Communications in Transportation Research 5 (2025) 100181

Contents lists available at ScienceDirect

Communications in Transportation Research

journal homepage: www.journals.elsevier.com/communications-in-transportation-research

Full Length Article

Traffic simulation optimization considering driving styles

Yunyang Shi [a], Tong Wu [b], Tan Guo [c], Jinbiao Huo [d], Ziyuan Gu [a,*], Yifan Dai [e], Zhiyuan Liu [a,**]

[a] School of Transportation, Southeast University, Nanjing, 210096, China
[b] Department of Civil Engineering, Monash University, Victoria, 3800, Australia
[c] Zhejiang Expressway Information Engineering and Technology Co., Ltd., Hangzhou, 310000, China
[d] Civil & Environmental Engineering Department, University of Wisconsin–Madison, Madison, WI, 53706, USA
[e] Suzhou Automotive Research Institute of Tsinghua University (Wujiang), Suzhou, 215200, China

ARTICLE INFO

Keywords:
Expected speed
parameter calibration
Bayesian optimization
Microscopic traffic simulation
Driving style

ABSTRACT

Parameter calibration is essential for ensuring the accuracy of microscopic traffic simulations. The expected speed is a critical parameter that characterizes behaviors of vehicles in most simulation models, which is influenced by road traffic conditions and the driving characteristics of different drivers. Most existing parameter calibration methods typically concentrate on micro-level parameters such as time headway and lane change motivation, while overlooking the calibration of vehicle expected speeds in consideration of driver behavior habits. This study combines data from highway electronic toll collection (ETC), gantries, and 100-m mileage average speed data, and proposes a method for calibrating vehicle expected speed that considers driving style clustering. The Gaussian mixture model (GMM) algorithm is used to develop driver models with three distinct driving styles: aggressive, moderate, and conservative. To ensure driving diversity and enhance parameter calibration efficiency, we rebuild vehicle driving models and representative parameters based on the classification results. Moreover, the Bayesian optimization algorithm is modified in conjunction with a microscopic traffic simulation model to perform automatic calibration of expected speeds. Experiments conducted on the Shanghai–Hangzhou–Ningbo highway demonstrate that the proposed method significantly reduces the mean absolute percentage error (MAPE) from 20.2% (using default parameters) to 3.1%. Additionally, in the model robustness test, the MAPE reaches 5.01%, indicating a certain level of stability and scalability. This method proposes a tailored calibration method accounting for the heterogeneous driving behaviors of micro-traffic simulation models, achieving satisfactory calibration results for simulation models in highway scenarios.

1. Introduction

The microscopic traffic simulation model, which uses computer programs to simulate the interactions and operations of elements within transportation systems, is an indispensable tool for analyzing and evaluating complex traffic systems. It has been widely recognized that the accuracy and effectiveness of microscopic traffic simulation models are significantly affected by parameters within the model (Shi et al., 2022). The parameters refer to the variables or inputs that characterize the operational, behavioral, and environmental aspects of the simulated traffic system. These parameters are crucial for defining how the traffic simulation behaves and interacts with real-world conditions and for ensuring accurate modeling of traffic flow, driver behavior, and system performance (Xu and Zheng, 2024). Recently, traffic systems are becoming increasingly complex, necessitating the use of microscopic traffic simulation models to support the analysis, operation, and management of transportation systems (Gu et al., 2023; Liu et al., 2023; Thonhofer et al., 2018). Despite these advancements, existing microscopic traffic simulation practical engineering prefers to apply default recommended parameters. This practice can result in simulation distortions and unreliable outcomes, particularly in traffic scenarios characterized by complex road conditions and diverse driving behaviors (Marina Martinez et al., 2018). Therefore, to advance transportation practice, it is essential to efficiently determine the value of parameters, generating accurate simulation outputs. The problem of increasing the accuracy of simulation models by determining the optimal values of parameters is referred to as parameter calibration problems in Chiappone et al. (2016).

* Corresponding author.
** Corresponding author.
E-mail addresses: ziyuangu@seu.edu.cn (Z. Gu), zhiyuanl@seu.edu.cn (Z. Liu).

https://doi.org/10.1016/j.commtr.2025.100181
Received 11 September 2024; Received in revised form 7 December 2024; Accepted 20 December 2024
Available online xxxx
2772-4247/© 2025 The Authors. Published by Elsevier Ltd on behalf of Tsinghua University Press. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

The parameter calibration for microscopic traffic simulation is inherently intricate and time-intensive. Due to the large number of parameters involved in microscopic traffic simulation, it is impractical to calibrate all of them simultaneously. Existing studies typically develop calibration methodologies for specific parameters, tailored to specific objectives or simulation scenarios. This study specifically focuses on simulation on highway scenarios. Compared to urban networks, highways equipped with better closure and data acquisition conditions, providing a solid foundation for high-precision simulations. Whereas, these data sources have not been fully used to enhance the accuracy of simulation models, and the simulation on freeways, when default parameters are used, usually generate unacceptable errors. In simulation, when two vehicles start at the same time and are detected simultaneously, real-world vehicles lag significantly behind their simulated counterparts when observed at a later point in time. This discrepancy becomes more pronounced as the vehicles cover more distance.

Existing parameter calibration for highway simulation usually focuses on parameters that describe driving behaviors, such as car following and lane changing. While these parameters are important, the expected vehicle speed, the maximum safe speed a driver aims to achieve with minimal or no impediment from other vehicles, remains largely unexplored. Huo et al. (2023) and practical applications typically use a uniform random distribution or the average travel speed derived from the electronic toll collection (ETC) data as the vehicle's expected speed. These methods are inherently flawed and lead to inaccurate simulations.

The expected speed varies with driver behavior characteristics. For instance, aggressive drivers reach their destinations more expeditiously than conservative drivers under optimal traffic conditions. Recently, with the extensive coverage of ETC on highways, the expected speed can be calibrated with greater precision by capturing detailed origin–destination (OD) information and accounting for diverse driving styles. Establishing a plausible expected speed for individual vehicles can significantly enhance the representation of driver heterogeneity, thereby improving the overall accuracy of the simulation.

The primary objective of this study is to develop a calibration method for the expected speeds within simulation on freeway scenarios. This will be accomplished by using single-vehicle data obtained from real highway ETC toll stations and 100-m mileage speed data from AutoNavi Company, combined with an efficient parameter calibration method. Different drivers' behavioral characteristics and styles are extracted using clustering algorithms, followed by calibration through an automatic parameter calibration algorithm that incorporates key parameters in microscopic traffic simulation. This continuous iterative process involves searching for an optimal set of input parameters to minimize the error between simulated and observed traffic measurements. Finally, the proposed method is verified using real data from Zhejiang highways in China. Developing a set of parameter calibration methods that account for different driving styles in real-world scenarios is also crucial for researching emerging technologies. This includes hybrid traffic flows composed of human drivers and intelligent networked vehicles. In a hybrid driving environment, accurately simulating human behavior for autonomous vehicles can lead to more consistent driving behavior on the road, thereby enhancing driving safety and reducing accidents.

In summary, the main contributions of this study are as follows.

1) A driving style recognition method based on GMM clustering is proposed, utilizing ETC data for typical highway scenes. Driving styles are classified into three categories: aggressive, moderate, and conservative, to reflect the behavior of different drivers under favorable traffic conditions.
2) Based on the driving style recognition method, a Bayesian optimization (BO) method for calibrating expected speeds is proposed.
3) Experiments based on real road networks show that the proposed method significantly reduces errors and demonstrates high universality and expansibility compared to traditional methods under non-congestion conditions.

The rest of this paper is organized as follows. Section 2 reviews previous studies related to driving style recognition and simulation parameter calibration. Section 3 introduces methods and details of GMM clustering and BO algorithm. Section 4 describes the specific experiments and results, including the processing of real data, the preparation and setting of microscopic traffic simulation, the testing of GMM clustering, and BO parameter calibration. Finally, the conclusions are drawn in Section 5.

2. Related works

2.1. Driving style recognition

Driving style encapsulates the synthesis of driving behaviors and habits influenced by various physiological factors such as individual driving tendencies, personality traits, age, fatigue levels, as well as external factors including traffic conditions and road characteristics (Li et al., 2017). Drivers exhibiting diverse driving styles demonstrate markedly distinct characteristics in their driving operational behaviors throughout the driving process (He et al., 2023). Consequently, drivers with varying styles can make distinct driving decisions when encountering the same traffic environment. Given that driving style information is not directly observable, quantitative analysis of different driving styles has emerged as a topic warranting thorough investigation. Researchers have undertaken extensive efforts from various perspectives to address this challenge.

Driving style is commonly categorized into two distinct types: short-term and long-term (Chu et al., 2023), both of which could be affected by the driver's habits, personality, and driving technical proficiency. Long-term driving style tends to be stable and may change gradually as driving proficiency improves. For instance, bus drivers often exhibit consistent driving behaviors when navigating the same route. Passengers' experiences riding a bus driven by the same driver are relatively uniform, reflecting consistent characteristics such as average acceleration, deceleration, and the frequency of lane changes along that route. According to Vaitkus et al. (2014), the long-term driving style of bus drivers with relatively consistent driving duties can be classified into two categories: normal driving style and aggressive driving style. Short-term driving style describes how drivers complete specific driving tasks over a short period. It is often influenced by the immediate driving task and scenario. Microscopic-level driving data is commonly employed in driving style recognition research. For instance, parameters such as brake pedal displacement, acceleration pedal position, and steering wheel angular speed provide direct insights into a driver's driving behavior, thus facilitating the recognition of driving tasks and the driver's short-term driving style, as noted in Khodairy and Abosamra (2021), which is more suitable for the recognition of a driving task and the recognition of the driver's short-term driving style.

For driving style recognition, existing approaches can be categorized into rule-based and machine learning methods, in which machine learning methods include unsupervised, supervised, and semi-supervised techniques. Table 1 outlines recent research efforts in this area. Different data conditions result in variations in driving style recognition methods. In recent years, unsupervised learning has become the mainstream approach for recognizing driving styles.

2.2. Simulation parameter calibration

The calibration of traffic simulation model parameters generally encompasses two essential steps. First, empirical data from real-world scenarios are utilized to validate various parameters within the simulation process, including road network attributes, simulation process parameters, and the optimal combination of simulation model parameters. Second, the disparity between the simulation outputs and real-world observations is rigorously evaluated and verified. This two-step process ensures that the simulation model is accurate and representative of actual traffic conditions (Gu et al., 2024).

Table 1
Summary of driving style recognition methods.

Category	Algorithm	Driving style	Data	Ref.
Rule-based	Fuzzy logic	Below normal, normal, aggressive, very aggressive	Vehicle GPS trajectory data	Aljaafreh et al. (2012)
	Fuzzy-rule	Aggressive, anxious, keen, sedate, economical	Driving simulator	Bär et al. (2011)
	Fuzzy Inference System	Comfortable, normal, sporty	Data available on the CAN bus	Dörr et al. (2014)
	Interval type-2 fuzzy inference system	Calm, moderate, aggressive	Argoverse (public dataset)	Gomes and Wolf (2021)
Unsupervised learning	*K*-means	Five different probability combinations of driving states	Large-scale ride-hailing GPS data	Li et al. (2023)
	K-means	Three driving styles	Data extracted from unmanned aerial vehicle	Li et al. (2024)
	K-means	Positive, relatively positive, relatively cautious, cautious	Data extracted from unmanned aerial vehicle	Wen et al. (2022)
	Modified latent Dirichlet allocation	Three different aggressive levels	CDBD (Chinese driving behavior database)	Qi et al. (2016)
	Kernel fuzzy *C*-Means model, latent dirichlet allocation	—	CDBD (Chinese driving behavior database)	Qi et al. (2019)
Supervised learning	Labeled latent dirichlet allocation	Aggressive, moderate, careful	SPMD (safety pilot model deployment) data	(Chen et al., 2021b)
	SVM, AdaBoost, random forest	—	Vehicle on-board diagnostic (OBD) data	Kumar and Jain (2023)
Semi-supervised learning	Graph-CNNs	Calm, moderate, aggressive	Vehicle GPS trajectory data	Chen et al. (2021a)
	SVM, multiple decision tree	Safety, moderation, risk	Driving simulator	Ma et al. (2019)
	Semi-supervised support vector machine	Aggressive and normal	Driving simulator	Wang et al. (2017)

The parameter calibration of traffic simulation is typically formulated as an optimization problem. The decision variables are the parameters of the traffic simulation model, and the objective function is defined as the discrepancy between actual data and simulation results (Wei et al., 2022). However, the multitude of model parameters in traffic simulation—including those related to roads, vehicles, and the simulation environment—interact with one another, thereby complicating the calibration process. To tackle this issue, extensive studies have been conducted to develop optimization methods for solving the parameter calibration problem. Meta-heuristic algorithms, such as genetic algorithm and simulation degradation, can find optimal solutions under nonlinear conditions (Bai et al., 2023; Yu and Qu, 2022). Henclewood et al. (2017) proposed a method based on Monte Carlo to generate candidate parameter sets for parameter calibration of microscopic-traffic simulation. Dai et al. (2022) combined vehicle trajectory data to calibrate human driving behavior based on associated Kalman filter. However, Microscopic-traffic simulation usually has a large number of parameters and strong randomness, which leads to the exponential increase of the solving time of a class of algorithms represented by heuristics with the increase of the number of parameters, which cannot be processed in a reasonable time. Therefore, such methods are usually only used to solve simulation parameter calibration problems in lower dimensions.

Gradient-based methods have also been applied to parameter calibration problems, and one representative algorithm is the simultaneous perturbation stochastic approximation (SPSA). Antoniou et al. (2016) proposed an improved W-SPSA method to introduce spatio-temporal correlations between parameters and measured values, thereby reducing the impact of noise from unrelated parameters on convergence. Qurashi et al. (2020) proposed the PC-SPSA method, which reduces the size and complexity of the problem through dimensionality reduction, and improves the solving efficiency of SPSA. Ho et al. (2023) combined the gradient of SPSA with the gradient of the differentiable meta-model function, which further improved the calibration efficiency compared with M-SPSA and SPSA algorithms. Sha et al. (2023) proposed a simulation parameter calibration framework combining traffic collision technology with multi-objective stochastic optimization technology. Seventeen key parameters were calibrated using SPSA algorithm and compared with vehicle trajectory extraction. These methods have satisfactory convergence speed, but they are easy to fall into local optimization. Moreover, the computational cost of SPSA methods can be expensive because in each iteration, three simulation evaluations are required to approximate the gradient information.

In recent years, surrogate model methods have gained prominence among researchers for addressing parameter calibration problems. These models construct an approximation of the complex objective function based on a limited set of model inputs and outputs. By optimizing this approximation function, surrogate models enhance the efficiency of the solution process, significantly improving the calibration accuracy and computational rate. For example, BO (Huo et al., 2023) and Kriging-based methods (Jung et al., 2022) have been applied to parameter calibration to approximate the simulation process and improve the efficiency of large-scale simulation sampling. Sha et al. (2020) proposed a microscopic-traffic simulation parameter calibration framework based on BO, and compared with the SPSA method. Results showed higher computational efficiency and better convergence, and was robust to the noise of the target equation. Patwary et al. (2021) developed a metamodel-based simulation optimization framework, taking account of the derivatives of link flows with respect to the calibration. This study attempts to develop parameter calibration method following the principle of surrogate-based optimization.

On the other hand, some researchers investigated the calibration of expected speeds in micro-traffic simulation. Zhang et al. (2016) created a simplified calibration method called the “five-point method” to calibrate the expected speed distribution by using travel speed, but ignore the heterogeneity of multiple driving styles and their behaviors. As a result, the ratio k between travel speed and expected speed of the same vehicle is fixed, ignoring the influence of different driving styles on different types of vehicles. On this basis, Maheshwary et al. (2020) used Latin Hypercube design to identify the sensitive parameters affecting driver behavior for each vehicle type, took the specific driving time of the vehicle class as the performance measurement standard, and adopted an optimization method based on genetic algorithm. The optimal parameter sets for different vehicle classes are obtained. Li et al. (2014) proposed the Driving Behavior Questionnaire (DBQ) to collect information about driving attitudes and behaviors and their possible impact on safety and traffic performance. They analyzed the driving behavior of 30 drivers in Changsha, China, through questionnaires and on-board tests, divided them into four driving types, and calibrated a micro-simulation model to reflect the expected speed and acceleration characteristics of these driving types (Li et al., 2015). However, limited by the amount of data and the subjectivity of the questionnaire, the representativeness and universality of the research results may be affected. By using 14 driving

parameters and the driver's age, gender, driving experience, annual mileage, and other attribute data, Chen and Chen (2019) identified three types of driving styles through cluster analysis and principal component analysis, and verified the influence of driver attributes on the classification of driving styles, but there was no further discussion on simulation speed calibration. Existing studies generally overlook the diversity of driving behaviors, failing to adequately consider the heterogeneity of multiple driving styles and their associated behaviors and preventing them from fully capturing the actual driving behaviors. To address these issues, this study employs the GMM clustering method to identify driver styles and calibrates the expected speed of vehicles through BO. This approach comprehensively considers the diversity of driving behaviors, thereby enhancing the accuracy of the simulation.

3. Methodology

Utilizing representative indicators and categorizing them through unsupervised learning is an effective approach to classify and explore the relationships among various driving styles. This section employs travel time and speed for GMM clustering to classify driver styles into three categories. Subsequently, BO is applied in combination with micro-simulation. Combing the clustering and the calibration, a parameter calibration method, the BOG, is developed. The following sections will introduce the theoretical principles and methodology behind GMM clustering and BO. The overall framework of this article is shown in Fig. 1.

3.1. Gaussian mixture model clustering method

GMM constitutes a statistical framework that encapsulates the probability density function of a dataset through the amalgamation of multiple Gaussian distributions, and each characterized by its respective weight. This modeling approach finds particular utility in datasets exhibiting potential heterogeneity, wherein the distribution of data within distinct subgroups can be effectively approximated by Gaussian distributions. In contrast to conventional hard clustering methodologies, GMM introduces a soft clustering mechanism, allowing for nuanced assignment of data points to multiple clusters with varying degrees of certainty. The underlying assumption of GMM posits that the dataset originates from a composite of multiple Gaussian distributions. Each Gaussian distribution, termed a "component", is parameterized by three essential attributes: mean μ_k, covariance matrix Σ_k, and weight π_k. These parameters collectively define the distribution of a given data point x across each component, whose probability density function (PDF) can be expressed as

$$p(x|\mu_k,\Sigma_k)=\frac{1}{\sqrt{(2\pi)^d|\Sigma_k|}}\exp\left(-\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)\right) \tag{1}$$

where d is the characteristic dimension of the data and$|\ \Sigma_k\ |$is the determinant of the covariance matrix.

The probability density function of the k-th Gaussian distribution is defined by its mean μ_k and the covariance matrix Σ_k, and each Gaussian distribution has a weight π_k associated with it. Thus, for a given data point x, the probability that it belongs to the k-th Gaussian distribution can be calculated using the probability density function of the Gaussian distribution. The probability distribution over the entire dataset is the weighted sum of these Gaussian distributions, expressed as

$$p(x|\Lambda)=\sum_{k=1}^{K}\pi_k p(x|\mu_k,\Sigma_k) \tag{2}$$

where Λ represents the set of all parameters of the model, including the mean, covariance matrix, and weights of all Gaussian distributions.

GMM clustering algorithm usually uses expectation maximization (EM) algorithm to estimate model parameters. The EM algorithm is an iterative method that consists of two main steps: the expectation step (E-step) and the maximization step (M-step). In E-step, the algorithm calculates the posterior probability of each data point belonging to each Gaussian distribution under the current model parameters, expressed as

$$r_{ik}=p(z_i=k|x_i,\Lambda)=\frac{\pi_k p(x_i|\mu_k,\Sigma_k)}{p(x_i|\Lambda)} \tag{3}$$

Then, in the M-step, the algorithm uses these posterior probabilities to update the model parameters, including the weights, means, and covariance matrices for each Gaussian distribution. Specifically,

Update of weight π_k:

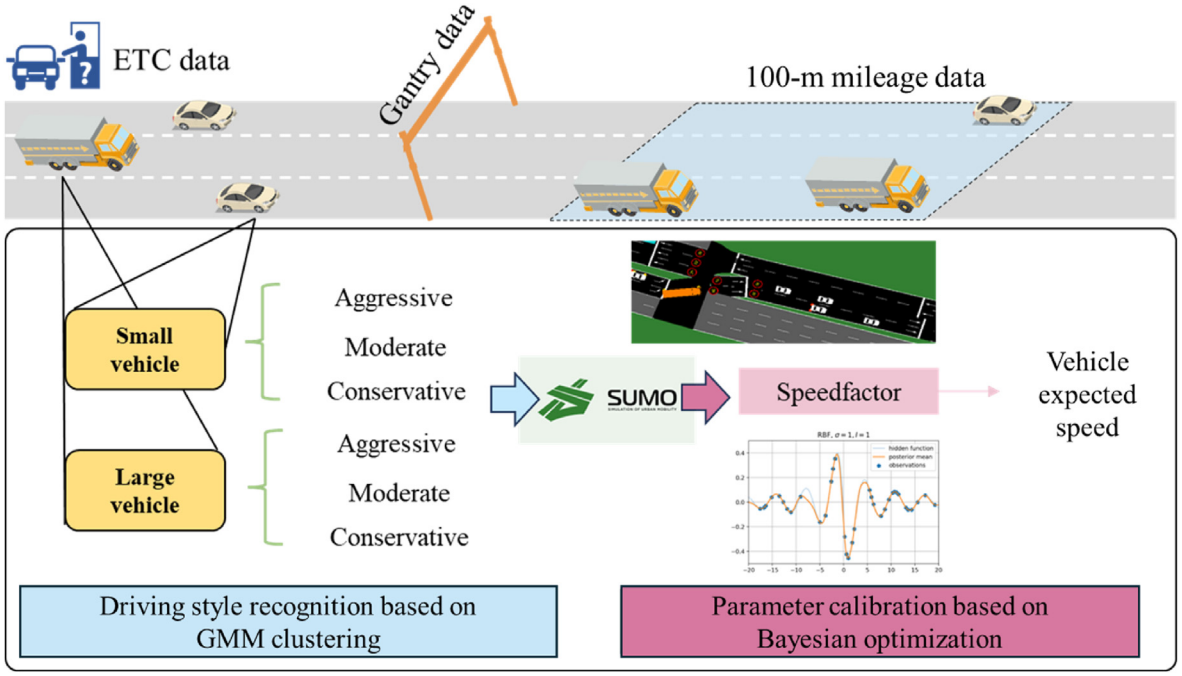

Fig. 1. BOG simulation parameter calibration framework.

$$\pi_k = \frac{1}{N}\sum_{i=1}^{N} r_{ik} \tag{4}$$

Update of mean μ_k:

$$\mu_k = \frac{\sum_{i=1}^{N} r_{ik} x_i}{\sum_{i=1}^{N} r_{ik}} \tag{5}$$

Update of covariance matrix Σ_k:

$$\Sigma_k = \frac{\sum_{i=1}^{N} r_{ik}(x_i - \mu_k)(x_i - \mu_k)^T}{\sum_{i=1}^{N} r_{ik}} \tag{6}$$

The update of weight π_k reflects the evaluation of the fitting degree of the entire dataset by the k-th Gaussian distribution of the current model. The update of the mean μ_k based on the difference between the data points and the currently estimated Gaussian distribution, while the update of the covariance matrix Σ_k takes into account the difference between the data points and their relationship to the current Gaussian distribution.

The iterative process of the GMM clustering algorithm continues until the change in model parameters falls below a preset threshold or a predetermined number of iterations is reached. At the conclusion of the algorithm, each data point is assigned a set of posterior probabilities indicating the degree to which it belongs to each Gaussian distribution. These posterior probabilities can be utilized for further analysis, such as determining the cluster membership of data points or assessing cluster uncertainty. A detailed description of the algorithm can be found in Algorithm 1.

Algorithm 1. Gaussian mixture model for driving style recognition

Input: D: dataset, K: number of Gaussian components, Epsilon: convergence threshold,
MaxIterations: maximum number of iterations
Output: Cluster labels for each data point in D

Step 1: Initialization
Randomly initialize the mean μ_k for each Gaussian component $k = 1$–K
Set the covariance matrices Σ_k to the identity matrix for each k
Equally distribute the weights π_k for each component
Step 2: EM algorithm iteration
For iteration = 1 to MaxIterations do
  E-step:
    For each data point i in D do
      Compute the posterior probabilities r_{ik} for each component k based on the current parameters
  M-step:
    Update the parameters (weights π_k means μ_k, and covariances Σ_k) based on the posterior probabilities r_{ik}
  Convergence Check:
    If the change in parameters is less than Epsilon, exit the loop
If the algorithm reaches MaxIterations without converging, issue a warning
Return the final cluster labels
End

GMM, as a statistical model founded upon probabilistic clustering principles, is distinguished by its mean vector and covariance matrix. In driver behavior data analysis and classification, GMM's adaptable and probabilistic nature makes it well-suited for capturing the distributional characteristics of data. This results in nuanced representation of different driving styles. Unlike rigid categorical allocation, GMM probabilistically assigns each data point to multiple clusters. This flexibility is crucial for accommodating the complex patterns and transitional states inherent in driving behavior data, allowing for a more accurate and detailed analysis of various driving styles.

Furthermore, using statistical metrics such as the bayesian information criterion (BIC), GMM autonomously determines the optimal cluster count, thereby mitigating the influence of subjective biases. Consequently, this study advocates for the adoption of GMM as the benchmark clustering algorithm for driving style recognition.

3.2. Bayesian optimization method for parameter calibration

This section introduces a parameter calibration method based on the BO algorithm. BO is widely recognized to be an efficient method for solving parameter calibration problems. The basic idea of BO is constructing a surrogate model (an analytical function that defines the distribution over the underlying objective function) from data acquired by evaluating the objective function at a few points (termed sampled points). The surrogate model is then used to estimate the location of the optimum, and suggest new sample points for model refinement. The efficiency of BO stems from its capability of incorporating prior beliefs of the problem to inform the sampling of points. The parameter calibration model is first established in Section 3.2.1. Then, the BO algorithm is introduced in Section 3.2.2.

3.2.1. Parameter calibration model

Consider the parameter calibration problem as

$$\min_{\theta} f(\boldsymbol{\theta}, \boldsymbol{p}, \boldsymbol{D}) \tag{7}$$

subject to

$$\boldsymbol{\theta}_l \le \boldsymbol{\theta} \le \boldsymbol{\theta}_u \tag{8}$$

where $f(\cdot)$ measures the difference between simulation outputs and field measurements; $\boldsymbol{\theta}$ is the decision vector, i.e., the simulation parameters to be tuned; $\boldsymbol{p}$ captures the exogenous variables (e.g., network topology) of the simulation model; $\boldsymbol{D}$ defines the field measurements (e.g., densities or travel speeds); $\boldsymbol{\theta}_l$ ($\boldsymbol{\theta}_u$) defines the lower (upper) bounds of $\boldsymbol{\theta}$.

The objective of the optimization problem is to minimize the difference between simulation outputs and field measurements. Let us first illustrate the intricacy of $f(\boldsymbol{\theta}, \boldsymbol{p}, \boldsymbol{D})$ (hereafter denoted as $f(\boldsymbol{\theta})$). Traffic simulators intricately embed detailed descriptions of complex travel behaviors. Specifically, they meticulously define car-following behaviors, lane-changing behaviors, and actions at ramp-merging segments for each vehicle, taking into account the surrounding traffic conditions. As results, $f(\boldsymbol{\theta})$ is non-convex and non-differentiable. The primary challenge of solving the parameter calibration problem lies in the fact that the objective function is highly nonlinear and without closed form expressions or derivatives with respect to the parameter $\boldsymbol{\theta}$. Thus, identifying satisfactory solutions that minimize the intricate objective function is a challengeable task.

3.2.2. Bayesian optimization method

BO is a sequential design technique for the global optimization of black-box functions (Jasper, 2012). It involves two iterative steps: (1) constructing surrogate models based on samples and (2) sampling new points to refine these surrogate models. The first step uses Gaussian processes regression (GPR) to define the prior and posterior distributions over the objective function. The second step involves optimizing acquisition functions to balance exploration and exploitation when searching for new samples.

BO algorithms use GPR model to construct the surrogate model for the objective function $f(\cdot)$. Specifically, GPR assumes that $f(\cdot)$ is a sample path of a Gaussian process f_{GP}, which has two main components, the mean function $\mu(\boldsymbol{\theta})$ and the covariance function $k(\boldsymbol{\theta}, \boldsymbol{\theta}')$. A GP prior is first assigned to the complicated function $f(\cdot)$:

$$f_{\mathrm{GP}}(\boldsymbol{\theta}) \sim N(\mu_0(\boldsymbol{\theta}), k_0(\boldsymbol{\theta}, \boldsymbol{\theta}')) \tag{9}$$

The prior distribution reflects the prior beliefs of the function $f(\cdot)$, which can be chosen by the users. Typically, the mean function is set to a constant, $\mu_0(\boldsymbol{\theta}) = 0$, and the covariance function is set as the radial bias

function, $k_0(\boldsymbol{\theta},\boldsymbol{\theta}') = \exp(-\|\boldsymbol{\theta}-\boldsymbol{\theta}'\|^2/2l^2)$, where $\|\cdot\|$ is the Euclidean norm and l is a hyperparameter of the function. The prior distribution is then updated based on measurements of the objective function. Let $X_n = \{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n\}$ denote n sample points, of which the associated objective function values are denoted by $Y_n = (f_1,\ldots,f_n)$. Given the observation dataset $\{X_n, Y_n\}$, the Gaussian process can be updated accordingly.

$$f_{GP}(\boldsymbol{\theta})|\{X_n,Y_n\} \sim N(\mu_n(\boldsymbol{\theta}), k_n(\boldsymbol{\theta},\boldsymbol{\theta}')) \tag{10}$$

specifically,

$$\mu_n(\boldsymbol{\theta}) = \mu_0(\boldsymbol{\theta}) + k_0(\boldsymbol{\theta},X_n)k_0(X_n,X_n)^{-1}(Y_n-\mu_0(X))$$

and

$$k_n(\boldsymbol{\theta},\boldsymbol{\theta}') = k_0(\boldsymbol{\theta},\boldsymbol{\theta}') - k_0(\boldsymbol{\theta},X_n)k_0(X_n,X_n)^{-1}k_0(X_n,\boldsymbol{\theta}')$$

where $k_0(\boldsymbol{\theta},X_n) = (k_0(\boldsymbol{\theta},\boldsymbol{\theta}_1),\ldots,k_0(\boldsymbol{\theta},\boldsymbol{\theta}_n))^T \in \mathbb{R}^{1\times n}$, $k_0(X_n,\boldsymbol{\theta}) = (k_0(\boldsymbol{\theta},\boldsymbol{\theta}_1),\ldots,k_0(\boldsymbol{\theta},\boldsymbol{\theta}_n)) \in \mathbb{R}^n k_0(X_n,X_n) = [k_0(\boldsymbol{\theta}_i,\boldsymbol{\theta}_j)]_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$, and $\mu_0(X_n) = (\mu_0(\boldsymbol{\theta}_1),\ldots,\mu_0(\boldsymbol{\theta}_n)) \in \mathbb{R}^n$. Given $\boldsymbol{\theta}$, $\mu_n(\boldsymbol{\theta})$ can be interpreted as the conditional estimation of $f(\boldsymbol{\theta})$, and $k_n(\boldsymbol{\theta},\boldsymbol{\theta})$ measures the uncertainty of the estimation.

Acquisition functions, derived from GPR models, propose new sample points that potentially enhance the estimation of the objective function. These functions are typically designed to balance exploration and exploitation in the search process. Numerous efficient acquisition functions have been developed and have shown promising results, including the expected improvement (EI) function, the Gaussian process upper confidence bound (GP-UCB) function, and the knowledge gradient (KG) function, among others. In this study, the EI function is adopted.

Let $f^* = \min\{f_1,\ldots f_n\}$. The basic idea of EI is to sample a new solution $\boldsymbol{\theta}_{n+1}$ to maximize the expected improvement over the current best sample that is

$$\max_{\boldsymbol{\theta}} \mathbb{E}[\max(f^* - f(\boldsymbol{\theta}),0)] \tag{11}$$

Given the posterior distribution $f_{GP}(\boldsymbol{\theta})|\{X_n, Y_n\}$, the optimization problem (11) is equivalent to the following problem:

$$\max_{\boldsymbol{\theta}}(f^* - \mu_n(\boldsymbol{\theta}))\Phi\left(\frac{f^*-\mu_n(\boldsymbol{\theta})}{k_n^{1/2}(\boldsymbol{\theta},\boldsymbol{\theta})}\right) + k_n^{1/2}(\boldsymbol{\theta},\boldsymbol{\theta})\phi\left(\frac{f^*-\mu_n(\boldsymbol{\theta})}{k_n^{1/2}(\boldsymbol{\theta},\boldsymbol{\theta})}\right) \tag{12}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution function (CDF) and probability density function (PDF) of the standard normal distribution, respectively. By iteratively formulating the GPR model and optimizing the EI function, BO algorithms successively approximating the original objective function, and finally optimize the problem through surrogate models. Details of the algorithm are summarized in Algorithm 2.

Algorithm 2. Bayesian optimization for parameter calibration

Input: observation data and the calibration problem.
Initialization: Randomly sample n feasible parameters $X_n = \{\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_n\}$. Conduct simulation under each parameter x_i, and obtain the associated objective function f, $i = 1,\ldots,n$. Construct GPR surrogate model based on the observation dataset $\{X_n, Y_n\}$.
While the stopping criterion of sampling has not been met **do**:
  Solve the EI function and obtain a new sample point $\boldsymbol{\theta}_{n+1}$.
  Conduct simulation and obtain, $f_{n+1} = f(\boldsymbol{\theta}_{n+1})$. Update $X_n = X_n \cup \{\boldsymbol{\theta}_{n+1}\}$, $Y_n = Y_n \cup \{f_{n+1}\}$, $n = n+1$.
  Update the GP surrogate model based on X_n and Y_n.
End While
Solve the optimization problem $\boldsymbol{\theta}_{\text{opt}} = \underset{\mathbf{x}}{\text{argmin}}(\mu_n(\boldsymbol{\theta}))$.
Return $\boldsymbol{\theta}_{\text{opt}}$.

The following conditions can be set as stopping criteria of sampling in Algorithm 2: (1) the sampling reaches computational budgets, (2) the minimal objective function value of samples reaches a predetermined value, and (3) the best-found solution remains unchanged for multiple iterations.

4. Experiment and results

4.1. Data processing

Compared to urban roads, highways offer a more controlled environment where drivers can focus primarily on their driving behaviors, thereby mitigating distractions from pedestrians, non-motor vehicles, and complex route selections typical of urban settings. Consequently, highways serve as a more suitable research domain for capturing drivers' distinctive driving style characteristics. Recent advancements in data acquisition equipment have led to the wider coverage, enhanced quality, and richer multi-source data available on highways. In this study, the Hangzhou–Ningbo highway in Zhejiang Province, China is chosen as the research area, leveraging ETC data, gantry passing data, and 100-m mileage data sourced from the AutoNavi map (which records the average speed of vehicles over a specified 100-m interval). The selected data milepost numbers for this study fall within the research area spanning from K188 to K215. Further details on the mapping and related data sources will be discussed in Section 4.3 and Fig. 5.

ETC data, collected through gantries (gate frames) and stations (toll stations), comprise a comprehensive set of eight valid fields, as delineated in Table 3. Such datasets provide valuable insights into the OD information of vehicles. Additionally, the travel time of vehicles can be inferred from regional toll station data. These metrics serve as essential parameters for evaluating driving styles, contributing significantly to the comprehensive analysis of driver behaviors (Table 4).

Gantry data comprise two main categories of information. Firstly, they encompass traffic data recorded by individual gantries, detailing the number of vehicles traversing each specific gantry in a specified time interval. Subsequently, they include data pertaining to the average vehicle speed beneath each gantry. Here, “gantryid” is the unique identifier for each gantry, while “aver” denotes the average vehicle speed per 10-min period in this study.

The 100-m mileage data provided by AutoNavi delineate the spatial average speed within the detection interval. The data are precisely articulated as the harmonic average speed of all vehicles traversing the spatial range at a given time, as outlined in Table 5. In addition to the above data, the remaining supplementary data can be found in the Electronic Supplementary Material.

4.2. Driving style recognition based on GMM clustering

In the study area, traffic flow exhibits significant fluctuations, resulting in considerable variance in running speeds across different segments (Table 6). Erratic traffic conditions significantly impact driving behaviors, especially during the formation of bottlenecks. In congestion, drivers' behaviors are constrained, making it difficult to observe individualized driving styles. Even if a driver intends to maintain a higher speed, congested road conditions inevitably suppress the expression of their typical driving behavior. Conversely, under smooth traffic conditions, drivers can adhere more closely to their preferred speeds, thereby more accurately reflecting their driving styles. During periods of low traffic flow, vehicles typically operate at free-flow speeds, reducing the mutual influence among drivers. Consequently, differences in driving

Table 3
Detailed description of ETC data.

Name	Description	Type
Vehicleplate	License plate number	Text
OriginEdge	Starting point (vehicle entering the section)	Text
DestinationEdge	End point (when the vehicle enters the section)	Text
StartTime	Starting time (when the vehicle enters the section)	Datetime
EndTime	End time (when the vehicle enters the section)	Datetime
VehicleClass	Vehicle class	Text
RoadID	Road ID	Text
Direction	Driving direction	Text

Table 4
Detailed description of Gantry data.

Name	Description	Type	Unit
Gantryid	Gantry or toll station ID	Text	—
StartTime	Starting time	Datetime	—
Count	Number of passes (per 10 min)	Integer	veh/10 min
Aver	Average speed (per 10 min)	Float	km/h
VrSpeed	Velocity variance (per 10 min)	Float	—

Table 5
Detailed description of 100-m mileage data.

Name	Description	Type	Unit
TimeFlag	Collecting time	Datetime	—
Speed	Space average speed	Float	km/h
StartMileage	Starting milepost number	Integer	km
EndMileage	Ending milepost number	Integer	km
Road	Road ID	Text	—
Direction	Driving direction	Text	—

Table 6
Description of vehicle speed control parameters in SUMO.

Attribute	Default	Range	Description
SpeedFactor	1.0	(0, 1]	The vehicle's expected multiplier of the edge speed limit and desiredMaxSpeed
SpeedDev	0.1	(0, 1]	Deviation of SpeedFactor

styles become less discernible. In such instances, the lack of substantial interaction and constraints tends to homogenize driving behavior, making it difficult to distinguish variations in driving styles among individuals.

To better capture drivers' driving styles and lay the groundwork for subsequent simulation parameter calibration efforts, this study selects appropriate data from AutoNavi's 100-m mileage dataset as the benchmark for calibration. Specifically, data from September 2023 is chosen, excluding instances of severe and occasional congestion observed on certain dates, as depicted in Fig. 2. September 6, 2023, is identified as the dataset for GMM clustering. During this period, traffic maintained a moderate flow without experiencing significant congestion, which serves as a robust basis for further analyses.

To achieve effective clustering of driving styles, it is imperative to discern and scrutinize the salient features that accurately capture driving behaviors. Prior research endeavors have explored a spectrum of factors, encompassing driver control metrics (e.g., brake pedal and accelerator pedal displacement, steering wheel angular speed), driver psychology, electroencephalography (EEG) signals, and meteorological conditions, among others. It is noteworthy, however, that these features often exhibit interrelations, potentially engendering redundant information that could escalate the computational cost of model training and impinge upon its predictive accuracy.

ETC data provides reliable and easily obtainable measurements of travel time, which directly reflect the variability in driving behaviors. In this study's context of non-congested highway scenarios, prioritizing travel time from vehicle OD data in various ETC datasets is an essential way to characterize drivers' behavioral tendencies. Moreover, the reduction in feature dimensionality serves to expedite model training while enhancing its efficacy.

In order to furnish a comprehensive portrayal of traffic dynamics and maximize the availability of vehicle velocity data, the dataset primarily focuses on vehicles with the maximum OD distance, indicating that these vehicles have traveled between the farthest two gantries within the network. Meticulous screening procedures are employed to exclude instances of local traffic congestion, thereby ensuring the robustness and applicability of vehicle expected speed calibration outcomes. Furthermore, cognizant of the inherent disparities among vehicle types, the analytical framework bifurcates into two distinct categories: small vehicles and large vehicles. The former encompasses passenger cars and minivans, with strong acceleration and maneuverability characteristics. Conversely, coaches and trailers belong to the latter category, characterized by slower acceleration profiles and longer braking distances due to their large size and weight. By embracing these inherent differentiations, the analytical framework aims to identify distinct clusters that reveal the nuanced driving behaviors among diverse vehicle types.

Figs. 3 and 4 delineate the outcomes of cluster analysis conducted through GMM, stratifying vehicle driving behavior into three discernible categories: aggressive, conservative, and normal. A detailed scrutiny of the travel speed distribution within each cluster, encompassing metrics such as median, quartile range, and outliers, unveils notable divergences among driving behaviors. In the context of small vehicles, conspicuous differentials emerge between each box plot, coupled with a sparse presence of outliers (notably, a mere six instances of high-speed outliers within cluster = 1). These findings vividly underscore the palpable distinctions in speed patterns across varied behavioral categories. Conversely, within the cohort of large vehicles, discernible disparities between box plots are pronounced, particularly accentuated by the substantial gap observed between cluster = 0 and cluster = 2. Such disparities in speed distribution signify a discernible discontinuity in driving behavior among large vehicles, indicative of significant

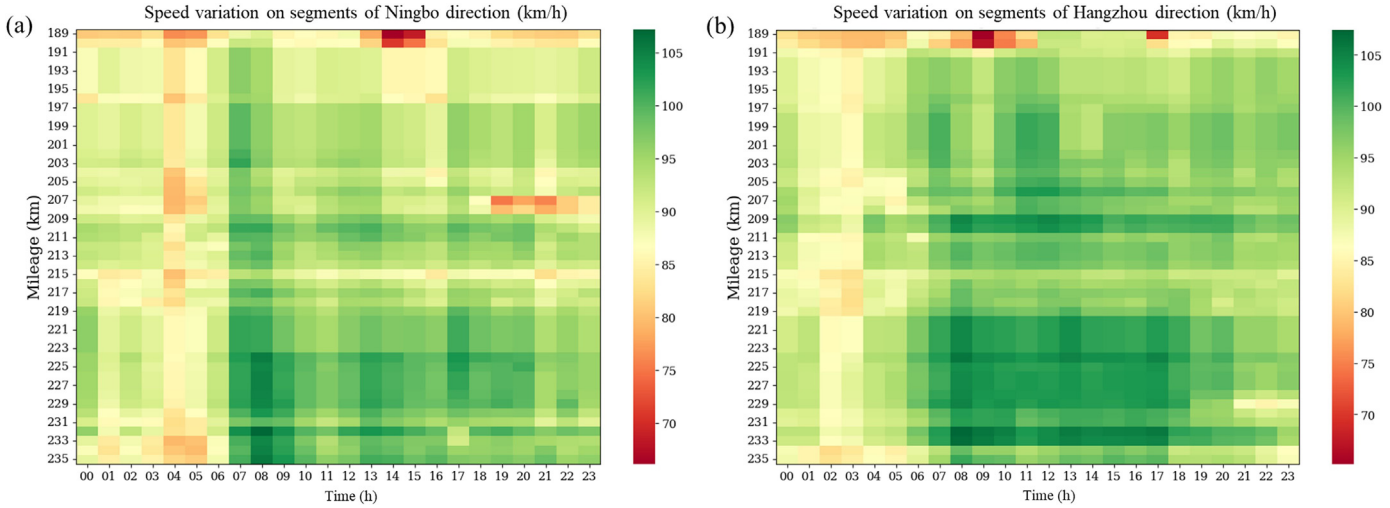

Fig. 2. Space–time diagram of speed variation.

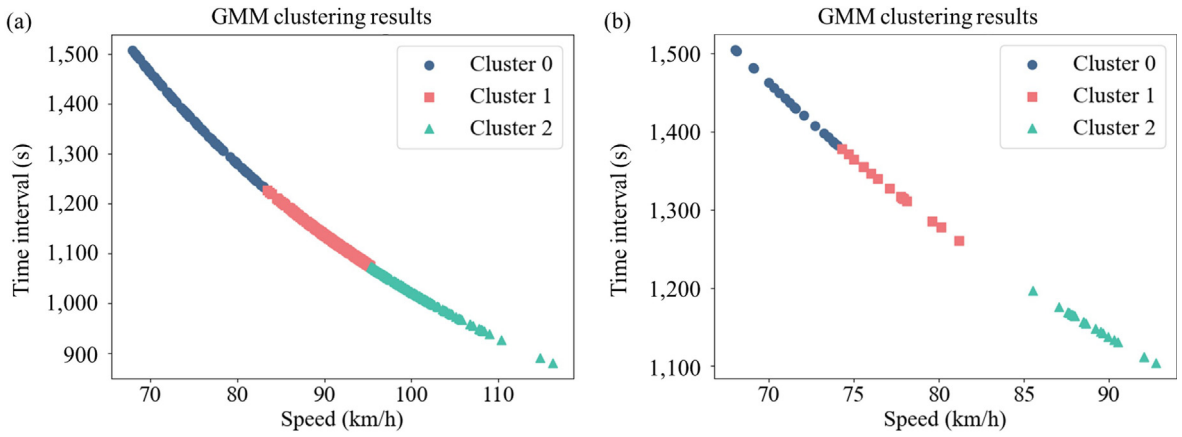

Fig. 3. Driving style clustering based on GMM.

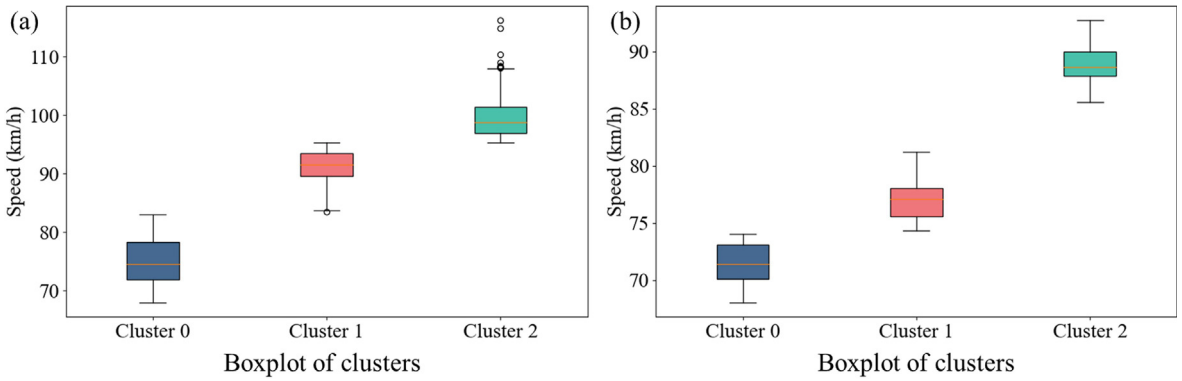

Fig. 4. Box plot of driving style clustering based on GMM.

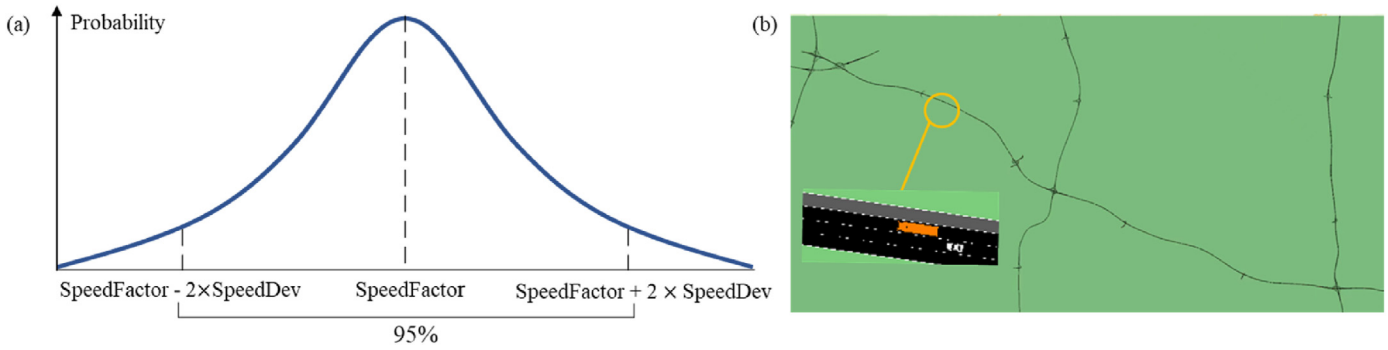

Fig. 5. Preparation for micro traffic simulation in SUMO.

differentials in selected driving speeds between radical and conventional drivers. A comprehensive examination of the clustering outcomes underscores the efficacy of the GMM clustering approach, leveraging ETC travel speed data, in elucidating and characterizing the distinctive driving behavior styles exhibited by small and large vehicles on highways under uncongested conditions. Moreover, these findings underscore the prospective utility of such clustering outputs as inputs for refining parameter calibration in microscopic traffic simulation endeavors.

4.3. Construction of microscopic traffic simulation

SUMO, a microscopic traffic simulation software, has garnered significant attention and adoption among researchers for a myriad of applications, encompassing model verification, parameter calibration, and co-simulation (Gressai et al., 2021; Gu et al., 2023; Lopez et al., 2018; Shi et al., 2023). In this study, SUMO is selected as the simulation tool to verify the proposed methods.

The definition of vehicle speed stands as a pivotal parameter within microscopic traffic simulation frameworks, wielding significant influence over simulation outcomes. Within SUMO, vehicle speed is meticulously modeled through the assignment of distinct SpeedFactor to each vehicle, a methodology seamlessly extendable to encompass road speed limits. This approach entails the multiplication of the road speed limit by the individual speed factor, culminating in the determination of the vehicle's expected free-flow speed. If the individual SpeedFactor exceeds unity, the vehicle's speed may surpass the stipulated road speed limit while adhering to the maximum permissible speed for the respective vehicle type. Hence, under conditions of unrestricted flow, the maximum

attainable speed for a vehicle is governed by the minimum value between the product of the SpeedLimit × SpeedFactor, and the designated MaxSpeed parameter. The SpeedFactor emerges as a pivotal parameter profoundly impacting the anticipated speed profile of vehicles during simulations conducted under free-flow conditions.

To effectively present the clustering results within the microscopic traffic simulation, expected speed of vehicles were converted into SpeedFactors based on the mean and variance derived from the clustering analysis. The SpeedFactor was determined by dividing the travel speed by a reference speed. In the simulation, the SpeedFactor is defined as a truncated normal distribution: SpeedFactor = normc "(mean, deviation, lowerCutOff, upperCutOff)". Different vehicle types necessitate distinct parameter configurations to enhance simulation accuracy. Drawing from the previously conducted GMM clustering of speed data for two vehicle categories (small vehicles and large vehicles) across three driving styles, we obtained the clustering means, variances, and upper and lower limits for each cluster. These parameters were then utilized to set the SpeedFactor for different vehicles in the simulation, thereby accurately representing the performance of drivers of various models and driving styles on the highway. The probability of each data point belonging to a particular cluster was iteratively calculated, and the mean, variance, and weight parameters of each cluster were continuously updated until convergence was achieved. For the purposes of this study, the reference speed was set at 120 km/h for small vehicles and 100 km/h for large vehicles, reflecting the respective speed limits for these vehicle types on the highway.

Based on each cluster's mean and bounds, we calculate the SpeedFactor's mean, upper, and lower bounds to highlight different driving behaviors. By setting a default deviation of speed, we identify various ranges by applying the upper and lower bounds. These SpeedFactor metrics facilitate the identification and detailed description of different driving behavior patterns, as delineated in Table 7.

The traffic characteristics of small and large vehicles vary greatly. For small vehicles, the SpeedFactor of aggressive drivers exhibits a large range of fluctuation, spanning 21.6 km/h, indicating significant variations in travel speed within this cohort. This variability may suggest a propensity for rapid acceleration and deceleration, often associated with aggressive driving styles. Meanwhile, conservative driving is characterized by a mean SpeedFactor of 0.62 and a fluctuation range around 15 km/h. This indicates that while these drivers maintain relatively stable speeds, the range of speed variation remains broad, reflecting cautious driving with periodic adjustments. For large vehicles, the fluctuating ranges of driving speeds are small and similar across the three categories of drivers. This denotes relatively stable travel speeds with a narrow range of fluctuation, characteristic of a more uniform and cautious driving style.

The statistical analysis of SpeedFactor parameters, as derived from the GMM clustering, provides a nuanced understanding of driving behaviors across different vehicle types and driver styles. This comprehensive approach enhances the fidelity and reliability of microscopic traffic simulations, allowing for more accurate modeling of real-world driving dynamics. Such insights are significant for traffic planning, policy formulation, and the development of targeted interventions to improve traffic flow and safety.

Following the GMM clustering analysis, we determined the speed factors for 6 distinct vehicle types, categorized into small and large vehicles, each exhibiting one of three driving styles: aggressive, moderate, and conservative. To validate the rationality of these clusters, we configured various vehicle types (Vtypes) in SUMO based on these six classifications. We then placed four sets of E1 detectors at corresponding real-world detector positions within the road network to compare the simulated speeds with actual speed data.

Table 7
SpeedFactor parameters of vehicles with different driving styles after GMM clustering.

Vehicle type	Driving characteristics	Speed factor
1 (Passenger and van)	Aggressive	Normc (0.83, 0.1, 0.79, 0.97)
Speed limit:120 km/h	Moderate	Normc (0.76, 0.1, 0.69, 0.79)
	Conservative	Normc (0.62, 0.1, 0.56, 0.69)
2 (Coach and trailer)	Aggressive	Normc (0.89, 0.1, 0.85, 0.93)
Speed limit:100 km/h	Moderate	Normc (0.77, 0.1, 0.74, 0.85)
	Conservative	Normc (0.71, 0.1, 0.68, 0.74)

At the start of the simulation, a preheating period is necessary for vehicles to enter the road network, mirroring real-world conditions rather than immediately inputting vehicles into an empty network. To ensure accuracy, we accounted for the preheating and startup time inherent in the simulation by discarding the initial preheating period. We selected five intervals, each spanning 10 min, for the comparative analysis. The actual speed values were compared with the simulated speeds derived from both the GMM algorithm and the specified method, using the mean absolute percentage error (MAPE) as the evaluation metric. MAPE is defined by

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - S_i}{A_i}\right| \tag{13}$$

where A_i represents the actual value and S_i represents the simulation value. This approach allows us to rigorously assess the performance and accuracy of the simulated speeds against real-world data, ensuring that the clustering method and resultant speed factors are both rational and effective for use in microscopic traffic simulation. This validation step is crucial for confirming that the model can accurately replicate real-world driving behaviors and conditions, thereby enhancing the reliability and applicability of the simulation outcomes. Based on the aforementioned methodology, we calculate the percentage error between the simulation speed and real speed using speed data collected from five gantries. Based on the aforementioned methodology, the MAPE value derived using the GMM algorithm is 6.5%. This result indicates that the clustering algorithm enhances the simulation accuracy to a significant extent. However, the presence of residual errors suggests that there is still room for optimization. These errors can be further reduced by integrating advanced parameter calibration algorithms, which would fine-tune the model parameters to better align simulated speeds with observed data.

4.4. Parameter calibration based on bayesian optimization

As the calibration source, data from the Ningbo direction on September 6, 2023, between 9:00 a.m. and 10:00 a.m., were chosen using the data screening procedure outlined in Section 4.2. Firstly, two sets of experiments are performed, including using BO directly and integrating BO with GMM. Two experiments were both configured to run for 30 iterations. Fig. 6 illustrates the error convergence curve as a function of the number of iterations.

Fig. 6a illustrates that the model converges rapidly and reaches an optimal value. The MAPE value derived using BO without GMM algorithm is 7.2%. Compared with the 6.5% error obtained when using only the GMM algorithm, as mentioned in Section 4.3, it is evident that the BO algorithm faces challenges in directly capturing the intrinsic relationships of driving characteristics among different drivers to identify optimized parameters. The integration of GMM and BO holds promise for addressing these issues.

Moreover, Fig. 6b shows that the method rapidly identified satisfactory parameters during the initial sample collection phase but subsequently encountered a local optimal solution. After 27 iterations, the algorithm discovered a new set of parameters with superior performance, reducing the combined error to 3.61%. This implies that utilizing the clustering results from the GMM algorithm in BO helps identify optimal parameter settings more effectively by leveraging inherent patterns and structures in the data, leading to a more accurate and efficient optimization process. Therefore, these errors can be further reduced by integrating advanced parameter calibration algorithms, which would fine-tune the model parameters to better align simulated speeds with observed data.

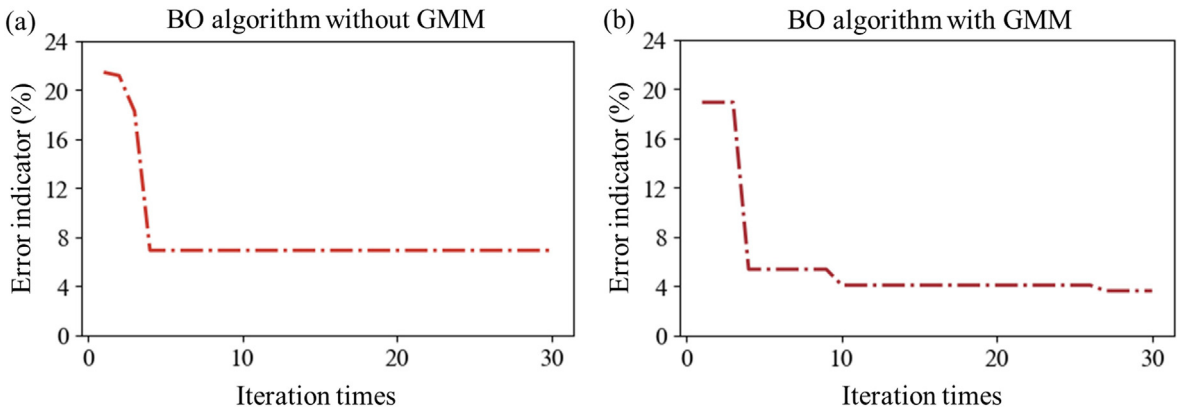

Fig. 6. Error convergence curve (9:00 a.m.–10:00 a.m.).

The parameter settings derived from driving style clustering and subsequently refined using the BO algorithm are detailed in Table 8. To further evaluate the effectiveness of this approach, we compare the simulation results using the default parameters with those obtained from the BOG method, alongside the real-world data (Table 9).

This comparative analysis aims to highlight the improvements in simulation accuracy achieved through the advanced calibration process. The default parameter settings often fail to account for the nuanced variations in driving behavior, leading to discrepancies between the simulated and actual traffic conditions. By contrast, the parameters adjusted through BOG, informed by GMM clustering, are expected to more accurately reflect the diverse driving styles and conditions observed in the real-world data.

When default parameters are utilized in the simulation, the average speed is approximately 1.25 times of the speed detected by real-world sensors, with a comprehensive MAPE of 20.2% (Fig. 7). This significant error highlights the substantial deviations inherent in default parameter settings, which fail to accurately represent actual traffic conditions and driving behaviors. In contrast, by leveraging Gantry data and considering parameter settings across three driving styles and six vehicle types, followed by calibration through the BOG method, the MAPE between the simulated and real values is reduced to 3.1%. This considerable reduction in error demonstrates a substantial improvement in simulation accuracy. Fig. 8 illustrates the scatter plot comparisons of speeds using default parameters, parameters adjusted for driving styles based on the GMM clustering algorithm, and parameters further refined with BO. The scatter plot shows the proximity of the data points to the central line, indicating the alignment between simulation results and actual data. The closer the points are to the middle line, the smaller the discrepancy between simulation and reality. Fig. 8 clearly shows that parameter adjustments using the BOG method yield a significant improvement over the default settings.

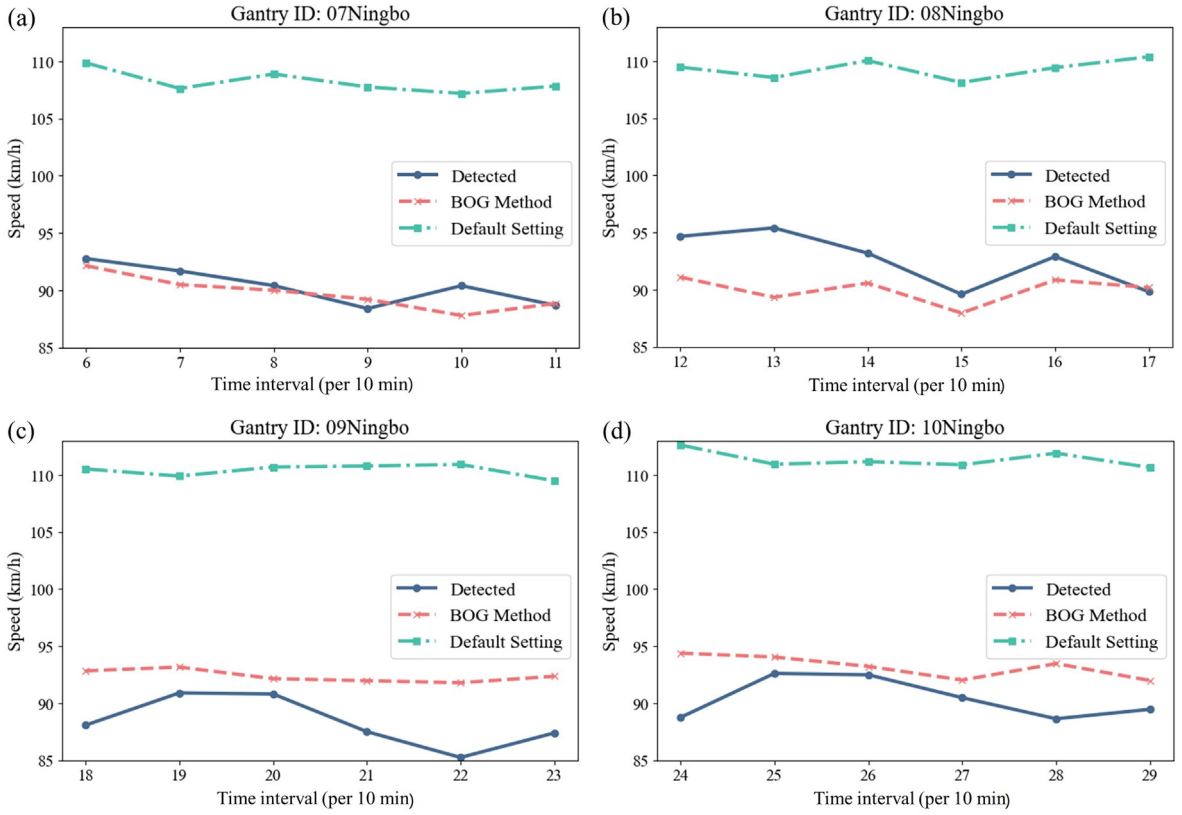

Fig. 7. Comparison of speed variation for different methods during observation time.

Table 8
SpeedFactor parameters of vehicles with different driving styles after BOG (9:00 a.m.–10:00 a.m.).

Vehicle type	Driving characteristics	Speed Factor
1 (Passenger and van)	Aggressive	Normc (0.79451, 0.1, 0.2, 2)
Speed limit:120 km/h	Moderate	Normc (0.72649, 0.1, 0.2, 2)
	Conservative	Normc (0.59859, 0.1, 0.2, 2)
2 (Coach and trailer)	Aggressive	Normc (0.73044, 0.1, 0.2, 2)
Speed limit:100 km/h	Moderate	Normc (0.63289, 0.1, 0.2, 2)
	Conservative	Normc (0.58514, 0.1, 0.2, 2)

Table 9
SpeedFactor parameters of vehicles with different driving styles after BOG (16:00 p.m.–17:00 p.m.).

Vehicle type	Driving characteristics	Speed factor
1 (Passenger and van)	Aggressive	Normc (0.71425, 0.1, 0.2, 2)
Speed limit:120 km/h	Moderate	Normc (0.65310, 0.1, 0.2, 2)
	Conservative	Normc (0.53812, 0.1, 0.2, 2)
2 (Coach and trailer)	Aggressive	Normc (0.74650, 0.1, 0.2, 2)
Speed limit:100 km/h	Moderate	Normc (0.64680, 0.1, 0.2, 2)
	Conservative	Normc (0.59800, 0.1, 0.2, 2)

To further validate the algorithm's transferability and generalizability across different traffic scenarios, we selected a dataset from 16:00 p.m. to 17:00 p.m. on September 6, 2023. The BOG method was employed to recalibrate the SpeedFactor parameters for vehicles exhibiting diverse driving styles, post-clustering. As illustrated in Fig. 9, the error accumulation curve indicates that the algorithm persistently seeks the optimal parameter combination through multiple iterations, achieving stability after approximately 16 rounds and ultimately reducing the overall error to 6.28%. Fig. 10 juxtaposes the real speeds recorded during 16:00 p.m.–17:00 p.m. with the simulated output speeds post-calibration. The calculated mean absolute percentage error (MAPE) for this interval is 5.1%. While this MAPE is slightly higher than the 3.1 % observed in the 9:00 a.m.–10:00 a.m. experiment, it remains within an acceptable range, demonstrating the algorithm's robustness and adaptability.

Moreover, Table 10 provides a comparative analysis of the MAPE between the actual vehicle speeds and the simulation outputs for each detector across the three methods. The MAPE for the Default Setting during 16:00 p.m.–17:00 p.m. exceeds 25%, indicating that the default parameters are unsuitable for practical engineering application. Conversely, the BO with GMM clustering (BOG Method) significantly reduces the MAPE to a range between 2% and 8% for the same period.

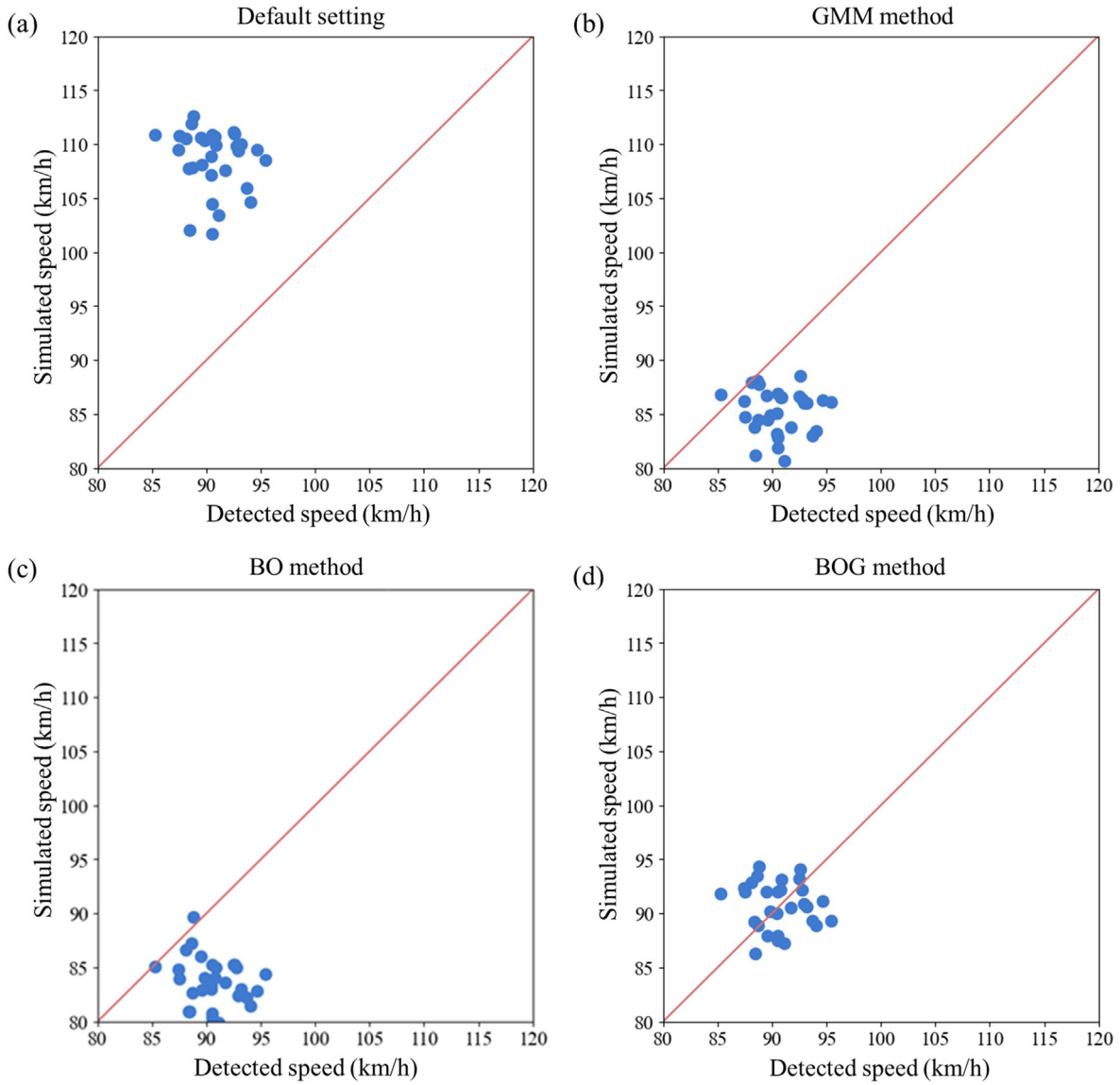

Fig. 8. Comparison of calibrated simulation results with real values of GMM Method, BO Method, and BOG Method.

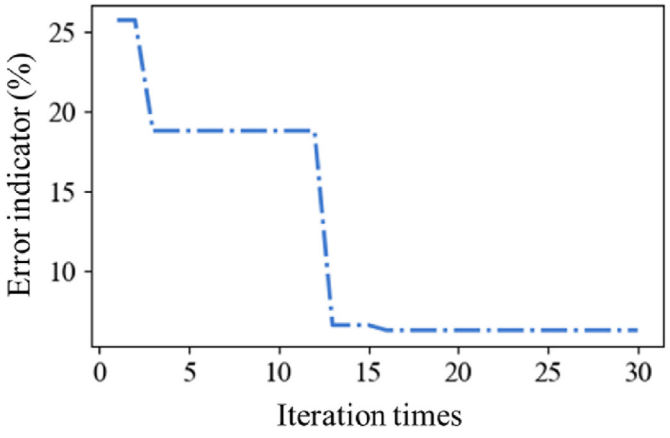

Fig. 9. Error convergence curve of Bayesian algorithm (16:00 p.m.–17:00 p.m.).

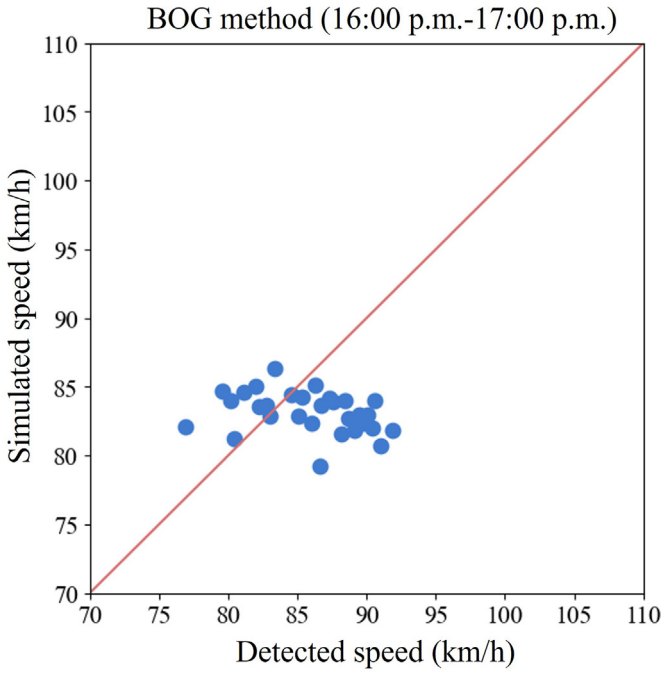

Fig. 10. Comparison of calibrated simulation results with real values (16:00 p.m.–17:00 p.m.).

Table 10
Comparison of simulation results in different scenarios.

Gantry ID	06Ningbo	07Ningbo	08Ningbo	09Ningbo	10Ningbo
Default setting (16:00 p.m.–17:00 p.m.)	26.07%	31.52%	26.88%	38.80%	37.17%
BOG method (16:00 p.m.–17:00 p.m.)	6.11%	6.24%	7.69%	2.58%	2.74%
BOG method (9:00 a.m.–10:00 a.m.)	3.85%	1.07%	2.90%	4.64%	3.11%

Additionally, the BOG Method for the 9:00 a.m.–10:00 a.m. period, which utilizes traffic data corresponding to the GMM cluster analysis, achieves an MAPE between 1% and 5%, demonstrating good calibration accuracy. Despite the difference in time periods, the experiment

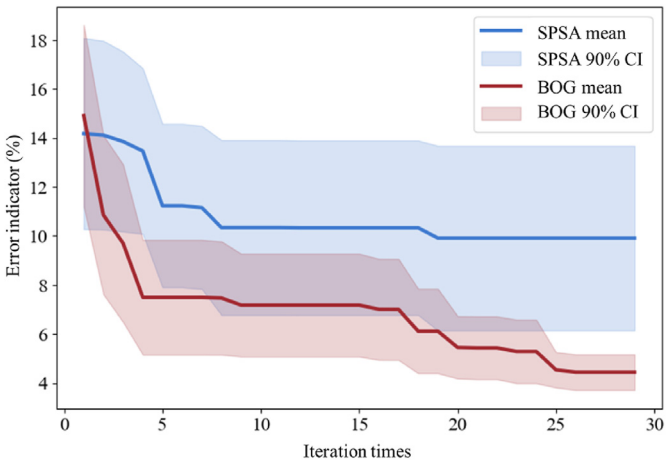

Fig. 11. Comparison of SPSA and BOG with multiple experiments.

conducted from 16:00 p.m. to 17:00 p.m. exhibited only a 2% increase in MAPE compared to the 9:00 a.m.–10:00 a.m. experiment. This minor increase in error highlights the strong portability and robustness of the proposed BOG automatic parameter calibration algorithm, which incorporates driving style considerations.

The simultaneous perturbation stochastic approximation (SPSA) algorithm is widely used for parameter optimization due to its efficiency in high-dimensional problems and its ability to handle noisy or uncertain environments (Jung et al., 2022). In order to consider the randomness of the simulation, a comparison experiment with the SPSA method was performed. Fig. 11 compares the mean convergence curves and their corresponding confidence intervals (CI) of the commonly used SPSA method and the proposed BOG method over 10 repeated experiments, showing the mean and 95% confidence intervals of the error indicator. After 30 iterations, the BOG algorithm demonstrates rapid convergence, stabilizing its error within the range of 4%–6%. In contrast, the SPSA algorithm exhibits less robust performance, with wider error variability ranging from 7% to 14%, indicating lower parameter reliability.

This disparity in performance can be attributed to BOG's ability to effectively leverage the information obtained from GMM clustering, which provides a more nuanced understanding of the underlying parameter relationships. By incorporating the clustering results, BOG not only optimizes the parameters more efficiently but also reduces the risk of converging to suboptimal solutions. In contrast, SPSA, despite its efficiency in high-dimensional spaces, lacks this mechanism to refine its search space, leading to higher variability in error and less reliable parameter estimates.

This consistent performance across different temporal datasets underscores the efficacy of the BO method, combined with GMM-based driving style clustering, in enhancing the precision and applicability of microscopic traffic simulations. The ability to maintain relatively low error rates despite varying traffic conditions suggests that this approach is well-suited for diverse operational contexts. The implications of these findings are significant for the field of traffic simulation. The application of the BO algorithm to parameter calibration highlights its potential to replace more traditional, resource-intensive methods such as sensitivity analysis and trial-and-error approaches. This methodological shift promises to streamline the calibration process, saving time and computational resources while enhancing accuracy. Moreover, integrating GMM clustering into the calibration process allows for a nuanced representation of driver behaviors, crucial for realistic traffic simulations. By considering various driving styles and their influence on traffic flow, this approach creates a more detailed and precise simulation environment. Furthermore, the BOG method proposed in this study provides a general method for mining data and constructing high-precision simulation for highways with extensive ETC equipment. Such advancements can

significantly improve traffic management and inform better planning decisions.

5. Conclusions

Advanced and intelligent parameter calibration algorithms are essential for enhancing the accuracy of microscopic simulation models, which could benefits traffic planning, design, and management decisions. This study highlights deficiencies inherent in default simulation parameter settings, which often lead to simulated vehicle speeds that exceed observed values at various detector positions. Such discrepancies can yield overly optimistic simulation results, potentially influencing research outcomes erroneously. Focusing specifically on highway microscopic traffic simulation, we utilized the GMM clustering algorithm to analyze and classify driver behaviors based on data from ETC, gantry, and 100-m mileage sources. This clustering identified three distinct driving styles—aggressive, moderate, and conservative—across different vehicle types. Subsequently, we developed a categorized vehicle simulation model that incorporates these driving styles and vehicle categories. To refine simulation parameters, we employed a BO algorithm to automatically determine optimal parameter values. Validation against empirical data from the Shanghai–Hangzhou–Ningbo highway demonstrated that the proposed method significantly reduced the MAPE from 20.2% (using default parameters) to 3.1%. Moreover, the method exhibited robust performance in mobility tests, achieving a MAPE of 5.01%.

In conclusion, the parameter calibration method introduced in this study significantly improves the accuracy and reliability of highway microscopic traffic simulations. By incorporating driving styles, this approach offers a more realistic reference for parameter settings, thereby enhancing the precision of simulations. Comparative analysis underscores that combining GMM clustering with BO notably enhances simulation accuracy, effectively overcoming the shortcomings of default parameter configurations. Future research can focus on extending this methodology to diverse road types and traffic conditions to further validate its efficacy and broaden its applicability.

CRediT authorship contribution statement

Yunyang Shi: Writing – review & editing, Writing – original draft, Project administration, Methodology, Formal analysis. **Tong Wu:** Writing – original draft, Visualization, Resources, Investigation, Data curation, Conceptualization. **Tan Guo:** Resources, Investigation, Formal analysis. **Jinbiao Huo:** Writing – review & editing, Visualization, Validation, Software. **Ziyuan Gu:** Validation, Supervision, Project administration, Funding acquisition. **Yifan Dai:** Validation, Visualization. **Zhiyuan Liu:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition.

Replication and data sharing

The data and the complete code for reproducing the simulations reported in the present paper are available at https://github.com/Shiyunyang-seu/BOG-method-for-traffic-simulation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (Nos. 72471057 and 52131203), the Natural Science Foundation of Jiangsu Province (BK20232019), and the Jiangsu Provincial Scientific Research Center of Applied Mathematics (BK20233002).

Electronic Supplementary Material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.commtr.2025.100181.

References

Aljaafreh, A., Alshabatat, N., Najim Al-Din, M.S., 2012. Driving style recognition using fuzzy logic. In: 2012 IEEE International Conference on Vehicular Electronics and Safety (ICVES 2012), pp. 460–463.

Antoniou, C., Barceló, J., Breen, M., Bullejos, M., Casas, J., Cipriani, E., et al., 2016. Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: design, demonstration and validation. Transport. Res. C Emerg. Technol. 66, 79–98.

Bai, C., Jing, J., Liu, B., Yao, W., Yang, C., Alagbé, A.J., Jin, S., 2023. Exploring heterogeneity in car-following behaviors based on driver visual characteristics: modeling and calibration. J. Adv. Transp. 2023, 583081.

Bär, T., Nienhuser, D., Kohlhaas, R., Zollner, J.M., 2011. Probabilistic driving style determination by means of a situation based analysis of the vehicle data. In: 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1698–1703.

Chen, K.T., Chen, H.W., 2019. Driving style clustering using naturalistic driving data. Transp. Res. Rec. J Transp. Res. Board 2673, 176–188.

Chen, C., Liu, Q., Wang, X., Liao, C., Zhang, D., 2021a. Semi-Traj2Graph: identifying fine-grained driving style with GPS trajectory data via multi-task learning. IEEE Trans. Big Data 8, 1550–1565.

Chen, D., Chen, Z., Zhang, Y., Qu, X., Zhang, M., Wu, C., 2021b. Driving style recognition under connected circumstance using a supervised hierarchical Bayesian model. J. Adv. Transp. 2021, 6687378.

Chiappone, S., Giuffrè, O., Granà, A., Mauro, R., Sferlazza, A., 2016. Traffic simulation models calibration using speed–density relationship: an automated procedure based on genetic algorithm. Expert Syst. Appl. 44, 147–155.

Chu, H., Zhuang, H., Wang, W., Na, X., Guo, L., Zhang, J., et al., 2023. A review of driving style recognition methods from short-term and long-term perspectives. IEEE Trans. Intell. Veh. 8, 4599–4612.

Dai, Q., Shen, D., Wang, J., Huang, S., Filev, D., 2022. Calibration of human driving behavior and preference using vehicle trajectory data. Transport. Res. C Emerg. Technol. 145, 103916.

Dörr, D., Grabengiesser, D., Gauterin, F., 2014. Online driving style recognition using fuzzy logic. In: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1021–1026.

Gomes, I.P., Wolf, D.F., 2021. Driving style recognition using interval type-2 fuzzy inference system and multiple experts decision making. Int. J. Fuzzy Syst. 26, 553–571.

Gressai, M., Varga, B., Tettamanti, T., Varga, I., 2021. Investigating the impacts of urban speed limit reduction through microscopic traffic simulation. Commun. Transp. Res. 1, 100018.

Gu, Z., Li, Y., Saberi, M., Rashidi, T.H., Liu, Z., 2023. Macroscopic parking dynamics and equitable pricing: integrating trip-based modeling with simulation-based robust optimization. Transp. Res. Part B Methodol. 173, 354–381.

Gu, Z., Yang, X., Zhang, Q., Yu, W., Liu, Z., 2024. TERL: two-stage ensemble reinforcement learning paradigm for large-scale decentralized decision making in transportation simulation. IEEE Trans. Knowl. Data Eng. 35, 13043–13054.

He, Y., Liu, Y., Yang, L., Qu, X., 2023. Deep adaptive control: deep reinforcement learning-based adaptive vehicle trajectory control algorithms for different risk levels. IEEE Trans. Intell. Veh. 9, 1654–1666.

Henclewood, D., Suh, W., Rodgers, M.O., Fujimoto, R., Hunter, M.P., 2017. A calibration procedure for increasing the accuracy of microscopic traffic simulation models. Simulation 93, 35–47.

Ho, M.C., Lim, J.M., Chong, C.Y., Chua, K.K., Siah, A.K.L., 2023. High dimensional origin destination calibration using metamodel assisted simultaneous perturbation stochastic approximation. IEEE Trans. Intell. Transport. Syst. 24, 3845–3854.

Huo, J., Liu, C., Chen, J., Meng, Q., Wang, J., Liu, Z., 2023. Simulation-based dynamic origin–destination matrix estimation on freeways: a Bayesian optimization approach. Transp. Res. Part E Logist. Transp. Rev. 173, 103108.

Jasper, S., Hugo, L., Ryan, P.A., 2012. Practical bayesian optimization of machine learning algorithms. In: Adv. Neural Inf. Process. Syst., 25, pp. 2951–2959. NIPS.

Jung, Y., Jo, H., Choo, J., Lee, I., 2022. Statistical model calibration and design optimization under aleatory and epistemic uncertainty. Reliab. Eng. Syst. Saf. 222, 108428.
Khodairy, M.A., Abosamra, G., 2021. Driving behavior classification based on oversampled signals of smartphone embedded sensors using an optimized stacked-LSTM neural networks. IEEE Access 9, 4957–4972.
Kumar, R., Jain, A., 2023. Driving behavior analysis and classification by vehicle OBD data using machine learning. J. Supercomput. 79, 1–20.
Li, J., van Zuylen, H., van der Horst, E., 2014. The driver behaviour questionnaire: an investigation study applied to Chinese drivers. In: Computer-based Modelling and Optimization in Transportation. Springer International Publishing, Cham, pp. 433–447.
Li, J., van Zuylen, H.J., Xu, X., 2015. Driving type categorizing and microscopic simulation model calibration. Transp. Res. Rec. J Transp. Res. Board 2491, 53–60.
Li, G., Li, S.E., Cheng, B., Green, P., 2017. Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities. Transport. Res. C Emerg. Technol. 74, 113–125.
Li, Y., Chen, Y., Bao, J., Xing, L., Tang, J., Dong, C., et al., 2023. A modified latent dirichlet allocation topic approach for driving style exploration using large-scale ride-hailing GPS data. J. Adv. Transp. 2023, 3203065.
Li, Y., Zhang, H., Wang, Q., Wang, Z., Yao, X., 2024. Study on driver behavior pattern in merging area under naturalistic driving conditions. J. Adv. Transp. 2024, 7766164.
Liu, Y., Wu, F., Liu, Z., Wang, K., Wang, F., Qu, X., 2023. Can language models be used for real-world urban-delivery route optimization? Innovation 4, 100520.
Lopez, P.A., Wiessner, E., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flotterod, Y.P., et al., 2018. Microscopic traffic simulation using SUMO. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), pp. 2575–2582.
Ma, Y., Li, Z., Li, Y., Li, H., Malekian, R., 2019. Driving style estimation by fusing multiple driving behaviors: a case study of freeway in China. Clust. Comput. 22, 8259–8269.
Maheshwary, P., Bhattacharyya, K., Maitra, B., Boltze, M., 2020. A methodology for calibration of traffic micro-simulator for urban heterogeneous traffic operations. J. Traffic Transp. Eng. Engl. Ed. 7, 507–519.
Marina Martinez, C., Heucke, M., Wang, F.Y., Gao, B., Cao, D., 2018. Driving style recognition for intelligent vehicle control and advanced driver assistance: a survey. IEEE Trans. Intell. Transport. Syst. 19, 666–676.
Patwary, A.U.Z., Huang, W., Lo, H.K., 2021. Metamodel-based calibration of large-scale multimodal microscopic traffic simulation. Transport. Res. C Emerg. Technol. 124, 102859.
Qi, G., Du, Y., Wu, J., Hounsell, N., Jia, Y., 2016. What is the appropriate temporal distance range for driving style analysis? IEEE Trans. Intell. Transport. Syst. 17, 1393–1403.
Qi, G., Wu, J., Zhou, Y., Du, Y., Jia, Y., Hounsell, N., et al., 2019. Recognizing driving styles based on topic models. Transport. Res. Transport Environ. 66, 13–22.
Qurashi, M., Ma, T., Chaniotakis, E., Antoniou, C., 2020. PC–SPSA: employing dimensionality reduction to limit SPSA search noise in DTA model calibration. IEEE Trans. Intell. Transport. Syst. 21, 1635–1645.
Sha, D., Ozbay, K., Ding, Y., 2020. Applying Bayesian optimization for calibration of transportation simulation models. Transp. Res. Rec. J. Transp. Res. Board 2674, 215–228.
Sha, D., Gao, J., Yang, D., Zuo, F., Ozbay, K., 2023. Calibrating stochastic traffic simulation models for safety and operational measures based on vehicle conflict distributions obtained from aerial and traffic camera videos. Accid. Anal. Prev. 179, 106878.
Shi, Y., Liu, Z., Wang, Z., Ye, J., Tong, W., Liu, Z., 2022. An integrated traffic and vehicle co-simulation testing framework for connected and autonomous vehicles. IEEE Intell. Transport. Syst. Mag. 14, 26–40.
Shi, Y., Gu, Z., Yang, X., Li, Y., Chu, Z., 2023. An adaptive route guidance model considering the effect of traffic signals based on deep reinforcement learning. IEEE Intell. Transport. Syst. Mag. 16, 21–34.
Thonhofer, E., Palau, T., Kuhn, A., Jakubek, S., Kozek, M., 2018. Macroscopic traffic model for large scale urban traffic network design. Simulat. Model. Pract. Theor. 80, 32–49.
Vaitkus, V., Lengvenis, P., Zylius, G., 2014. Driving style classification using long-term accelerometer information. In: 2014 19th International Conference on Methods and Models in Automation and Robotics (MMAR), pp. 641–644.
Wang, W., Xi, J., Chong, A., Li, L., 2017. Driving style classification using a semisupervised support vector machine. IEEE Trans. Human-Mach Syst. 47, 650–660.
Wei, S., Feng, S., Ke, J., Yang, H., 2022. Calibration and validation of matching functions for ride-sourcing markets. Commun. Transp. Res. 2, 100058.
Wen, S., Shahd Omar, X.G., Jin, X., He, Z., 2022. Analysis of vehicle driving styles at freeway merging areas using trajectory data. In: 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), pp. 3652–3656.
Xu, Z., Zheng, N., 2024. Integrating connected autonomous shuttle buses as an alternative for public transport–A simulation-based study. Multimodal. Transp. 3, 100133.
Yu, Y., Qu, X., 2022. Development of parametric eco-driving models for fuel savings: a novel parameter calibration approach. Int. J. Transp. Sci. Technol. 11, 268–282.
Zhang, C., Liu, S., Ogle, J., Zhang, M., 2016. Micro-simulation of desired speed for temporary work zone with a new calibration method. Promet-Traffic & Transp. 28, 49–61.

Yunyang Shi received the M.S. degree of traffic and transportation engineering from Southeast University, Nanjing, China, in 2018. He is currently pursuing the Ph.D. degree of traffic and transportation engineering in Southeast University. His research interests include traffic and autonomous vehicle simulation, agent-based simulation.

Tong Wu received the B.S. degree in transportation engineering from Chang'an University, Xi'an, China, in 2023. He is currently pursuing the M.S. degree in transportation systems at the Department of Civil Engineering, Monash University. His research interests focus on traffic simulation, transportation big data modeling, and transportation electrification.

Tan Guo received the B.S. degree in traffic engineering from Chang'an University in 2020 and the M.S. degree in traffic engineering from Southeast University in 2023. He is currently working at Zhejiang Expressway Information Engineering and Technology Co., Ltd. His focus areas include AI large models, smart highways, and traffic simulation.

Jinbiao Huo received the B.E. degree of traffic and transportation engineering from Southeast University, Nanjing, China, in 2020. He is currently pursuing a Ph.D. degree of traffic and transportation engineering in Southeast University. His research interests include traffic simulation and simulation-based transportation problems.

Ziyuan Gu received the B.E. degree from Nanjing Tech University, China; the M.E. degree from Southeast university; and the Ph.D. degree from the University of New South Wales, Australia. He is currently an Associate Professor with the School of Transportation at Southeast University. His research focuses on machine learning and its application in transportation, traffic simulation and simulation-based optimization, etc. He has published more than 50 peer-reviewed papers in internationally renowned journals.

Yifan Dai received the Ph.D. degree in 2013 from Tsinghua University. He is currently serving as the chief engineer at the Suzhou Automotive Research Institute of Tsinghua University. His research interests include autonomous driving technology and vehicle dynamics control technology.

Zhiyuan Liu received the Ph.D. degree in transportation engineering in 2011 from National University of Singapore. He is a Professor in the School of Transportation, and the Youth Chief Professor of Southeast University. His research interests include transport network modeling, public transport, and intelligent transport system.