k + 1
- θ_k= vector of initial parameters at iteration k + 1
- a_k= gain coefficient at iteration k + 1 calculated using Eq. 5
- $g_k\theta_k$= estimated gradient at iteration k + 1.

$$a_k = \frac{a}{(k + 1 + A)^{\alpha}} \tag{5}$$

where a, A, and α are empirical non-negative coefficients. These coefficients affect the convergence of the SPSA algorithm. The simultaneous perturbation and gradient estimate are represented by $g_k\theta_k$, and is calculated using Eq. 2.

Here, c_k is calculated using Eq. 6 where c and γ are empirical non negative coefficients.

$$c_k = \frac{c}{(k + 1)^{\gamma}} \tag{6}$$

where, $c = 2.7598$ and $\gamma = 0.1666$.

The elements in the random perturbation vector are Bernoulli-distributed, with a probability of one-half for each of the two possible outcomes (Eq. 7).

$$\Delta k = [\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \Delta_{k3}^{-1}, \ldots, \Delta_{kp}^{-1}]^T \tag{7}$$

The SPSA algorithm is implemented using the following steps [18]:

- Step 1: Set counter k equal to zero. Initialization of coefficients for the gain function a, A, and α and calibration parameters θ_0.
- Step 2: Generation of the random perturbation vector Δ_k.
- Step 3: Evaluation of the objective function plus and minus the perturbation.
- Step 4: Evaluation of the gradient approximation $g_k\theta_k$.
- Step 5: Update the vector of calibration parameters using Eq. 4 along with the corresponding constraints denoted by Eq. 3.
- Step 6: Check for stopping criteria. If criteria is achieved, stop; otherwise, set counter $k = k + 1$ and repeat Steps 1–6.
- Convergence is achieved when all the criteria in Table 1 is satisfied or the maximum number of iterations is reached.

2.4 Stopping Criteria

Stopping criteria is reached when the inequality in Eq. (4) is satisfied or a user pre-specified maximum number of iterations is reached. At convergence, the calibration criteria are expected to be satisfied or a significantly better model is obtained.

$$\frac{\sum_{k-n+1}^{k} \sqrt{(NRMS_{AV} - NRMS_k)^2}}{n} < \rho \tag{8}$$

where,

- $NRMS_{AV}$ = average NRMS of the last n iterations
- $NRMS_k$ = NRMS at k iteration
- k = iteration counter
- n = pre-specified integer = 10, and
- ρ = pre-specified convergence condition = 0.015.

3 Experiments and Results

3.1 Micro-simulation Model

The proposed methodology was tested using CORSIM, a tool that integrates two different models to represent a complete traffic system, FRESIM for freeways and NETSIM for surface streets [36, 42]. The Traffic Analysis Toolbox Volume IV: Guidelines for Applying CORSIM Micro-simulation Modelling Software [5] describes a procedure for the calibration of micro-simulation traffic flow models, with a focus on CORSIM. The suggested procedure in these guidelines uses three sequential and iterative steps, including the calibration of (i) capacity at key bottlenecks, (ii) traffic volumes, and (iii) system performance. However, the guidelines do not suggest any particular methodology to perform the calibration in an efficient and effective manner. For example, issues associated with convergence and stability of the solutions are not discussed. Nevertheless, alternative studies have proposed and developed practical procedures to accelerate the calibration process, which typically is time consuming [43]. However, stability and convergence still are issues.

3.2 Calibration Parameters for CORSIM Models

The calibration of CORSIM models can involve Driver Behaviour and Vehicle Performance parameters [36, 42]. These parameters can be defined exclusively for surface streets or freeways or both models simultaneously. In addition, the resolution of these parameters can be global or link-based defined. This study considered all types of parameters and levels of resolution. In addition, parameters related to demand patterns were included. Table 1 shows all the different parameters used for the calibration of CORSIM models. Several studies have conducted sensitivity analysis for the calibration of CORSIM models [8]. These studies have showed that the maximum non-emergency deceleration rate, for example, does not affect the outcomes of a specific FRESIM model. However, the specific vehicle distributions improve the accuracy of the model [8]. Driver behaviour parameters were found to affect the time to breakdown and the flow on ramps. Flow related parameters showed low effects.

Table 1 Calibration parameters for NETSIM and FRESIM models

NETSIM model surface streets		
Driver behaviour	Vehicle performance	Demand patterns
• Queue discharge headway	• Speed and acceleration characteristics	• Surface street turn movements
• Start-up lost time	• Fleet distribution and passenger occupancy	
• Distribution of free-flow speed by driver type		
• Mean duration of parking manoeuvres		
• Lane change parameters		
• Maximum left and right turning speeds		
• Probability of joining spillback		
• Probability of left turn jumpers and laggers		
• Gap acceptance at stop signs		
• Gap acceptance for left and right turns		
• Pedestrian delays		
• Driver familiarity with their path		
FRESIM model-freeways		
• Mean start-up delay at ramp meters	• Speed and acceleration characteristics	• Freeway turn movements
• Distribution of free flow speed by driver type	• Fleet distribution and passenger occupancy	
• Incident rubbernecking factor	• Maximum deceleration	
• Car-following sensitivity factor		
• Lane change gap acceptance parameters		
• Parameters that affect the number of discretionary lane changes		

The calibration parameters have different effects for specific networks and conditions. The interaction between these parameters is very complex and might vary from model to model. As a starting point, the proposed methodology uses a set of default CORSIM values for the parameters listed in Table 1. This decreases the effort during the selection of the calibration parameters and set-up. During calibration, the value of the selected parameters is adjusted while constraining their boundaries in order to avoid unrealistic values.

3.3 Experimental Set-Up and Results

Three experiments were designed to test the capabilities of the proposed methodology to calibrate simultaneously, using vehicle counts and speeds. A software tool was developed to implement the proposed calibration methodology. The tool was developed using a basic layered architecture were each layer handles a group of related functions. A Graphical User Interface (GUI) provides access to the entire software capabilities. The entire software was developed in Java; it includes more than 5,000 lines of code.

System Specifications

- Operative System: Windows Server, Standard Edition, 2007, Service Pack 2 64Bit
- System: Intel Xeon CPU E7450 2.4 GHz (4 processors)
- Ram memory: 32 GB

First Experiment: Pyramid Highway in Reno, Nevada, USA

In this experiment a CORSIM model for a portion of the Pyramid Highway in Reno, Nevada, was calibrated. This portion of highway is located between Milepost 1.673 and 5.131. This calibration focused on speeds and link counts for the entire simulation. The weight factor in the objective function was set to 0.7. This value is constant for the first two experiments because link counts were obtained using more accurate data collection methods compared to speeds. The model included 126 arterial links, and no freeways were included. Link counts and speeds were only available for 45 of these links. Coefficients for the SPSA algorithm were selected using guidelines from the literature [18]). These values affected the convergence of the algorithm. The time required for calibration was 25.5 min.

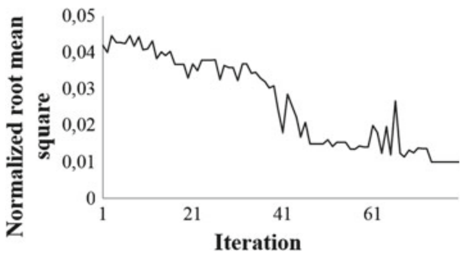
Figure 1a shows a Google map of the Pyramid Highway. Figure 1b illustrates the corresponding CORSIM model. Figure 2 illustrates how the objective function was minimized. The noisy trajectory was a consequence of the stochastic perturbation applied to all calibration parameters to obtain the gradient approximation at each iteration. The characteristics of the traffic model made the function noisier due to rounding. The NRSM was 0.042 before calibration and 0.010 after calibration. The calibration process stopped around the 80th iteration, when a stable region was found.

Figure 3a shows the actual and simulated counts and speeds before calibration. These values present poor initial conditions, especially for the volumes over 1500 vehicles per hour (vph). Figure 3b shows the actual and simulated counts and speeds after calibration. The proposed methodology is able to reduce the gap between actual and simulated counts. The results illustrate larger improvements for the large counts. Figure 3a clearly shows that links with counts over 1500 vph were improved, while the values with good initial conditions were slightly modified.

As illustrated in Fig. 3a, simulated speeds are far from actual speeds. The simulation model underestimates many speed values. After calibration (Fig. 3a), the speeds were improved for 23 of the links. The rest of the speeds were kept close to the initial values with a variation less than 1 mile per hour (mph). This can be associated to the relative large value of the weight assigned to the counts in the objective function (W = 0.7). In addition, the experimental results show that link counts are more sensitive than speeds to changes in the calibration parameters. The GEH statistics for the models before and after calibration are shown in Table 2. This statistic is included in our analysis because it is recommended by the Traffic Analysis Tool-box [5]. It is

Fig. 1 Pyramid highway, Reno, Nevada, USA (**a**) google map and CORSIM model (**b**) for the first experiment

Fig. 2 Objective function for the first experiment



clear that the calibration model significantly improves the GEH statistic. All the links reach a GEH statistic less or equal to 5, thereby satisfying the calibration criteria. The results show that the three calibration criteria are satisfied. In general, the proposed methodology was able to improve significantly the model outcomes.

Table 2 summarizes the calibration results for the first experiment. The total difference between actual and simulated link counts is 6 % for all links in the network.

A sensitivity analysis was conducted using the Pyramid Highway model. With W = 0.5 and W = 1.0 the difference between simulated and link counts increased significantly.

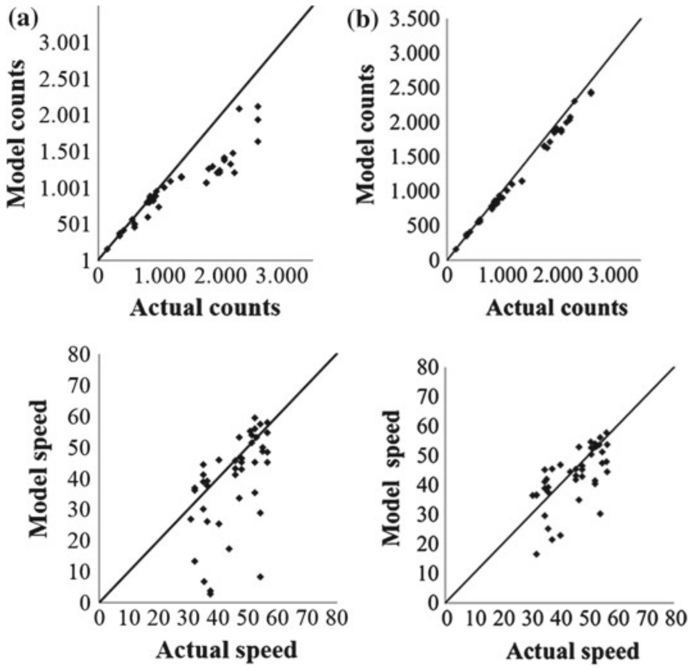


Fig. 3 Actual versus simulated counts and speeds before (**a**) and after (**b**) calibration, for the first experiment

Table 2 Summary of calibration results for the first experiment

	NRMS	Total link counts	GEH
Before calib.	0.042	45,359	<5 for 74 % of the cases
After calib.	0.010	55,882	<5 for 100 % of the cases
Actual		59,610	

Second Experiment: I-75 in Miami, Florida, USA

In this experiment, a portion of I-75 in Miami, Florida was calibrated. A total of 375 freeway ramps and 334 arterial links were included in the model. Data was available for 353 freeway ramps and 59 arterial links for a morning peak period of one hour. The coefficients of the SPSA algorithm were the same as those used in the first experiment. All the calibration parameters in the network were included as well as the turning volumes for freeways and arterials. The weight factor in the objective function was set to 0.7. The time required for calibration was 125 min.

Figure 4a shows the Google map of I-75 highway in Miami, Florida, USA. Figure 4b illustrates the corresponding CORSIM model.

Figure 5 illustrates the trajectory of the objective function for this experiment. The NRMS goes from 0.270 to 0.245.

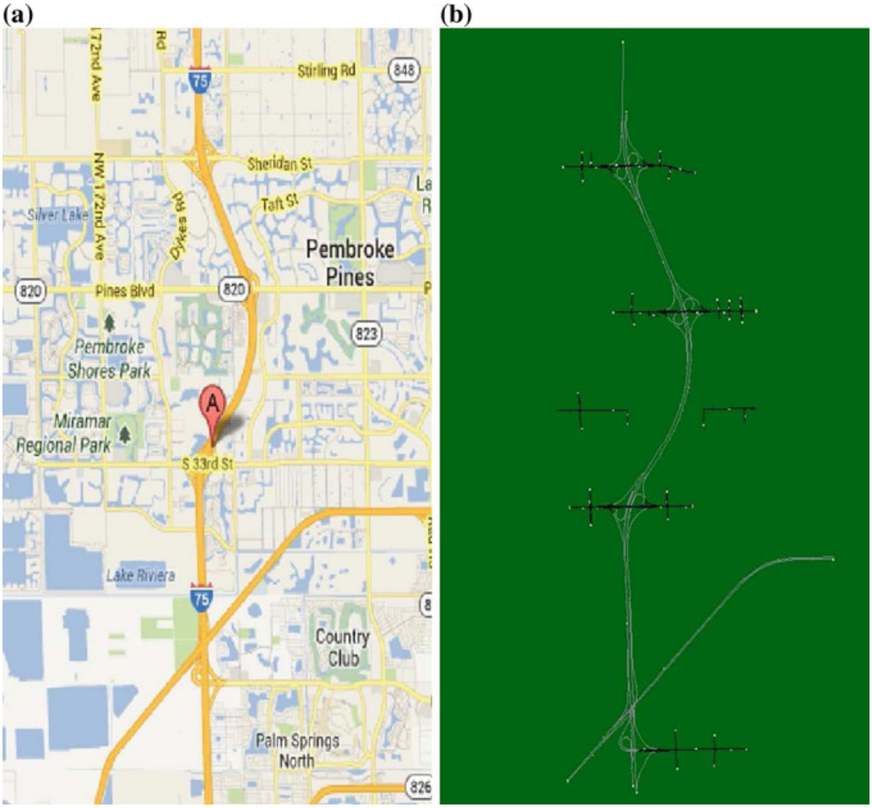


Fig. 4 I-75 in Miami, Florida, USA (**a**) google map and CORSIM model (**b**), for the second experiment

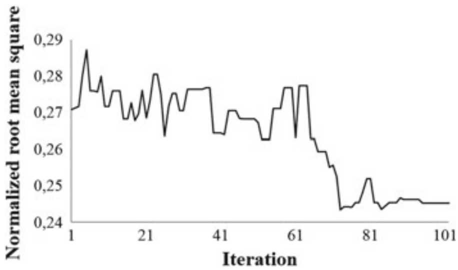


Fig. 5 Objective function for the second experiment

Figure 6a illustrates the link counts for the ramp segments in the model before calibration. Figure 6b shows the link counts for the ramps after calibration. These results clearly show that the calibration process significantly reduces the difference between actual and simulated link counts. It is clear that the calibration model significantly improves the GEH statistic. 99.6 % of the links reach a GEH statistic less or equal to 5, thereby satisfying the calibration criteria.

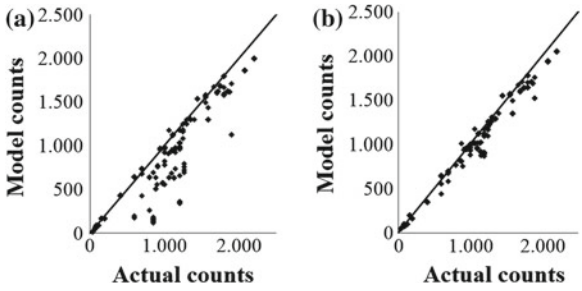


Fig. 6 Links counts before (**a**) and after (**b**) calibration for freeway ramps in the network (second experiment)

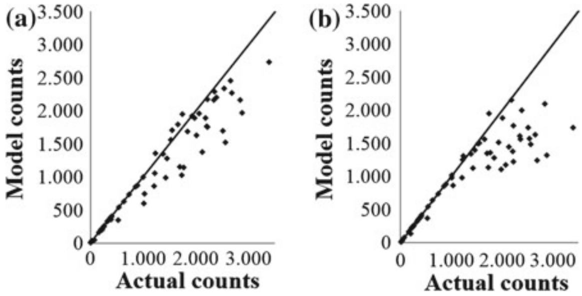


Fig. 7 Links counts before (**a**) and after (**b**) calibration for arterials in the network (second experiment)

Figure 7a illustrates the link counts for the arterials before calibration. Figure 7b shows the link counts for the ramps after calibration. These results show that there is significant improvement for links with large link counts. The calibration model significantly improves the GEH statistic. Seventy-six percent (76 %) of the freeway ramp links reach a GEH statistic less or equal to 5.

Figures 6 and 7 together show that the calibration methodology provides better results for freeway ramps than for arterials. This could be a consequence of having more data available for freeway ramps than for arterials, thereby giving more weight to the ramps.

Table 3 shows the 'before' and 'after' GEH statistics. As illustrated, the calibration improves the statistics, especially for the highest GEHs. However, some GEH values need to be improved because they are over 5.

Table 3 Summary of calibration results for the second experiment

		Total link counts (vph)	GEH
Freeway	Before calib.	234,928.2	<5 for 86 % of the cases
	After calib.	257,454.1	<5 for 99.6 % of the cases
	Actual	271,908	
Arterials	Before calib.	61,097	<5 for 66 % of the cases
	After calib.	68,927	<5 for 76 % of the cases
	Actual	80,524	

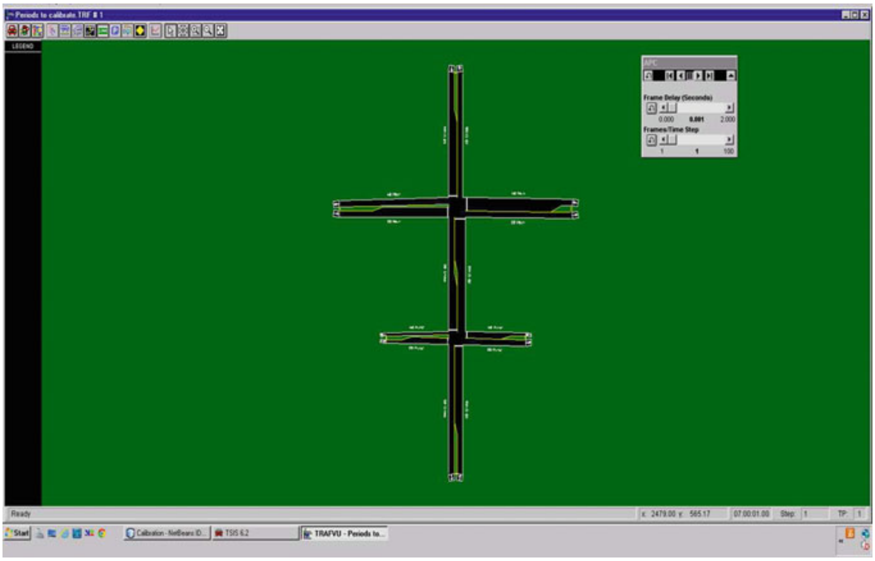


Fig. 8 CORSIM Model for the third experiment: network from McTrans sample datasets

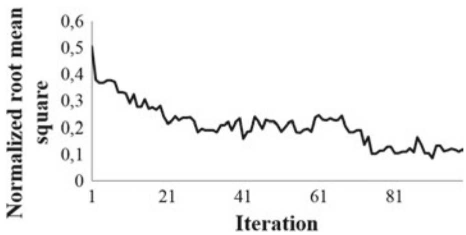
Third Experiment: Network from McTrans Sample Datasets

In this experiment, a network with arterials from McTrans official web page was calibrated. A total of 20 arterial links were included in the model. Data was available for all arterial links. Figure 8 shows the CORSIM model for this experiment. The time required for calibration was 10 min.

The total simulation time was 1 h divided in 4 time periods t of 15 min each (T = 4). In this experiment, all parameters for all links for all four time periods were updated. The coefficients of the SPSA algorithm were the same as those used in the previous experiments. All the calibration parameters in the network as well as the turning volumes were included. The weight factor in the objective function was set to 0.7.

Figure 9 illustrates the trajectory of the objective function corresponding to the third experiment. The initial NRMS value is 0.51, while the minimum obtained after 100 iterations of the optimization algorithm is 0.09.

Fig. 9 Objective function for the third experiment



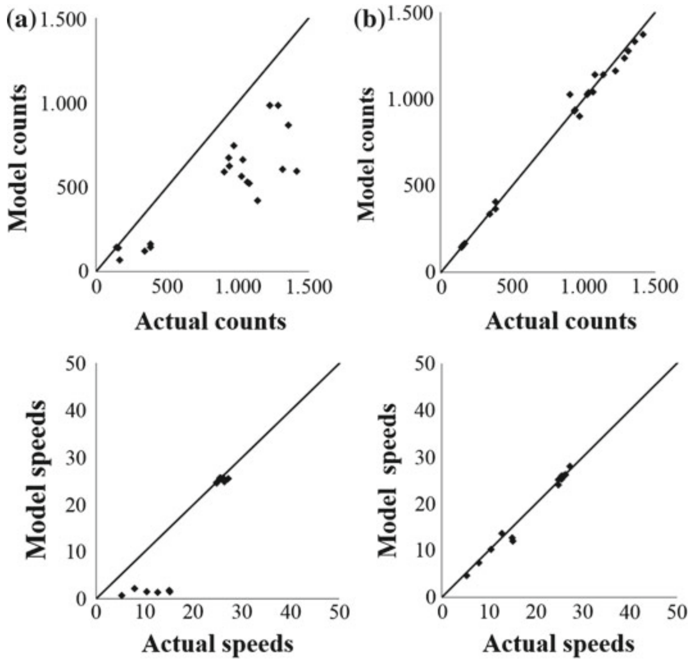


Fig. 10 Actual versus simulated counts and speeds before (**a**) and after (**b**) calibration for time period 1, (third experiment)

Figure 10 illustrates the link counts and speeds before and after the calibration results for all links in the network for the first time period of the simulation. These results clearly show that the calibration process significantly reduces the difference between actual and simulated link counts and speeds.

Similar to Fig. 10, Table 4 shows the summary of link counts and speeds for all links in the network for the second, third, and fourth simulation time period, respectively. The calibrated results are significantly closer to the actual values, relative to the 'before calibration' results. In addition, all links have a GEH statistic below the threshold limit of 5 for all time periods. Speeds were improved for most links especially for values less than 20 mph.

In this experiment, optimal parameters for the model were determined in order to reproduce time-dependent link counts and speeds. The calibrated parameters took a single value during the entire simulation process; that is, they were not time-dependent. In contrast, the link counts and speeds were time-dependent. These results illustrate the ability of the proposed calibration methodology to adjust model parameters so as to calibrate the time-dependent link counts and speeds.

The summary of the results are showed in Table 4.

Table 4 Summary of the calibration results for the third experiment

Goalkeeper	GK	Total link counts (vph)	GEH
Time period 1	Before calib.	10,126	<5 for 10 % of the cases
	After calib.	17,136	<5 for 100 % of the cases
	Actual	17,276	
Time period 2	Before calib.	13,498	<5 for 10 % of the cases
	After calib.	22,625	<5 for 100 % of the cases
	Actual	22,891	
Time period 3	Before calib.	10,502	5 for 0 % of the cases
	After calib.	17,820	<5 for 100 % of the cases
	Actual	18,767	
Time period 4	Before calib.	10,533	<5 for 0 % of the cases
	After calib.	17,939	<5 for 95 % of the cases
	Actual	19,013	

4 Conclusions

This study proposed a methodology for the calibration of micro-simulation traffic flow models. The design and implementation of this methodology seeks to enable the calibration of generalized models. The proposed calibration methodology was developed independent of characteristics for any particular microscopic traffic flow simulation model. It minimizes the difference between actual and simulated time dependent link counts and speeds by considering all model parameters and turning volumes simultaneously.

The methodology used the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm to determine the calibrated set of model parameters. Previous studies have proposed the use of the SPSA algorithm for the calibration of vehicular traffic systems; however, few parameters were considered, and the calibration typically was based on a single performance measure, usually link counts. During the experiments developed, the proposed algorithm always reached convergence and stability.

The proposed methodology was tested using CORSIM models. However, there is nothing preventing the implementation of the proposed methodology for the calibration of other models. Three different vehicular traffic systems were calibrated, taking into consideration all their model parameters by using various performance measures, including link counts and speeds. The first experiment included arterials, using as performance measures link counts and speeds. The second system included both arterials and freeways. Considering arterials and freeways represented a significant challenge because two different models with different parameters needed to be considered simultaneously. The third experiment included time-dependent link counts and speeds for four time periods during this experiment; in addition, global, individual, and time-dependent parameters were considered. Further analysis was required to determine the weight factor, W. This value was set constant because link counts were obtained using more accurate data collection methods compared to speeds. Information about the data collection and data quality can be used to set the weight factor.

The experimental results illustrated the effectiveness of the proposed methodology. The three vehicular traffic systems used in this study were successfully calibrated; specifically, the calibration criteria were satisfied after the calibration was performed. The results from the first and third experiment showed that speeds were improved after the calibration. The quality of the second vehicular traffic system improved significantly. However, further sensitivity analysis of the parameters used by the SPSA algorithm is required to achieve better results and satisfy the calibration criteria. These parameters were chosen using sensitivity analysis. A pattern to find optimal values for the SPSA parameters was not found. Further, as the number of parameters required for calibration increases, the complexity of the optimization problem also increases as well as the complexity to determine the set of required optimization coefficients.

The same set of calibration parameters was used in all the experiments. Therefore, any effort during parameter selection has been reduced. The results were improved for the entire model. All calibrated parameters were within reasonable boundaries. Similarly, no irregularities were observed using the graphical user interface. The calibration software developed in this study can be downloaded, along with a user's guide and examples, using this link: http://faculty.unlv.edu/apaz/files/CalibrationToolDemo.zip. Hence, the reviewers can replicate the results from this study.

The calibration tool developed as part of this study used an optimization algorithm that required a set of coefficients to find the appropriate set of CORSIM model parameters. A time-consuming sensitivity analysis of these coefficients was required to achieve desired results.

A bi-level optimization framework is required to enable the simultaneous calibration of traffic flow and SPSA parameters. The first level of the bi-level framework represents the existing calibration tool developed as part of the existing project, whose objective was the calibration of CORSIM models under saturated conditions. Here, and Simultaneous Perturbation Stochastic Approximation (SPSA) optimization algorithm was used to determine the appropriate calibration parameters. The second level of the proposed bi-level framework corresponds to future research, whose objective is to automate the sensitivity analysis that is required to find the right set of optimization coefficients for the SPSA algorithm. A parallel paper currently under review describes the proposed bi-level framework.

Acknowledgments The authors thank the Nevada Department of Transportation for sponsoring this project. The authors also appreciate Carlos Gaviria's and Cristian Arteaga's help and support during the implementation stage of this study.

References

1. Punzo V, Ciuffo B, Montanino M (2012) Can results of car-following model calibration based on trajectory data be trusted? Transp Res Rec J Transp Res Board 2315(1):11–24
2. Anderson RE, Hicks C (2011) Highlights of contemporary microsimulation. Soc Sci Comput Rev 29(1):3–8
3. Breški D, Cvitanić D, Lovrić I (2006) Sensitivity analysis of the corsim simulation model parameters. Gradevinar 58(7):539–548
4. Brockfeld E, Kühne RD, Wagner P (2005) Calibration and validation of microscopic models of traffic flow. Transp Res Rec J Transp Res Board 1934:179–187
5. Holm P, Tomich D, Sloboden J, Lowrance C (2007) Traffic analysis toolbox volume iv: guidelines for applying corsim microsimulation modeling software. (no. fhwa-hop-07-079). itt industries
6. Kim KO, Rilett L (2003) Simplex-based calibration of traffic microsimulation models with intelligent transportation systems data. Transp Res Rec J Transp Res Board 1855(1):80–89
7. Kondyli A, Soria I, Duret A, Elefteriadou L (2012) Sensitivity analysis of corsim with respect to the process of freeway flow breakdown at bottleneck locations. Simul Model Pract Theory 22:197–206
8. Schultz GG, Rilett L (2004) Analysis of distribution and calibration of car-following sensitivity parameters in microscopic traffic simulation models. Transp Res Rec J Transp Res Board 1876(1):41–51
9. Schultz GG, Rilett LR (2005) Calibration of distributions of commercial motor vehicles in corsim. Transp Res Rec J Transp Res Board 1934(1):246–255
10. Paz A, Molano V, Gaviria C (2012) Calibration of corsim models considering all model parameters simultaneously. In: 15th international IEEE conference on intelligent transportation systems (ITSC), pp 1417–1422. doi:10.1109/ITSC.2012.6338841
11. Balakrishna R, Antoniou C, Ben-Akiva M, Koutsopoulos HN, Wen Y (2007) Calibration of microscopic traffic simulation models: methods and application. Transp Res Rec J Transp Res Board 1999(1):198–207
12. Chin DC (1997) Comparative study of stochastic algorithms for system optimization based on gradient approximations. IEEE Trans Syst Man Cybern B Cybern 27(2):244–249
13. Lee JB (2008) Calibration of traffic simulation models using simultaneous perturbation stochastic approximation (spsa) method extended through bayesian sampling methodology. ProQuest, Ann Arbor
14. Maryak JL, Spall JC (2005) Simultaneous perturbation optimization for efficient image restoration. IEEE Trans Aerosp Electron Syst 41(1):356–361
15. Spall JC (1995) Stochastic version of second-order (Newton-Raphson) optimization using only function measurements. In: Proceedings of simulation conference, pp 347–352
16. Spall JC (1998) Implementation of the simultaneous perturbation algorithm for stochastic optimization. IEEE Trans Aerosp Electronic Syst 34(3):817–823
17. Spall JC (1998) An overview of the simultaneous perturbation method for efficient optimization. Johns Hopkins APL Tech Dig 19(4):482–492
18. Spall JC (2003) Introduction to stochastic search and optimization: estimation, simulation, and control. Wiley, New York
19. Lee JB, Ozbay K (2009) New calibration methodology for microscopic traffic simulation using enhanced simultaneous perturbation stochastic approximation approach. Transp Res Rec J Transp Res Board 2124(1):233–240
20. Ma J, Dong H (2007) Calibration of microsimulation with heuristic optimization methods. Transp Res Rec J Transp Res Board 1999(1):208–217
21. Cunha AL, Bessa Jr JE, Setti JR (2009) Genetic algorithm for the calibration of vehicle performance models of microscopic traffic simulators. In: Progress in artificial intelligence, Springer, pp 3–14

22. Toledo T, Ben-Akiva ME, Darda D, Jha M, Koutsopoulos HN (2004) Calibration of microscopic traffic simulation models with aggregate data. Transp Res Rec J Transp Res Board 1876(1):10–19
23. Henclewood D, Suh W, Rodgers M, Hunter M, Fujimoto R (2012) A case for real-time calibration of data-driven microscopic traffic simulation tools. In: Simulation conference (WSC), proceedings of the 2012 winter, pp 1–12. doi:10.1109/WSC.2012.6465294
24. Rahman M (2013) Application of parameter estimation and calibration method for car-following models
25. Bando M, Hasebe K, Nakayama A, Shibata A, Sugiyama Y (1995) Dynamical model of traffic congestion and numerical simulation. Phys Rev E 51(2):1035
26. Cheu RL, Jin X, Ng KC, Ng YL, Srinivasan D (1998) Calibration of FRESIM for Singapore expressway using genetic algorithm. J Transp Eng 124(6):526–535
27. Holm P, Tomich D, Sloboden J, Lowrance C (2007) Traffic analysis toolbox volume iv: guidelines for applying corsim microsimulation modeling software. Technical Report. http://ops.fhwa.dot.gov/trafficanalysistools/tat_vol4/index.htm
28. Lee DH, Yang X, Chandrasekar P (2001) Parameter calibration for paramics using genetic algorithm. In: 80th annual meeting of the transportation research board, Washington, DC
29. Jha M, Gopalan G, Garms A, Mahanti BP, Toledo T, Ben-Akiva ME (2004) Development and calibration of a large-scale microscopic traffic simulation model. Transp Res Rec J Transp Res Board 1876(1):121–131
30. Li Z, Liu H, Li J (2010) A calibration and validation procedure for microscopic simulation model. In: 13th international ieee conference on intelligent transportation systems (ITSC), pp 563–568. doi:10.1109/ITSC.2010.5625018
31. Quadstone PARAMICS V4.2 (2003) Modeller reference manual. Quadstone Ltd, Edinburgh
32. Weinan H, Jian S (2009) A ngsa-ii based parameter calibration algorithm for traffic microsimulation model. In: International conference on measuring technology and mechatronics automation, 2009. ICMTMA '09, vol 3, pp 436–439. doi:10.1109/ICMTMA.2009.437
33. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: Nsga-ii. IEEE Trans on Evol Comput 6(2):182–197
34. VISSIM 5.10 (2011) User manual. PTV, Karlsruhe, Germany
35. Ma T, Abdulhai B (2002) Genetic algorithm-based optimization approach and generic tool for calibrating traffic microscopic simulation parameters. Transp Res Rec J Transp Res Board 1800(1):6–15
36. Center M (2010) TSIS-CORSIM
37. Omrani R, Kattan L (2013) Simultaneous calibration of microscopic traffic simulation model and estimation of origin/destination (od) flows based on genetic algorithms in a high-performance computer. In: 16th international IEEE conference on intelligent transportation systems—(ITSC), pp 2316–2321. doi:10.1109/ITSC.2013.6728573
38. Wang Z, Koutsopoulos HN (2011) Calibration of urban rail simulation models: a methodology using SPSA algorithm. IEEE, pp 3699–3709
39. Ma X, Huang Z, Koutsopoulos H (2014) Integrated traffic and emission simulation: a model calibration approach using aggregate information. Environ Model Assess 19(4):271–282
40. Cipriani E, Florian M, Mahut M, Nigro M (2011) A gradient approximation approach for adjusting temporal origin destination matrices. Transp Res Part C 19(2):270–282
41. Ben-Akiva M, Gao S, Wei Z, Wen Y (2012) A dynamic traffic assignment model for highly congested urban networks. Transp Res Part C Emerg Technol 24:62–82
42. McTrans (2012) Traffic software integrated system—corridor simulation. http://mctrans.ce.ufl.edu/
43. Hourdakis J, Michalopoulos PG, Kottommannil J (2003) Practical procedure for calibrating microscopic traffic simulation models. Transp Res Rec J Transp Res Board 1852(1):130–139