CALIBRATION OF MICROSCOPIC TRAFFIC MODEL: CROSS-ENTROPY METHOD AND PROBABILISTIC SENSITIVITY ANALYSIS

R. X. Zhong, Ph.D. *
Associate Professor
Research Center of Intelligent Transportation System,
SUN YAT-SEN University, Guangzhou 510006, China
Tel: 86-20-39332772-503
Fax: 86-20-39332775
Email: zhrenxin@mail.sysu.edu.cn

K. Y. Fu, Master student
Research Center of Intelligent Transportation System,
SUN YAT-SEN University, Guangzhou 510006, China
Tel: +86-15876583089
Fax: 86-20-39332775
Email: 649734746@qq.com

D. Ngoduy, EPSRC Career Acceleration Fellow
Institute for Transport Studies, University of Leeds, Leeds LS2 9JT,
United Kingdom
Tel: +44 (0) 1133435345
Email: D.Ngoduy@leeds.ac.uk

A. Sumalee, Professor
Department of Civil Engineering, King Mongkut's Institute of Technology Ladkrabang,
Bangkok, Thailand

William H.K. Lam, Professor
Department of Civil and Environmental Engineering,
The Hong Kong Polytechnic University, Hong Kong SAR, China

* Corresponding author

Word count: text without references (5494) + 4* figure (250) +2 * table (250) = 6994

Submitted for Presentation Only at the 95rd Annual Meeting of the Transportation Research Board
Submission Date: 2015-7-31

ABSTRACT

Calibration and validation techniques pave the way towards the descriptive power of car-following models and their applicability for analyzing traffic flow. However, calibrating these models is never a trivial task because of the existing of unobservable parameters and erroneous traffic data. This contribution puts forward a new calibration framework of car-following models based on the Cross-Entropy Method and Probabilistic Sensitivity Analysis. Cross-Entropy Method is able to identify parameters of car-following models by formulating it as a stochastic optimization problem and to analyse the parameter estimations statistically while Probabilistic Sensitivity Analysis is used to identify the important parameters so as to reduce the complexity, data requirement and computational effort of the calibration process. Empirical results of calibration of intelligent driving model indicate the power of Cross-Entropy Method for searching global optimum for the case of synthetic data and next generation simulation datasets. Furthermore, adopting several termination criteria indicates better property of convergence of CEM than genetic algorithm.

Keywords: Car-following model, Model calibration, Cross-Entropy Method, Probabilistic Sensitivity Analysis

INTRODUCTION

For most planning and operational applications, accurate representation of realistic driving behaviors offers a great help to transportation analysts. Along this stream, microscopic traffic models especially Car Following (CF) models are widely adopted to simulate complex traffic scenarios such as traffic incident, signal control, public transport priority wherein analytical methods are unlikely to work due to the complexity. A large number of CF models have been developed to describe CF behavior under a wide range of traffic conditions in the past decades (1-2). Model calibration would heavily affect the reliability of the results achieved by using the underlying model as well as its applicability in traffic engineering practice. Calibration is vital for microscopic traffic models, yet it can be rather difficult since these models often contain a wide range of variables. Moreover, some parameters, such as reaction time, desired spacing, desired time headway, desired speed, are generally not directly observable from traffic data, and this makes them hard to be identified. Nonetheless, the parameters are scenario specific, i.e. they are not transferable to other situations (different locations, periods of the day such as morning rush hours and evening rush hours, etc.). On the other hand, driving behavior and local traffic rules, which the microscopic traffic flow models intend to describe, are variable in time and space (3-4). Therefore, many of the microscopic traffic models were neither empirically calibrated nor validated using real traffic data until recently, a considerable amount of research have been devoted to the calibration and validation of microscopic models (2, 4, 5). However, guidance on the systematic and rigorous calibration and validation of traffic flow models is still lacking (2, 6). Nevertheless, traffic data might subject to various errors and noises. The calibration of stochastic microscopic models is more complicated as they will require multiple runs to reduce the noise in the objective function (6). To tackle these difficulties, this paper aims to develop a new framework for calibrating microscopic traffic models with various uncertainties to maximize the model's descriptive power based on representative traffic data.

Conventionally, deterministic search methods which aim to minimize the discrepancy between the model prediction and observed data are common approaches to access model calibration for both microscopic and macroscopic traffic models. As a consensus in the literature, such kind of methods will result in a large number of local optima due to different combinations of the set of parameters (6-8). Therefore, a popular and convenient approach to compensate this is to use random search techniques. The basic idea behind such methods is to systematically partition the feasible region into smaller subregions and then to move from one subregion to another based on information obtained by random search (6-7). Well-known examples include simulated annealing, genetic algorithm, tabu search, and ant colony methods. All these methods are reported to find a good local optimal solution (while some also claimed global optimal solution can be obtained) but there is not as yet a fully accepted method. Noticing the optimization nature of the calibration problem, Ciuffo and Punzo (7) applied the no free lunch theorems to the calibration problem of microscopic traffic models to access the performance of various algorithms ranging from heuristic optimization methods to meta heuristic searching methods. Remarkably, the analysis reveals that the Genetic Algorithm (GA), which is probably the most widely used algorithm type for the calibration of microscopic traffic simulation models, outperforms the others globally for the tested cases in (7).

However, it is known that the GA is generally computationally intensive and of no convergence proof. Moreover, the presence of random noise would affect the optimization procedures. Ngoduy and Maher (6) applied the Cross Entropy Method (CEM), which is a generic Monte Carlo technique with importance sampling for reducing the computational burden, to the calibration purposes of a second order macroscopic model. The empirical results have verified several merits of such method including attaining globally optimal solution, computationally efficient and convergent. Furthermore, Maher et al. (9) extended this CEM framework for signal optimization to consider the effect of noise in the evaluation process.

Because of the unobservable parameters of the CF models, inherent noise of traffic data, complicated human factors to be modeled and traffic scenarios to be simulated, one of the main challenges arising in the calibration process concerns the selection of the most important parameters and the identification of their probabilistic characteristics. Selecting the most important parameters also helps guiding data collection by limiting the number of input parameters to be observed rather than the whole parameter set (2). To avoid the potential negative effects caused by subjectively chosen set of calibration parameters, a Sensitivity Analysis (SA) of the parameters is essentially required. SA is a procedure to explore the relationship between the simulation output and input parameters considering the potential uncertainty (including noise, un-modeled characteristics and parameter variations etc.) on model response. SA is also a crucial procedure to individuate the most important sources of modeling uncertainties. Despite of its importance, a proper SA for traffic models is barely performed in common practice (10-11).

One commonly adopted SA approach is the ANalysis Of Variance (ANOVA) which is able to quantify the uncertainty of a model and to capture the interactions among different factors (12). A more efficient SA method based on the decomposition of the variance has been adopted to the SA of CF models (13). However, the ANOVA method is not suitable to measure the dispersion of a variable with a heavy-tail or a multimodal distribution which may be the case of calibration of CF models (as it will be shown later in Empirical Study). Furthermore, this method is computational burden which often requires a large number of model evaluations.

To combine the calibration and SA into a unified framework, this paper puts forth a Cross-Entropy Method (CEM) (14-16) and the Probabilistic Sensitivity Analysis (PSA) based approach for model calibration and identification of important parameters. The CEM is able to find a global optimal solution to the calibration problem while the PSA reveals the probabilistic behavior of a model's response with respect to its parameter uncertainties. Results from PSA can be used to assist engineering design from various aspects, such as to help reducing the dimension of a design problem, to investigate potential improvements on a probabilistic response (17). Since this method is based on the relative entropy, it better fits into the framework of CEM. To be more specific, the PSA can be conducted in conjunction with the importance sampling technique of the CEM to reduce the computational burden of the optimization whilst identifying the important parameters and the effect of noise.

CAR-FOLLOWING MODELS

The Intelligent Driving Model (IDM) proposed by Treiber et al. (18) is one of the widely-applied CF models. This model considers both the desired speed and the desired space headway to model the acceleration/deceleration strategy of subject vehicle with respect to the preceding vehicle,

$$\dot{v}_{IDM}(v_f, s, s^*(v_f, \Delta v)) = a[1-(\frac{v_f}{v_0})^4 - (\frac{s^*(v_f, \Delta v)}{s})^2] \tag{1}$$

where a is the maximum acceleration/deceleration of the subject vehicle, v_0 the desired velocity, v_f the velocity of following vehicle, s the spacing between two vehicles measured from the front edge of the subject vehicle to the rear end of the preceding vehicle, and $s^*(\cdot)$ the desired spacing.

When preceding vehicle is far away, the third term in this equation becomes negligible small and the model performs as a free flow model where the desired speed of the driver governs the acceleration, i.e., $\dot{v}_{free} = a[1-(v_f/v_0)^4]$. When the subject vehicle approaches the preceding vehicle, the braking strategy $\dot{v}_{brake}(s, v_f, \Delta v) = -a(s^*/s)^2$ will be dominant to ensure s to approach the desired minimum gap s^*, which depends on several factors: speed, speed difference (Δv), the maximum acceleration a, a comfortable deceleration (or the desired deceleration which will be active in non-stationary traffic) b, the minimum spacing at the standstill situation s_0, and the desired time gap T. To be specific,

$$s^*(v_f, \Delta v) = s_0 + v_f T - \frac{v_f \Delta v}{2\sqrt{ab}} \tag{2}$$

Remarkably, the deceleration of IDM will be quite strong if the current gap becomes too small to ensure collision-freeness. Reaction time is ignored in this model.

It is worth noticing that parameters of IDM are related to different traffic conditions. For example, the maximum acceleration a and the desired deceleration b are related to stop-and-go traffic flow while the desired time gap T is mainly involved in the steady-state car-following period, the desired velocity v_0 is observed in free-flow traffic condition while the creeping and standing traffic situation is crucial for the identification of the minimum distance s_0. In other words, the data sources used to calibrate IDM are suggested to contain all traffic regimes mentioned above.

Some parameters of the CF models, such as the desired time gap, desired spacing, comfortable deceleration etc., are subject to human factors, which mainly features in the inter-driver and intra-driver heterogeneity. Generally, the inter-driver heterogeneity implies that different drivers behave in different ways even if they follow the same vehicle due to individual driving habits, while the intra-driver heterogeneity indicates that a driver shows differential response to the same change of driving situation at different time or under different conditions (20). Inter-driver heterogeneity is easily explained by the traffic oscillations caused by the aggressive and timid driver behaviors (21). The intra-driver variability accounts for a large part of the deviations between simulations and empirical observations (19). Several recent studies (2, 22, 23) argued that the parameters of CF models differ from drivers because drivers are different so as their driving styles and risk-taking capabilities. All these suggest that the randomness in the CF models can be interpreted as uncertainty in the parameters (24). Furthermore, the desired parameters such as desired time headway, desired velocity are generally unobservable in nature, which renders the parameter estimation problem more challenging (2). Therefore, it is necessary to consider such randomness in the calibration process.

THE CROSS-ENTROPY METHOD

As discussed in Introduction, a cross-entropy method based approach is proposed to solve the calibration problem. Simply speaking, the CEM approach can be broken down into two key steps:

1. Generate a number of trial parameter sets randomly according to a chosen distributions.
2. Based on the values of the objective function associated with each trial parameter set, update the probability distribution used to generate the random trial sets according to the principle of 'importance sampling'.

For a general optimization problem, obtaining a (global) optimum solution can be regarded as a rare event. The CEM is a general Monte Carlo approach to solve rare event probability estimation problems. In this sense, we can reformulate an optimization problem in terms of cross entropy method as follows. Without loss of generality, consider the following minimization problem:

$$\gamma^* = \min_{x \in \chi} S(x) \tag{3}$$

where γ^* represents the minimum of $S(x)$ and x is defined in a function space χ. By the above analogy, the cross-entropy method may first formulate a family of Probability Density Functions (PDF) distributed in χ, denoted by $f(x;v)$, parameterized by v. For the minimization of $S(x)$, by defining the minimum γ^* as a threshold (or some $\gamma \geq \gamma^*$ but sufficient close to γ^*), we can define a rare event as $S(x) \leq \gamma$. To this end, we can define

$$\ell(\gamma) = P_u(S(X) \leq \gamma) = E_u(I_{\{S(X) \leq \gamma\}}) \tag{4}$$

where γ is a threshold and $X = (X_1, X_2, ..., X_n)$ is a random vector generated by PDF with parameter v set to u (i.e., a realization) in $f(x;v)$. P_u denotes the probability, E_u denotes the expectation, and $I(\cdot)$ is the indicator function, i.e., $I_{\{S(X) \leq \gamma\}} = 1$ if and only if $S(X) \leq \gamma$ is true, 0 otherwise. By this, we convert the original optimization problem into a rare event probability estimation problem. Detailed routine for solving this estimation problem by the CEM can be found in the cross entropy tutorial (15).

To apply the CEM optimization algorithm to the calibration problem, it is necessary to apply the time-discretization scheme to convert the infinite-dimensional functional optimization problem into a finite-dimensional parametric optimization problem. A simple way to achieve this is to divide the simulation time span into $\xi - 1$ subintervals. The parametric optimization problem yielded belongs to continuous multi-extremal optimization problems, i.e. each decision variable is real valued and the decision vector belongs to a subset of Euclidean space. A CEM approach based on kernel density estimation method is proposed by Ngoduy and Maher (6) to solve such optimization problems. However, the drawback of such kernel based CEM is its high computational demand. On the other hand, noting that the sampling distribution of the CEM can be quite arbitrary, and does not need to be related to the function that is being optimized (25). Adopting Gaussian (mixture) distributions is convenient and can reduce the computational effort. The relevant pseudo code for such modified CEM is summarized in the next second section.

PROBABILISTIC SENSITIVITY ANALYSIS

Noting that the relative entropy (also known as Kullback-Leibler (K-L) distance) is calculated directly from the probability distribution function and thus provides a more general measure of output variability. For discrete probability distributions p_k and q_k, the K-L distance of q from p is defined to be

$$D_{KL}(p \| q) = \sum_{k=1}^{K} p_k \ln \frac{p_k}{q_k} \tag{5}$$

Based on the relative entropy, Liu et al. (17) proposed a new Probabilistic Sensitivity Analysis (PSA) approach to evaluate the impact of a random variable on the performance index function by measuring the K-L distance between two probability density functions of the performance index function, obtained before and after the variation reduction of the chosen random variable. This method has been recently extended to study the SA of complex stochastic processes (26-27).

Assume the model of interest admits a mapping of the following form

$$Y = h(X_1, X_2, ..., X_n) \tag{6}$$

where $X_1, X_2, ..., X_n$ are the random inputs that refer to intrinsic model parameters of the CF models in this paper. We assume that the model output has a PDF of $f(y)$. Global sensitivity analysis aims to rank the inputs $X_1, X_2, ..., X_n$ according to the degree to which they influence the output, individually and conjointly. If one particular parameter X_i is fixed to its nominal value (or its mean value or one of its realizations/observations) say replacing it with $\bar{x}_i$, then the yielded PDF of Y is denoted as $g(y|\bar{x}_i)$. From the above paragraph, the K-L distance $D(g(y|\bar{x}_i) \| f(y))$ measures the difference between the divergence between two probability density functions of the output obtained before and after the variation reduction of the random variable X_i. Specifically,

$$D_{KL_{X_i}}(g(y|\bar{x}_i) \| f(y)) = \sum g(y|\bar{x}_i) \ln \frac{g(y|\bar{x}_i)}{f(y)} \tag{7}$$

for a discrete case. This equation quantifies the change of the PDF of the model output after eliminating the variability in X_i. Therefore, the K-L distance $D_{KL_{X_i}}(g(y|\bar{x}_i) \| f(y))$ is regarded as the total sensitivity index of X_i. The larger value $D_{KL_{X_i}}(g(y|\bar{x}_i) \| f(y))$ is, the more important X_i is (which implies that Y is more sensitive to X_i). More details and applications of PSA can be found in (17).

THE PSEUDO CODE OF CEM AND PSA

Note that the PSA requires PDF estimations which can be achieved by the Kernel density estimation method. To this end, we combine the CEM for continuous optimization in terms of Normal updating by (25) and the Kernel-based updating by Ngoduy and Maher (6) to reduce computational burden, which is summarized as follows:

Step 1. Discretize the simulation time span into $\xi - 1$ equal sub-intervals $[\tau_1, \tau_2], ..., [\tau_{\xi-1}, \tau_\xi]$.

Step 2. Set $t = 0$, and choose $\hat{u}_0$ and $\hat{\sigma}_0^2$. In empirical study we choose $\hat{u}_0 = [3,3,20,2.5,4]^T$ and $\hat{\sigma}_0^2 = [100,100,100,100,100]^T$, with each element corresponds to the maximum acceleration a, desired deceleration b, desired speed v_0, desired time gap T, and minimum gap s_0 of IDM, respectively.

Step 3. Increase t by 1. Generate a set of random samples $X_1, ..., X_N$ from $\mathbb{N}(\hat{u}_{t-1}, \hat{\sigma}_{t-1}^2)$ distribution, which is an n−dimensional normal distribution with independent components. N is set to 1000 in our cases.

Step 3.1. For a given sample of the parameters, solve the IDM and calculate the objective function $S(X_i)$ for each sample.

Step 3.2. Order $S(X_i)$ from smallest to the largest as $S_{(1)} \le S_{(2)} \le ... \le S_{(N)}$ and finally evaluate the $(1-\rho)100\%$ sample percentile $\hat{\gamma}_t$ of the sample scores,

$$\hat{\gamma}_t = S_{(\lceil (1-\rho)N \rceil)} \tag{8}$$

where ρ is a small real number and set to 0.01 here.

Step 3.3. Let Γ be the indices of the N^{elite} best performing samples.

Step 3.4. For all $j = 1, 2, ..., n$, let

$$\tilde{u}_{t,j} = \sum_{i\in\Gamma} X_{i,j} \big/ N^{elite} \tag{9}$$

$$\tilde{\sigma}^2_{t,j} = \sum_{i\in\Gamma} (X_{i,j} - \tilde{u}_{t,j})^2 \big/ N^{elite} \tag{10}$$

Step 4. Smooth: $\hat{u}_t = \beta\tilde{u}_t + (1-\beta)\hat{u}_{t-1}, \hat{\sigma}_t = \beta\tilde{\sigma}_t + (1-\beta)\hat{\sigma}_{t-1}$. Here β is set to 0.7.

Step 5. If $\max_j \{\hat{\sigma}_{t,j}\} < \varepsilon$ stop and return $u = \hat{u}_t$ (or the overall best solution generated by the algorithm) as the approximate solution to the optimization, where $\varepsilon \in R_+$ is a preset tolerance and set to 10^{-6} here. Otherwise, return to Step 3.

Step 6. PSA: Generate a new set of random samples X from $\mathbb{N}(u, \hat{\sigma}_0)$ distribution and regard it as original set of random samples.

Step 6.1. Fix one parameter x_j to its optimal solution u_j to produce conditional samples, i.e. $X | x_j$, $j = 1, 2, ...5$ for the IDM.

Step 6.2. For original and conditional sets of samples $X, X | x_j$, $j = 1, 2, ..., 5$, simulate the IDM, then calculate the objective function $S(X_i)$ accordingly. Then determine the PDFs $f(y)$ and $g(y | \bar{x}_i)$ using Kernel density estimation method.

Step 6.3. Calculate the corresponding sensitivities by solving Equations (7) and return.

Constrains on each parameters of IDM are adopted into CEM to ensure a plausible optimum solution. In our case, the feasible region $\bar{\chi}$ includes $a \in [0.1, 6](m/s^2)$, $b \in [0.1, 6](m/s^2)$, $v_0 \in [0.1, 35](m/s)$, $T \in [0.1, 5](s)$, $s_0 \in [0.1, 8](m)$.

EMPIRICAL STUDY

Data Processing

Recent research has found that the trajectory data derived from the Next Generation SIMulation (NGSIM) project contains kinds of errors ranging from measurement to traffic flow characteristics (28). A multi-step procedure proposed by Punzo et al. (28) is adopted to eliminate possible biased trajectories in the I80 dataset which is a 15 minutes time frame observation (4:00 pm to 4:15 pm, on April 13, 2005) on a stretch of Interstate 80 in San Francisco, California. Here, we focus on the car-following behaviors of the middle lane, which is termed lane 3 in I80 dataset and hereafter, in order to maximize the fraction of pure car-following situations.

Calibration through Synthetic Data

As suggested in Ossen and Hoogendoorn (29), to assess how well parameters can be identified by the calibration procedure, the ‘ground truths’ of these parameters must be known since parameters minimizing the objective function do not necessarily capture following dynamics best. To show the proposed method can actually find the global optimum, a set of synthetic data is used. The synthetic data also help verifying whether the data contain enough information to estimate the parameters of interest which helps in analyzing the cause for deviations from the real parameters when performing calibration with real data. Calibrating the synthetic data can also provide guidance on choosing suitable variables in the objective function for better model calibration.

The synthetic data is obtained from simulating the IDM with trajectory of the leading vehicle and a set of parameters. To ensure that the synthetic data resemble real data as much as possible, the data of leading vehicle is derived from I80 dataset randomly. Here a parameter set $x^* = [1.5, 0.8, 20, 1.25, 4.5]^T$, with each element is the same as $\hat{u}_0$, is given as the ground truth. With this set of synthetic data and the same profile for the leading vehicle, calibration is carried out to find out the ability of CEM for searching the optimal parameters from this synthetic data set. Motivated by the empirical studies by Ossen and Hoogendoorn (29); Paz et al. (30) that including gap and speed in objective function would improve the calibration result. Thus, the authors adopted the following Combined Objective Function (COF):

$$\hat{x}^* = \arg\min_{x \in \bar{\chi}} \lambda \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(s_t^{cali}(x) - s_t^{data})^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}s_t^{cali}(x)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(s_t^{data})^2}} + (1-\lambda)\frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(v_t^{cali}(x) - v_t^{data})^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}v_t^{cali}(x)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(v_t^{data})^2}} \quad (11)$$

where $s_t^{cali}(x)$ and s_t^{data} denote the calibrated and observed gap of following vehicle towards preceding vehicle at time t, and $v_t^{cali}(x)$ and v_t^{data} denote the calibrated speed derived from simulation of IDM and observed speed, respectively. $0 \le \lambda \le 1$ is the weighting coefficient.

By varying λ from 1 to 0, the difference between calibrated parameters and their 'ground truth' counter-parts are present in Table 1.

TABLE 1 The Solutions and Corresponding Relative Errors (RE) with respect to Different Weighting Coefficients

	Parameters	$a(m/s^2)$	$b(m/s^2)$	$v_0(m/s)$	$T(s)$	$s_0(m)$	OV[a]
$\lambda=1$	Solution	1.43	0.89	18.55	1.29	4.31	1.34E-5
	RE[b]	4.67%	11.25%	7.25%	3.20%	4.22%	NV[c]
$\lambda=0.8$	Solution	1.43	0.86	19.14	1.29	4.31	1.41E-5
	RE	4.67%	7.50%	4.30%	3.20%	4.22%	NV
$\lambda=0.5$	Solution	1.45	0.84	19.97	1.29	4.33	1.28E-5
	RE	3.33%	5.00%	0.15%	3.20%	3.78%	NV
$\lambda=0.2$	Solution	1.47	0.81	20.47	1.28	4.38	7.81E-6
	RE	2.00%	1.25%	2.35%	2.40%	2.67%	NV
$\lambda=0.1$	Solution	1.48	0.81	20.39	1.27	4.42	4.82E-6
	RE	1.33%	1.25%	1.95%	1.60%	1.78%	NV
$\lambda=0.01$	Solution	1.50	0.80	20.06	1.25	4.49	6.53E-7
	RE	0.00%	0.00%	0.00%	0.00%	0.00%	NV
$\lambda=0.001$	Solution	1.50	0.80	20.00	1.25	4.50	6.82E-8
	RE	0.00%	0.00%	0.00 %	0.00%	0.00%	NV

[a]: The Value of the Objective function

[b]: $RE = |\hat{x}^* - x^*| / x^*$

[c]: Not a Value

As observed, increasing the weighting of the speed fit in objective function, i.e. decreasing λ, yields more accurate parameter estimation. Particularly, when the objective function contains almost only speed fit, the calibrated parameters are almost identical to their ground truth. The reason may be that when the solution is approaching closely to the optimum, the space gap is sufficiently small and/or stationary, i.e. the effect of the gap fit part is very small. Under this situation, looking into its first order derivative, i.e. the speed gap, would help moving the solution towards to the optimum. Moreover, the speed evolution of the IDM depends on the following vehicle only while the derivative of the space gap is relative to the difference between the speeds of the preceding and the following vehicles, i.e.,

$$\hat{s}_t = (v_t^p - \hat{v}_t) * \Delta t + \hat{s}_{t-1} \tag{12}$$

where $\Delta t = 0.1$ is the sample interval of I80 dataset. As the speed profile of the proceeding vehicle is given, the space gap will decrease by reducing the speed gap given that there is no offset which can be achieved easily by initializing the IDM with the corresponding measured location.

Calibration with Actual Single Trajectory

The goal of this test is to reproduce the chosen actual single trajectory data by identifying the five parameters of the IDM using cross entropy method whilst to evaluate the impact of each parameter on objective function by probabilistic sensitivity analysis. To this end, the authors randomly select a pair of following-leading vehicle from I80 dataset and set λ to be 1, 0.5, 0.01, 0 respectively in the COF to assess the performance of the proposed calibration method.

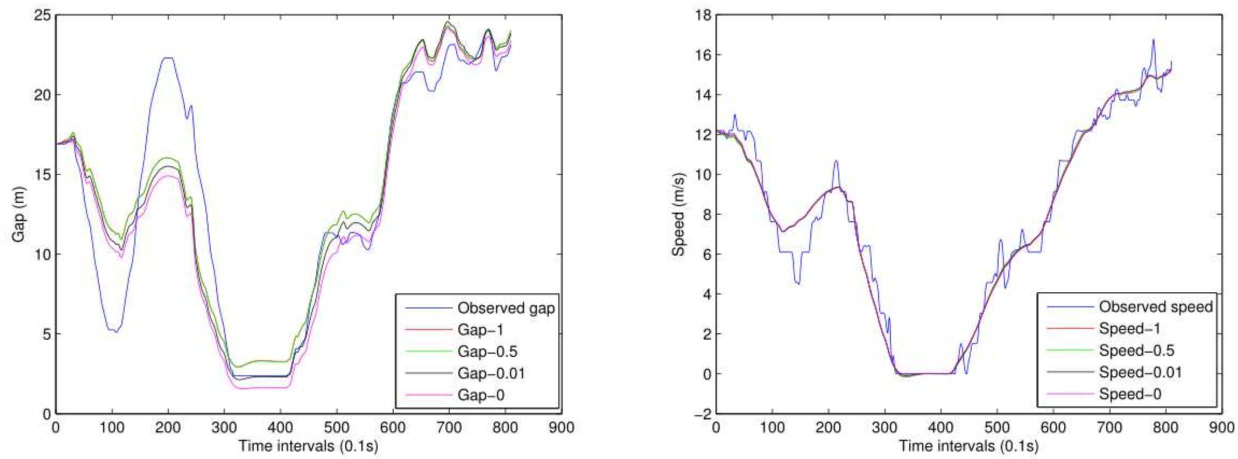
With the default settings of CEM algorithm, the calibration results are presented in Table 2.

TABLE 2 Solutions to IDM for Actual Dataset with respect to Different Value of λ

Parameters	$a(m/s^2)$	$b(m/s^2)$	$v_0(m/s)$	$T(s)$	$s_0(m)$	OV
$\lambda = 1$	1.20	6	35	1.25	3.28	0.0170
$\lambda = 0.5$	1.19	6	35	1.25	3.25	0.0105
$\lambda = 0.01$	1.07	6	35	1.29	2.31	0.0042
$\lambda = 0$	1.04	6	35	1.31	1.62	0.0040
GA	1.07	6	35	1.29	2.31	0.0042

The optimum solutions of four cases seem plausible except the same high values of desired speed $v_0 = 35m/s(126km/h)$ and desired deceleration $b = 6m/s^2$, which indeed hit the specified boundaries. This may be because that this actual single trajectory is not sufficient for calibrating all parameters which are relevant to different traffic regimes rather than the detailed structure of the objective function. To be specific, this single trajectory contains neither free flow nor approaching traffic regime which is necessary to calibrate desired speed and desired deceleration, respectively. Treiber et al. (4) regarded this phenomenon as data incompleteness. As a consensus in the literature of calibration of microscopic models, the choice of suitable inputs (inputs have to sufficiently excite the system in order to be informative) is critical.

For a more visible comparison, the simulated trajectory is plotted against the measured one in terms of gap and speed measures in Figure 1.



(a) (b)

FIGURE 1 Comparison of three calibrated and observed data for (a) gap and (b) speed ('Gap-1' means the case of $\lambda = 1$).

The high consistency between the calibrated and observed measured of four cases indicates that the calibration results by the CEM is plausible. Remarkably, a better identification of standing traffic situation during time intervals from 300 to 400 is observed when $\lambda = 0.01$ rather than $\lambda = 0$. Although the global optimum can be found with synthetic data when there is only speed fit in objective function, i.e. $\lambda = 0$, a small weighting of the space gap is more suitable for calibration with actual trajectory data in this case study. The reason may be that, in contrast with the synthetic data, actual data such as NGSIM contains kinds of errors which renders considering speed fit only in the objective function is unsuitable for calibration (4, 29).

To show the convergence of the CEM, the evolution of the PDF of the desired time gap for the case of $\lambda = 0.01$ is shown in Figure 2.

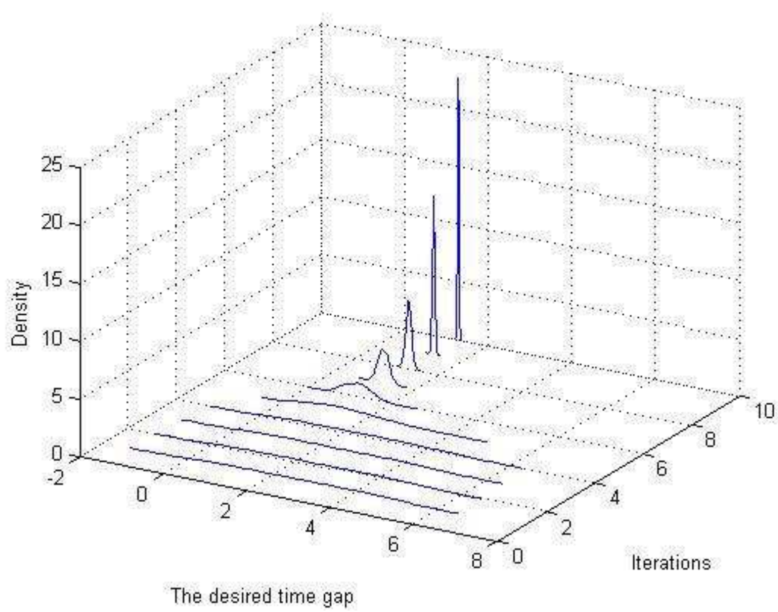


FIGURE 2 PDF estimation of T over first 10 iterations for $\lambda = 0.01$.

The change on the horizontal axis values of the PDF over iterations indicates that the PDF of the desired time gap becomes more and more concentrated around the optimal solution over iterations, and the optimal parameter is found when the standard deviations of this distribution is approximately zero (i.e. the density function is spiked).

Probabilistic sensitivity analysis is conducted to provide an intuitive view of data incompleteness and identify the importance of certain parameters. For the case of $\lambda = 0.01$, the PDF by fixing certain parameter x_i to its optimal value while the other four parameters vary in pre-specified ranges, which is denoted as $g(y|\overline{x}_i)$ previously, is presented in Figure 3. Note that $g(y|x_2)$, where x_2 is the desired deceleration, is almost the same as $f(y)$ which implies that the variation of other parameters except the desired deceleration contributes almost the total uncertainty to the objective function. As a result, the desired deceleration is identified as an insignificant parameter under this single trajectory case. In contrast, the PDF of the objective function by fixing the desired time gap T presents a sharp spike. This distinct difference between the original PDF and $g(y|x_4)$ indicates that the variance of objective function concentrates around the range of 0 when desired time gap T is fixing to its optimum. That is to say the variance of desired time gap contributes a lot to the uncertainty of the objective function under this single trajectory. Moreover, the low impact of certain parameters on objective function may be regarded as an index to data incompleteness such as the desired deceleration case.

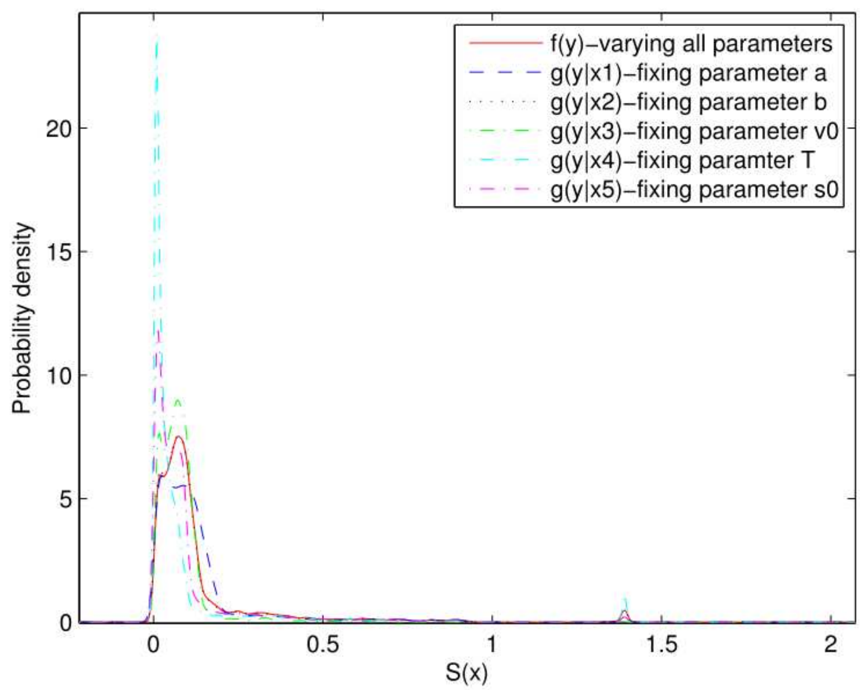


FIGURE 3 The results of PSA.

The corresponding discrepancy of each $g(y|\overline{x}_i)$ to $f(y)$ is estimated by K-L distance, with results of $[28, 2, 21, 84, 38]^T$ for $i = 1, 2, ..., 5$. On the one hand, small values of K-L distance of the desired deceleration and desired speed indicate that these two parameters are relevant insignificant inputs to objective function under this single trajectory, which may also correspond to the conclusion of data incompleteness of this single trajectory. On the other hand, the desired time

T and minimum gap s_0, which are related to safe distance of individual driver in the IDM, are regarded as the most significant parameters beyond other parameters as indicated by their large values of K-L distance. Their variation contributes a lot to the fluctuation in the difference between calibrated and observed gap, which is consistent with the finding of Kesting et al. (19).

Performance Assessment of the CEM and GA for Calibration

In microscopic calibration literature, Genetic Algorithm (GA) is a widely adopted algorithm, which is also regarded as the benchmark solution algorithm, to solve the optimization problem (10, 31). The GA is a stochastic-based global optimization method which mimics the natural biological evolution mechanisms, including selection and reproduction, cross over, and mutation mechanisms. GA is also one of the most popular optimization algorithms because not only any derivative information for optimization but also differentiable property for objective function are not required during process. With 1000 individuals and 100 maximum iterations, the optimal solution to the calibration problem of the IDM with single trajectory with respect to the case of $\lambda = 0.01$ using GA is shown in the bottom of Table 2. As observed, the parameters calibrated by the CEM and the GA are almost the same. However, their convergence properties are quite different.

To explore the convergence of CEM and GA, two commonly used termination criteria are adopted:

1. Standard deviation of objective function S of current iteration k is less than ε:

$$\mathrm{TermCond}_1 : \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_{ki} - \bar{S}_k)^2} \le \varepsilon \tag{13}$$

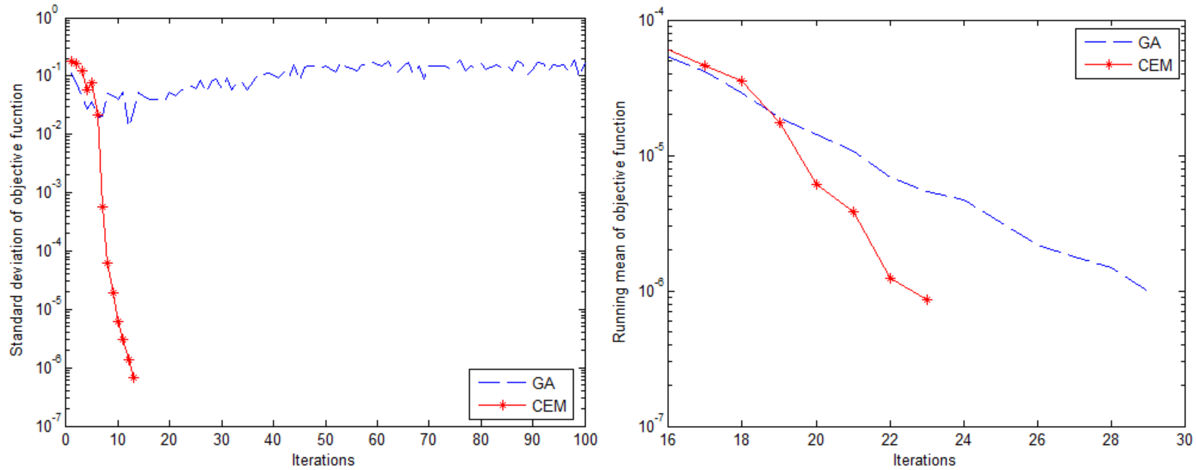
2. The difference between the current optimal value of the objective function S_k^{best} and the average of the optimal values of the objective function of last (several) iterations t_{last} is less than ε, which is also termed as running mean:

$$\mathrm{TermCond}_2 : \| S_k^{best} - \frac{1}{t_{last}} \sum_{i=k-t_{last}+1}^{k} S_i^{best} \| \le \varepsilon \tag{14}$$

In the following cases, the above two termination criteria are used to calibrate the single trajectory shown in Figure 3, with respect to COF with $\lambda = 0.01$, to investigate different convergent properties of these two algorithms.

As indicated in Figure 4 (a) that in terms of the standard deviation convergence criterion the CEM converges within 13 iterations while the GA seems to fail to converge. That means the CEM also converges within several iterations in terms of the mean objective value. In contrast, the standard deviation of GA decreases in the first 20 iterations but trends to fluctuate in a large scale later on. This may be due to that the mutation mechanism of GA then tries to change certain property of several samples to explore new potential best samples but fails. This failure results in several extremely high values of the objective function. As a consequence, the standard deviation of objective function also oscillates and cannot meet the standard deviation convergence criterion.

As for the second termination criterion, Figure 4 (b) depicts that both the running means of the GA and CEM are decreasing and less than ε after finite iterations. Therefore, the running mean termination criterion is fulfilled. The CEM takes less iterations to converge than the GA does. In terms of computational time, the CEM requires 11 seconds to converge while the GA takes 15 seconds. For a multi-trajectories case including 111 pairs of trajectories, the CEM takes 727 seconds to converge while the GA needs 1055 seconds, indicating a better performance of the CEM than the GA.



(a) (b)

FIGURE 4 The convergence property of CEM and GA with respect to termination criteria of (a) 'Standard deviation' and (b) 'Running mean'.

CONCLUSIONS

This paper proposed a new framework based on CEM and PSA to calibrate CF models with uncertainties such as noisy data and heterogeneous driver behaviors. Converting the calibration problem into a rare event detection problem, the CEM based optimization method is proposed to identify the parameters of car-following models. To identify important parameters, PSA is applied to investigate the probabilistic behavior of the model output response with respect to its parameter uncertainties. The PSA can handle the case that the PDF of interest is nonlinear and/or non-Gaussian, and not only in a few moments, due to its heavy tail. Furthermore, since the CEM and PSA are both based on the K-L distance, they can be simultaneously integrated into a unified framework to reduce the computational burden.

Several empirical studies were conducted to illustrate the performance of the proposed calibration framework. Firstly, a synthetic data is used to show the effectiveness of the CEM to search a global optimum. Furthermore, the results suggested to find a better structure of COF for better performance of certain CF models to avoid subjectively decision on the variables in objective function. Secondly, in terms of actual trajectory dataset, the unrealistic value of certain parameter, which may be caused by insufficient information, has been quantified by the low value of the relative entropy through the lens of PSA. For IDM, the low impact of the desired velocity may be because that most of the drivers would expect more or less the same desired velocity. On the other hand, the desired time gap and the minimum spacing are the two most significant factors by the PSA. This is consistent with the observation in huge heterogeneity in drivers' desired time gap and minimum spacing. Finally, comparing with GA, the CEM shows its superiority in terms of accuracy, computational efficiency, and convergence property under different termination criteria.

In the future, more trajectories will be used to attempt to increase significance of the desired deceleration on objective function. Besides, more experiments with noise in the objective function are needed to confirm the robustness of our framework.

ACKNOWLEDGEMENTS

Financial support from National 'Twelfth Five-year Plan' Science & Technology Pillar Program under Grant No. 2014BAG01B05**.** The authors are grateful to the NGSIM Program for data availability.

REFERENCES

1. Barceló, J., Fundamentals of traffic simulation. Springer, 2010.
2. Saifuzzaman, M., Z. Zheng. Incorporating human-factors in car-following models: A review of recent developments and research needs. Transportation Research Part C 48, 2014, pp. 379-403.
3. Hoogendoorn, S., and R. Hoogendoorn. Calibration of microscopic traffic-flow models using multiple data sources. Philosophical Transactions Royal Society A: 368, 2010, pp. 4497-4517.
4. Treiber, M., A. Kesting. Microscopic calibration and validation of car-following models - a systematic approach. Procedia-Social and Behavioral Sciences 80, 2014, pp. 922-939.
5. Kesting, A., M. Treiber. Calibrating car-following models using trajectory data: methodological study. In Transportation Research Record: Journal of the Transportation Research Board, No. 2088, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 148-156.
6. Ngoduy, D., M. Maher. Calibration of second order traffic models using continuous cross entropy method. Transportation Research Part C 24, 2012, pp. 102-121.
7. Ciuffo, B., V. Punzo. 'No Free Lunch' Theorems applied to the calibration of traffic simulation models. IEEE Transactions Intelligent Transportation Systems, 15(2), 2014, pp. 553-562.
8. Zhong, R., Chen, C., Andy, C., Pan, T., Yuan, F., Z. He. Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models. Journal of Advanced Transportation, 2015.
9. Maher, M., Liu, R., and D. Ngoduy. Signal optimisation using the cross entropy method. Transportation Research Part C 27, 2013, pp. 76-88.
10. Ciuffo, B., and C. Azevedo. A sensitivity-analysis-based approach for the calibration of traffic simulation models. IEEE Transactions Intelligent Transportation Systems, 2014, pp. 1298-1309.
11. Ge, Q., Ciuffo, B., M. Menendez. An exploratory study of two efficient approaches for the sensitivity analysis of computationally expensive traffic simulation models. IEEE Transactions Intelligent Transportation Systems, 15(3), 2014, pp. 1288-1297.
12. Saltelli, A., Ratto, M., T. Andres. Global sensitivity analysis: the primer, 1st ed. John Wiley & Sons, Inc., New York, 2008.
13. Ma, J., Dong, H., H. Zhang. Calibration of microsimulation with heuristic optimization methods. In Transportation Research Record: Journal of the Transportation Research Board, No. 1999, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 208-217.
14. R. Rubinstein. Optimization of computer simulation models with rare events. European Journal of Operations Research, No. 99, 1997, pp. 89-112.
15. De Boer, P., Kroese, D., Mannor, S., R. Rubinstein. A tutorial on the cross-entropy method. Annals of Operations Research, No. 134, 2005, pp. 19-67.

16. Kroese, D., Rubinstein, R., and P. Glynn. The cross-entropy method for estimation. Handbook of Statistics, Vol 31. Chennai: Elsevier B.V., 2013, pp. 19-34.
17. Liu, H., Chen, W., A. Sudjianto. Relative entropy based method for probabilistic sensitivity analysis in engineering design. Transactions of the ASME, Vol. 128, 2006, pp. 326-336.
18. Treiber, M., Hennecke, A., D. Helbing. Congested traffic states in empirical observations and microscopic simulations. Physics Review E 62 (2), 1805, 2000.
19. Kesting, A., M. Treiber. Calibration of car-following models using floating car data. Springer, Berlin Heidelberg, 2009, pp. 117-127.
20. Treiber, M., A. Kesting. Traffic flow dynamics: data, models and simulation. Springer, 2013.
21. Laval, J., L. Leclercq. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. Philosophical Transactions Royal Society A: 368, 2010, pp. 4519-4541.
22. Ossen, S., S. Hoogendoorn. Car-following behavior analysis from microscopic trajectory data. In Transportation Research Record: Journal of the Transportation Research Board, No. 1934, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 13-21.
23. Duret, A., Buisson, C., N. Chiabaut. Estimating individual speed-spacing relationship and assessing ability of newells carfollowing model to reproduce trajectories. In Transportation Research Record: Journal of the Transportation Research Board, No. 2088, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 188-197.
24. Chiabaut, N., Leclercq, L., C. Buisson. From heterogeneous drivers to macroscopic patterns in congestion. Transportation Research Part B 44 (2), 2010, pp. 299-308.
25. Kroese, D., Porotsky, S., R. Rubinstein. The cross-entropy method for continuous multi-extremal optimization. Methodology and Computing in Applied Probability, No. 8, 2006, pp. 383-407.
26. Komorowski, M., Costa, M., Rand, D., D. Stumpf. Sensitivity, robustness, and identifiability in stochastic chemical kinetics models. Proceedings of the National Academy of Sciences of the United States of America, 108, 2011, pp. 8645-8650.
27. Pantazis, Y., M. Katsoulakis. A relative entropy rate method for path space sensitivity analysis of stationary complex stochastic dynamics. The Journal of Chemical Physics, 138, 2013, 054115.
28. Punzo, V., Borzacchiello, M., B. Ciuffo. On the assessment of vehicle trajectory data accuracy and application to the next generation simulation (NGSIM) program data. Transportation Research Part C: 19, 2011, pp. 1243-1262.
29. Ossen, S., S. Hoogendoorn. Calibrating car-following models using microscopic trajectory data: A critical analysis of both microscopic trajectory data collection methods, and calibration studies based on these data. Techical report, Delft University of Technology, 2008.
30. Paz, A., Molano, V., Martinez, E., Gaviria, C., C. Arteaga. Calibration of traffic flow models using a memetic algorithm. Transportation Research Part C 55, 2015, pp. 432-443.
31. Rahman, M., Chowdhury, M., Khan, T., P. Bhavsar. Improving the efficacy of car-following models with a new stochastic parameter estimation and calibration method. IEEE Transactions Intelligent Transportation Systems, 2015, inpress.