



Wang, Z., Liu, Z., Cheng, Q., & Gu, Z. (2024). Integrated self-consistent macro-micro traffic flow modeling and calibration framework based on trajectory data. *Transportation Research Part C: Emerging Technologies*, 158, [104439].
<https://doi.org/10.1016/j.trc.2023.104439>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1016/j.trc.2023.104439](https://doi.org/10.1016/j.trc.2023.104439)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



Integrated self-consistent macro-micro traffic flow modeling and calibration framework based on trajectory data

Zelin Wang^a, Zhiyuan Liu^a, Qixiu Cheng^{b,*}, Ziyuan Gu^{a,*}

^a Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transportation, Southeast University, Nanjing 210096, China

^b University of Bristol Business School, University of Bristol, Bristol BS8 1PY, UK

ARTICLE INFO

Keywords:

Self-consistency
Multi-resolution modeling
Car-following
Fundamental diagram
Multi-objective optimization
Deep learning

ABSTRACT

Calibrating microscopic car-following (CF) models is crucial in traffic flow theory as it allows for accurate reproduction and investigation of traffic behavior and phenomena. Typically, the calibration procedure is a complicated, non-convex optimization issue. When the traffic state is in equilibrium, the macroscopic flow model can be derived analytically from the corresponding CF model. In contrast to the microscopic CF model, calibrated based on trajectory data, the macroscopic representation of the fundamental diagram (FD) primarily adopts loop detector data for calibration. The different calibration approaches at the macro- and microscopic levels may lead to misaligned parameters with identical practical meanings in both macro- and micro-traffic models. This inconsistency arises from the difference between the parameter calibration processes used in macro- and microscopic traffic flow models. Hence, this study proposes an integrated multiresolution traffic flow modeling framework using the same trajectory data for parameter calibration based on the self-consistency concept. This framework incorporates multiple objective functions in the macro- and micro-dimensions. To expeditiously execute the proposed framework, an improved metaheuristic multi-objective optimization algorithm is presented that employs multiple enhancement strategies. Additionally, a deep learning technique based on attention mechanisms was used to extract stationary-state traffic data for the macroscopic calibration process, instead of directly using the entire aggregated data. We conducted experiments using real-world and synthetic trajectory data to validate our self-consistent calibration framework.

1. Introduction

Longitudinal car-following (CF) models are critical in traffic flow studies, particularly in describing driving behavior and predicting vehicle trajectories. Since the 1950s, research on CF models has evolved significantly and has been widely implemented (Brackstone and McDonald, 1999, Saifuzzaman and Zheng, 2014). Numerous microscopic models have been developed to replicate traffic flow patterns using various approaches. The stimulus-response model (Chandler et al., 1958, Gazis et al., 1961), safety distance model (Kometani and Sasaki, 1959, Gipps, 1981), optimal velocity model (OVM) (Bando et al., 1995, Jiang et al., 2001), desired measures model (Helly, 1959, Treiber et al., 2000), and psycho-physiological model (Michaels, 1963, Wiedemann, 1974), among others, are

* Corresponding authors.

E-mail addresses: zelinwang@seu.edu.cn (Z. Wang), zhiyuanyan@seu.edu.cn (Z. Liu), qixiu.kevin.cheng@gmail.com (Q. Cheng), ziyuangu@seu.edu.cn (Z. Gu).

<https://doi.org/10.1016/j.trc.2023.104439>

Received 16 May 2023; Received in revised form 16 September 2023; Accepted 25 November 2023

0968-090X/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

some of the most representative categories of microscopic models developed to replicate traffic flow patterns (Wang et al., 2023). Classical CF models typically feature parameters with practical physical significance. The macroscopic traffic stream model can be derived analytically from a specific CF model where the driver's behavior is in equilibrium under steady-state¹ conditions (Gu et al., 2022). In this state, the speed difference is zero, and the spacing headway is fixed for vehicles. It is well recognized that when traffic flow is in a stationary state, the density of the macroscopic traffic flow and the space headway in microscopic traffic flow models are reciprocally inverse. The parameters in microscopic CF models and their associated macroscopic fundamental diagram (FD)² models have well-defined physical meanings and coexist. However, the calibration results show significant differences (Wang et al., 2019). To demonstrate this, we calibrated the parameters in the CF model first and then applied them after parameter transformation to the FD model. As shown in Fig. 1, if the parameters estimated from the CF model are directly substituted into the corresponding steady-state FD model, the model fitting performance deviates significantly from the FD curve estimated from the speed-density scatter, as indicated by the green dashed line in Fig. 1.

This study has identified two primary inconsistencies in current traffic flow modeling and calibration methods, leading to inconsistent outcomes. One of the major contributors to these inconsistent calibration results is the use of divergent datasets with varying acquisition time intervals and measurement accuracies. While the CF model calibration uses trajectory data, FD models are usually calibrated using data collected from loop detectors, with only a few studies using probe vehicle trajectory data (Seo et al., 2019). Nevertheless, the assumption of traffic flow being in equilibrium state also warrants equal attention. Although there are limited studies on identifying and extracting nonstationary and stationary states of traffic flow from raw data (Castillo and Benítez, 1995, Cassidy, 1998, Yan et al., 2018, Wang et al., 2019), almost all related studies on calibrating FD models use aggregated macroscopic quantities. This simplistic approximation is questionable because it involves a certain number of scatters in a non-steady state, which, to some extent, interferes with the ultimate calibration results. To address the underlying issues mentioned earlier, it is crucial to establish a self-consistent calibration framework that uses the same dataset at both the macro- and microscopic levels and which filters out nonstationary macroscopic quantities during FD calibration.

The aim of this paper is to propose a self-consistent calibration framework that can overcome the limitations of existing studies. To maintain macro-to-micro consistency, previous research has attempted to connect these two types of models (Bourrel and Lesort, 2003, Leclercq, 2007, Ma et al., 2011), primarily by using loop detector data to calibrate microscopic models (Rakha and Wang, 2009, Ni et al., 2016). However, owing to the limitations of aggregated macro data in reflecting intricate microscopic traffic phenomena, errors and inconsistencies can occur if parameters estimated by a micro (macro) model are directly substituted for those of the corresponding macro (micro) model (Joueiai et al., 2013). Combined with the present traffic simulation mainly focuses on micro or macro aspects only, without integrating both of them. The proposed physics-informed framework further improves the effectiveness and persuasiveness of traffic simulation, in view of the fact that it takes into account both the critical characteristics of macroscopic traffic stream (free-flow speed, jam density, maximum capacity, etc.) and the vehicle dynamics of microscopic traffic flow. Based on the same trajectory dataset, it incorporates the features of both macro and micro traffic flow and bridges the macro and micro models during the calibration process, which supplements the lack of information caused by focusing on a single level and concentrates on parameters with the identical physics implications. Additionally, another motivation lies in the extensibility and transferability of the calibration framework under different traffic scenarios. The self-consistent calibration method is generalizable to typical CF and FD models rather than limited to a specific model.

To the best of our knowledge, no previous studies have focused on integrating macroscopic traffic characteristics into the calibration of microscopic CF models. This motivated us to propose a self-consistent calibration framework, which provides the following contributions:

- First, we present the concept of self-consistency in the new methodology to address the inconsistent calibration results caused by distinct datasets in separated macro- and microscopic calibrations. The microscopic model and its corresponding macroscopic traffic stream model, as well as the data adopted in the calibration procedures of the CF and steady-state FD models, should be consistent. On this basis, we convert the conventional single-objective problem into a multi-objective optimization problem in the CF calibration problem, which incorporates macro- and micro-level objective functions. How to discover the Pareto frontier (PF) in an intricate multi-objective problem is still an open issue that requires continuous investigation. Based on the White Shark Optimizer (WSO) (Braik et al., 2022), we furthermore designed an improved multi-objective White Shark Optimizer (IMOWSO) in the self-consistent calibration framework.
- Second, to resolve the steady-state traffic flow extracting issue, we proceeded to slice the spatio-temporal trajectory images into equal-sized space-time regions and developed a deep learning methodology for the extraction of space-time regions at steady state based on the attention mechanism, where the deep residual network (ResNet) model (He et al., 2016), the EfficientNet model (Xie et al., 2017) and the state-of-the-art Transformer model (Vaswani et al., 2017) are embedded. In this study, both empirical and synthetic data generated by microscopic simulations were used to demonstrate the applicability and robustness of the proposed physics-informed framework.

¹ In the existing literature, there are various vocabularies for describing the traffic flow that fulfills the steady conditions (such as equilibrium, steady-state, stationary-state, etc.).

² It is particularly important to note here that the macroscopic FD described in this study is the macroscopic representation of traffic flows at the specific location of a road, rather than the network fundamental diagram (NFD) or macroscopic fundamental diagram (MFD) observed by Mahmassani et al. (1984), Geroliminis and Daganzo (2008) and Gu et al. (2018).

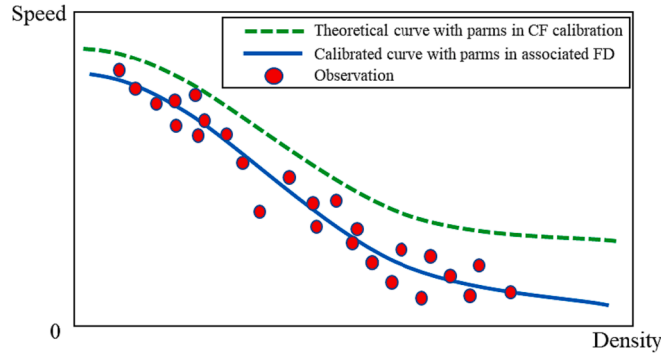


Fig. 1. Demonstration of inconsistent calibration results.

This paper is organized as follows: [Section 2](#) presents the literature review of existing investigation about calibration methods. [Section 3](#) describes the integrated self-consistent calibration framework, consisting of a deep learning-based stationary state detection method. In [Section 4](#), we introduce an improved metaheuristic algorithm to resolve these calibration problems. In [Section 5](#), the proposed framework is implemented to calibrate the CF models in the real world and generate data to validate their applicability and efficiency. Finally, [Section 6](#) concludes the paper.

2. Literature review

2.1. Review of FD calibration

The FD depicts a well-defined relationship between three macroscopic quantities -flow rate, density, and space-mean speed- in a steady state of traffic. It is considered one of the cornerstones of traffic flow theory, and several relevant research achievements have emerged in recent decades, including the Greenshields model ([Greenshields et al., 1935](#)), the Greenberg model ([Greenberg, 1959](#)), and the Underwood model ([Underwood, 1961](#)). Macro- and micro-traffic flow models exhibit explicit inter-transformation relationships, implying that the FD can be rationally obtained from the corresponding CF models under steady-state traffic conditions ([Cheng et al., 2021](#), [Cheng et al., 2024](#), [Pan et al., 2024](#)). Considerable efforts have been devoted to calibrating FD models, and one widely recognized macroscopic traffic flow model is the Gazis-Herman-Rothery (GHR) model ([Chandler et al., 1958](#), [Herman et al., 1959](#), [Gazis et al., 1961](#)), which has many variants that can be obtained by varying parameter values. Other analytical CF models also have consistent macroscopic traffic flow models. However, the collected loop detector data usually exhibit a nonuniform distribution, which can cause calibration problems. To address this issue, [Qu et al. \(2015\)](#) developed a weighted least-squares approach, in which the weight of each observation depends on the adjacent distance. In addition, [Zhang et al. \(2018\)](#) proposed a calibration framework based on the least-squares method that generates uniformly distributed measurement samples from raw data.

To deal with non-i.i.d. noise in FD calibration procedures, data-driven investigations, such as the Gaussian Process (GP), are also commonly applied ([Yuan et al., 2021](#), [Cheng et al., 2022](#), [Würth et al., 2022](#)). To address the under-fitting problem, [Liu et al. \(2022\)](#) promoted the utilization of the Hat Kernel and developed a GP-based traffic flow model that incorporates demand-side and supply-side factors ([Liu et al., 2023](#)). [Bramich et al. \(2022\)](#) provided a comprehensive review of the evolution of the functional form of FD over the past few decades and proposed a universal framework for modelling empirical FD that incorporates several sophisticated noise components ([Bramich et al., 2023](#)).

2.2. Review of CF model calibration

Even the most accurate models cannot fully represent all microscopic traffic flow phenomena using a fixed parameter set ([Treiber and Kesting, 2013a](#)). Thus, individual regional driving preferences and traffic conditions at different times need to be considered when estimating parameters in CF models. Calibration of the CF model based on vehicle trajectory data has recently become a widely investigated complex nonconvex and nonlinear optimization problem. In this process, the decision variables refer to the parameters in the given CF model, and the objective function quantifies the discrepancy between field measurements and estimated outcomes. Calibration usually involves the measurement of performance (MoP) and goodness of function (GoF) and adopting a global optimization algorithm ([Punzo et al., 2012](#)). A comprehensive overview of the three aforementioned components was presented ([Punzo et al., 2021](#)). For instance, Root Mean Square Error (RMSE) ([Xu and Laval, 2020](#), [Montanino et al., 2021](#)), Root Mean Squared Percentage Error (RMSPE) ([Saifuzzaman et al., 2015](#), [Sharma et al., 2019](#)), Theil's Inequality Coefficient U (Theil's U) ([Zhong et al., 2016](#), [Chen et al., 2020](#)), Mean Absolute Error (MAE) ([Milanés and Shladover, 2014](#), [Zheng et al., 2016](#)), and Mean Absolute Percentage Error (MAPE) ([Chong et al., 2013](#), [Zheng et al., 2016](#)) are the most frequently used GoFs in existing calibration methods. The primary MoPs are velocity ([Papathanasopoulou and Antoniou, 2015](#), [Li et al., 2020](#)), intervehicle spacing ([Kurtc and Treiber, 2020](#), [Montanino et al., 2021](#)), acceleration ([Hao et al., 2016](#), [Pei et al., 2016](#)) and speed standard deviation ([Huang et al., 2018](#), [Tian et al., 2019](#), [Alhariqi et al., 2022](#)). In this field, several appropriate optimization algorithms have been applied. Random search algorithms such as the

genetic algorithm (GA) (Jin et al., 2014, Hamdar et al., 2015), simplex method (Ossen and Hoogendoorn, 2008), OptQuest Multistart (Punzo and Simonelli, 2005, Ciuffo and Punzo, 2010), and downhill simplex (Brockfeld et al., 2004, Kim and Mahmassani, 2011) have been widely implemented in the literature to search for optimal solutions, compared with greedy search algorithms, which often converge a local optimum. Li et al. (2016) proposed a combined two-stage algorithm for CF calibration, including global and local searches, while Zhong et al. (2016) developed a novel framework for microscopic calibration based on cross-entropy and sensitivity analyses.

Notably, empirical datasets significantly impact the comparison of the CF model performance. With the continuous advancement of state-of-the-art acquisition technologies, such as cameras (Federal Highway Administration, 2010, Ztd, 2018), drones (Krajewski et al., 2018, SEU, 2019), and global navigation satellite systems (GNSS) (Gunter et al., 2021, Makridis et al., 2021), large amounts of trajectory datasets from the highway and urban sections are readily available for investigating calibration problems. While real-world trajectory data lacks ground truth parameters for validation, synthetic data or generated data produced by the CF model with a predefined parameter set serves as an alternative dataset. In this regard, the global optimum is known, enabling the verification of the efficiency and robustness of the calibration framework, making it the only approach to address an optimization problem (Punzo et al., 2012).

2.3. Several existing CF models and their relevant traffic stream models

Under steady-state conditions, where drivers are in equilibrium and acceleration and speed differences of any vehicle of interest in a specific segment are zero, macroscopic FD models can be derived from a specific microscopic CF model. Our preliminary analysis indicates that the specific macroscopic model incorporates several variables with explicit physical meanings that are shared with the corresponding CF model. It is necessary to note that the microscopic CF model can be transformed between its macroscopic representation by fulfilling the above-mentioned equilibrium state conditions. In this section, we present a concise derivation process for obtaining related FD models. Notably, our proposed modeling and physics-informed calibration framework is generalizable to other typical CF models rather than model-specific. Owing to length constraints, we provide a brief overview of two CF models: the Intelligent Driver Model (IDM) and OVM. In this study, IDM and OVM are selected for demonstration due to their representativeness and superior performance. In previous studies, they are typically chosen as benchmark models to analyze the effectiveness of the physics-informed calibration framework (Hoogendoorn and Hoogendoorn, 2010, Li et al., 2016, Wang et al., 2019).

a. IDM and its macroscopic representation

IDM is one of the most widely used time-continuous models (Treiber et al., 2000), which synchronously considers the desired speed and space headway. The acceleration of the vehicle i of interest is governed by the following expression:

$$a_i(t) = a_{\max} \left[1 - \left(\frac{v_i(t)}{v_0} \right)^\delta - \left(\frac{s^*(v_i(t), \Delta v_{i-1,i}(t))}{s_i(t)} \right)^2 \right] \quad (1)$$

$$s^*(v_i(t), \Delta v_{i-1,i}(t)) = s_0 + \max \left(0, v_i(t) \cdot T_0 - \frac{v_i(t) \cdot \Delta v_{i-1,i}(t)}{2\sqrt{a_{\max} \cdot b}} \right) \quad (2)$$

where $s_i(t) = x_{i-1}(t) - x_i(t) - L_{i-1}$ represents the gap distance of the object vehicle i at time t . L_{i-1} is the length of the vehicle $i-1$, and $\Delta v_{i-1,i}(t) = v_i(t) - v_{i-1}(t)$ denotes the speed difference between the leading vehicle $i-1$ and its follower vehicle i . s^* is the desired gap distance, δ is the free acceleration exponent, and usually the value is taken as 4. a_{\max} represents the maximum acceleration, b denotes the comfortable deceleration, v_0 is the desired speed. s_0 represents the minimum safe gap distance and T_0 is the desired time gap. The last five parameters need to be calibrated. The traffic flow is considered a continuous compressible fluid at the macro level. The traffic flow q , density k , speed v , gap distance s and average vehicle length L are introduced. To derive the relevant macroscopic FD model, the steady-state conditions, $a_i(t) = 0$, $\Delta v_{i-1,i}(t) = 0$, as well as $s + L = 1/k$, can be substituted into Eqs. (1) and (2).

Combined with the classical corollary $q = kv$, the macroscopic representation of the IDM can be obtained after omitting the subscripts i and $i-1$, as follows:

$$k = \frac{\sqrt{1 - (v/v_0)^\delta}}{s_0 + vT_0 + L\sqrt{1 - (v/v_0)^\delta}} \quad (3)$$

$$q = \frac{v\sqrt{1 - (v/v_0)^\delta}}{s_0 + vT_0 + L\sqrt{1 - (v/v_0)^\delta}} \quad (4)$$

b. OVM and its representation

The OVM (Bando et al., 1995) occupies a significant role in CF-related research. Mathematically, the acceleration of the i^{th} object vehicle can be defined as:

$$a_i(t) = \frac{v_{opt}(s_i(t)) - v_i(t)}{\tau} \quad (5)$$

$$v_{opt}(s) = v_0 \frac{\tanh(s_i(t)/\Delta s - \beta) + \tanh\beta}{1 + \tanh\beta} \quad (6)$$

where $v_{opt}(s)$ represents the optimal speed, which depends on the spacing headway; τ denotes the adaptation time; v_0 is the desired speed; β and Δs are the form factors and transition width in the OVM's optimal speed function, respectively, and need to be calibrated.

Similarly, to obtain the macroscopic FD model for the OVM, we substitute the traffic flow equilibrium state conditions ($a_i(t) = 0$, $\Delta v_{i-1,i}(t) = 0$, and $s + L = 1/k$) into Eqs. (5) and (6). Considering that the OVM's theoretically consistent FD expression differs from the IDM, the density k is the independent variable in the density-speed expression. Consequently, the macroscopic representation of OVM, as shown in Eq. (7) and (8):

$$v = v_0 \frac{\tanh\left(\frac{1-kL}{k\Delta s} - \beta\right) + \tanh\beta}{1 + \tanh\beta} \quad (7)$$

$$q = kv_0 \frac{\tanh\left(\frac{1-kL}{k\Delta s} - \beta\right) + \tanh\beta}{1 + \tanh\beta} \quad (8)$$

c. Highlights of some deficiencies

Theoretically, the CF models share several parameters with the same physical significance as the associated macroscopic FD models. We denote the parameters in the microscopic CF model and its theoretically consistent FD macroscopic model as θ_{micro} and θ_{macro} , respectively. Then, the shared parameter $\tilde{\theta}$ existing in both the microscopic CF model and macroscopic FD model can be expressed as:

$$\tilde{\theta} = \theta_{micro} \cap \theta_{macro} \quad (9)$$

For instance, as discussed in this section, $\tilde{\theta} = [v_0, s_0, T_0]^T$ in the IDM, and $\tilde{\theta} = [v_0, \Delta s, \beta]^T$ in OVM. However, existing approaches implement these common parameters from a one-sided macro or micro perspective, resulting in inconsistent calibration outcomes. To resolve these deficiencies, we propose a self-consistent framework for calibrating the common parameter $\tilde{\theta}$ in the next section.

3. Integrated self-consistent calibration framework

In this section, we develop an integrated self-consistent methodology combining the macro- and micro-traffic flow models during calibration. The self-consistency idea is mainly reflected in two aspects. Firstly, we harmonize the calibration results of parameters with identical physical implications in macro-micro models. Secondly, we adopt the same trajectory dataset in the calibration procedure for the macro-micro models, which largely eliminates errors arising from separate acquisition processes. In addition, as previously discussed, the FD model is formulated based on steady-state traffic conditions. Therefore, we propose a supervised deep-learning-based steady-state detection method to identify whether the vehicle stream in a space-time region is in a steady state. The proposed framework can be used in various traditional CF models such as the GHR, Gipps (Kometani and Sasaki, 1959, Gipps, 1981), OVM, and IDM provided that the steady state conditions are met, owing to its sound scalability and interpretability. As shown in Fig. 2, the proposed methodology incorporates three modules. The implementation of each component is elaborated in this section, and the

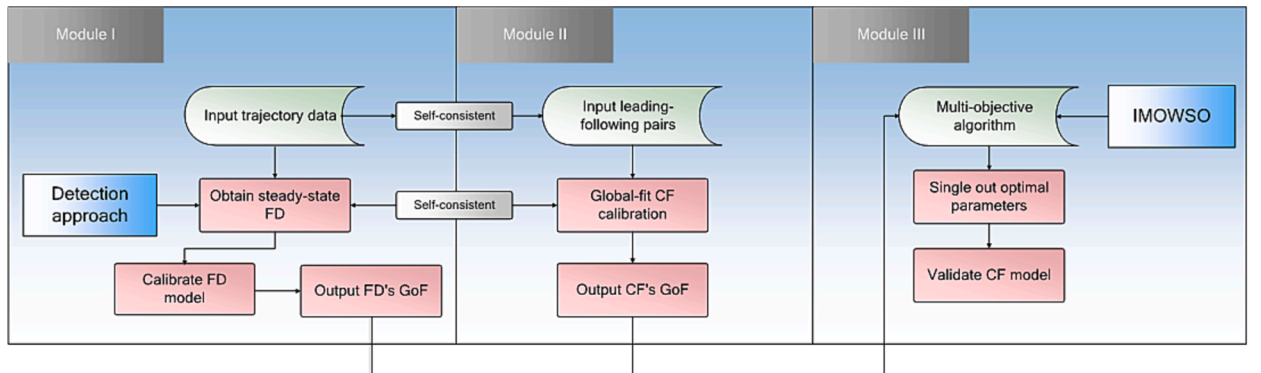


Fig. 2. Main structure of the self-consistent calibration framework.

notations used in this section are summarized as follows:

Notation	Explanation
ω	Space-time region
$Q(\omega)$	Flow in region ω
$K(\omega)$	Density in region ω
$V(\omega)$	Speed in region ω
θ_{macro}	Vector of variables in macro model
θ_{micro}	Vector of variables in micro model
Ψ_1	Solution space of θ_{macro}
Ψ_2	Solution space of θ_{micro}
T^*	Updated time interval
T_{\min}	Minimum following time
s_{\max}	Maximum headway during following
LB	Low bound of decision variables
UB	Upper bound of decision variables
τ	Reaction time
$v_i(t \theta_{\text{micro}})$	Velocity of the object i vehicle at time t
$x_i(t \theta_{\text{micro}})$	Position of the object i vehicle at time t
$\Delta x_{i-1,i}(t)$	Spacing headway difference at time t
$\Delta v_{i-1,i}(t)$	Speed difference at time t
$\hat{x}_i(t)$	Observed position of vehicle i at time t
$\hat{v}_i(t)$	Observed velocity of vehicle i at time t
$\bar{\theta}$	Set of candidate solutions
θ^*	Optimal solution

3.1. Macroscopic FD calibration (Module I)

A. Steady-state detection method.

The fundamental diagram represents the mathematical relationship between macroscopic speed-density, that is, $\hat{v}(\cdot)$. As shown in Fig. 3, the traffic state, including flow, density, and speed can be determined from the complete space-time trajectory by dividing it into uniform space-time regions ω , lengths in space-time regions $L(\text{m})$ and $T(\text{s})$.

We propose a steady-state detection architecture comprising three blocks, as shown in Fig. 4 (a). The steady state detection framework is developed for fine-grained recognition tasks, where the difference between space-time regions is characterized by the local region information, and is modified from classical Bilinear CNN model (Lin et al., 2015). Based on the typical Bilinear CNN model, multiple extractors are designed to accurately obtain the local information of each time-space region, and it should be noted that the inputs of all three extractors are the same, i.e., a single time-space region. The input images of the uniformly divided space-time region were first resized to the same identical dimension before passing through the extractors. *ResNet*, *EfficientNet* and *Transformer* models, which are commonly used in image recognition and classification, were employed as blocks to extract feature information from the space-time region. *ResNet* is a variant on CNN that tackles the problem of gradient vanishing by incorporating residual blocks (He et al., 2016), each of which contains multiple convolutional layers and a “shortcut” across a number of layers, allowing information to be passed directly from one of the previous layers to one of the later ones. Besides, *EfficientNet* introduces the “Compound Scaling” method and the structure of “Mobile Inverted” to decrease the number of parameters (Xie et al., 2017), enhance feature representation and provide superior performance and effectiveness relative to the CNN model by utilizing computational resources more productively. The two above-mentioned models are representatives of the advanced models widely used to address the image recognition problems (Chen et al., 2019, Kaur et al., 2019, Sarwinda et al., 2021, Yang et al., 2021, Gong et al., 2023). *Transformer* model is a state-

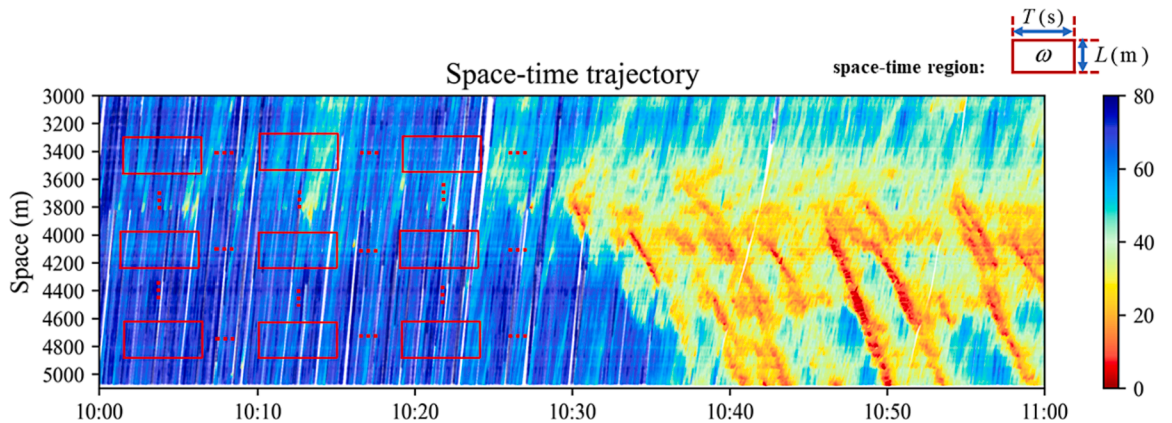
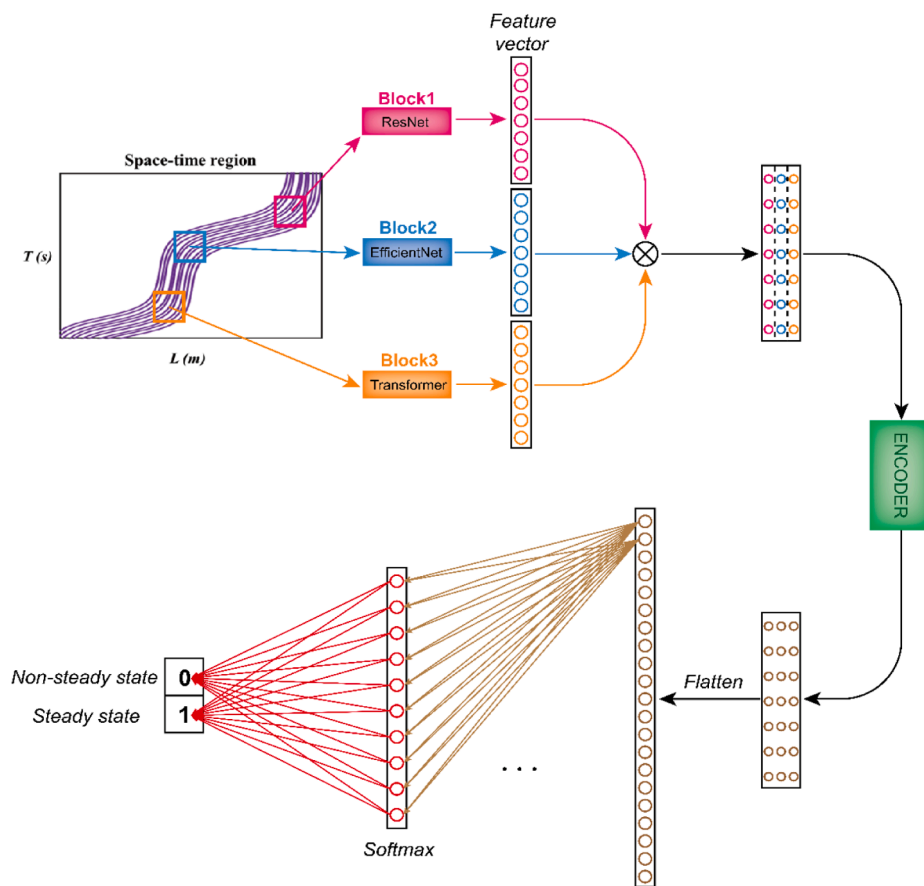
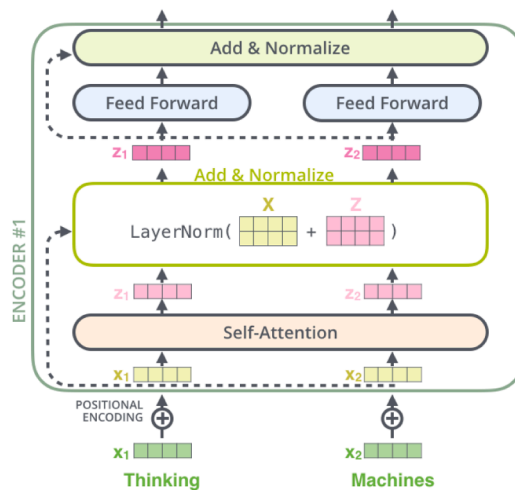


Fig. 3. Representation of space-time regions in trajectory.



(a)



(b)

Fig. 4. Deep-learning-based detection approach. (a) Framework of Steady-state Detection Model. (b) Main Architecture of the Encoder (Source from Alammari, 2018).

of-the-art neural network architecture based on a self-attention mechanism, and has been proven in existing studies to have excellent performance in the field of image recognition (Liu et al., 2021, Messina et al., 2021, He et al., 2022). The output feature vectors from these three blocks were stacked to create an image descriptor that could be further processed by an attention-based encoder, which is an important component in our steady-state detection method and can be regarded as a resource allocation scheme to address the information overload problem. In the case of limited computational power, it can utilize the limited computational resources to extract more important information, commonly used in issues related to image classification (Zheng et al., 2017, Qu et al., 2020, Niu et al., 2021, He et al., 2022). The encoder was designed to effectively process and capture non-steady information and consisted of positional encoding, multi-head attention, and a residual block. The primary architecture of the encoder is illustrated in Fig. 4(b). The final prediction of the time-space region's state is accomplished by integrating the three feature vectors. The resulting output was then processed through the classification layer to predict the state of the specific space-time region as either non-steady or steady, determining whether it would be incorporated into subsequent macroscopic FD calibration.

B. Calibration of FD models using empirical trajectory.

According to Edie's definition of the traffic state in a space-time region, ω (Edie, 1963),

$$Q(\omega) = \frac{\sum_{n \in N(\omega)} d_n(\omega)}{|\omega|} \quad (10)$$

$$K(\omega) = \frac{\sum_{n \in N(\omega)} t_n(\omega)}{|\omega|} \quad (11)$$

$$V(\omega) = \frac{\sum_{n \in N(\omega)} d_n(\omega)}{\sum_{n \in N(\omega)} t_n(\omega)} = \frac{Q(\omega)}{K(\omega)} \quad (12)$$

where $Q(\omega)$, $K(\omega)$ and $V(\omega)$ represent flow, density and speed in region ω , respectively. $N(\omega)$ denotes a set of all the vehicles involved in ω . $d_n(\omega)$ and $t_n(\omega)$ represent the total distances traveled by all vehicles in ω and the total time spent by all vehicles in ω and $|\omega|$ is the area of ω .

Consider a list of P (number of divided time-space regions) observations of the density-speed pair acquired from the empirical trajectory and pass the steady-state detection, speed $\mathbf{v} = [v_1, v_2, \dots, v_P]^T$, and density $\mathbf{k} = [k_1, k_2, \dots, k_P]^T$. Regarding the calibration of the FD, the objective is to minimize the squared distance between the field measurement v_i and the estimated speed obtained from the speed-density relationship $\hat{v}(k_i)$. *RMSE* is adopted as the GoF in the macro calibration, which can be formulated as follows:

$$\min_{\theta_{\text{macro}} \in \Psi_1} f_1 = \sqrt{\frac{1}{P} \cdot \sum_{i=1}^P (v_i - \hat{v}(k_i | \theta_{\text{macro}}))^2} \quad (13)$$

where θ_{macro} is the vector of the speed-density function parameters and Ψ_1 is the solution space of θ_{macro} . v_i and $\hat{v}(k_i | \theta_{\text{macro}})$ denote the observed and predicted speeds related to the density k_i , respectively.

3.2. Microscopic CF calibration (Module II)

A. Filtering criteria of leading-following pairs.

The primary basis for the CF models in the calibration procedure was to extract the leading-following pairs. It is common practice to simulate a series of state variations of the rear vehicle by fixing the information of the front vehicle, including its position and speed. The selection of leading-following pairs was crucial for accurate calibration results. Most CF models involve parameters that describe the driving state, where the desired speed characterizes the free-flow state, and the minimum headway depicts the congestion state. If the selected vehicle trajectory data does not include the aforementioned states, the resulting calibration may lack physical validity and be deemed implausible. To ensure the rationality and precision of the calibration results, we recommend using the following criteria for screening the following pairs before conducting the microscopic calibration:

- The following time should exceed T_{\min} (s) with no lane changing during the period
- The maximum spacing between leading and following vehicles is less than s_{\max} (m)
- The following period should include at least one acceleration and deceleration behavior to reflect related parameters, such as maximum acceleration and comfortable deceleration.

where T_{\min} and s_{\max} are the predefined parameters in the specific calibration problem.

B. Calibration of CF models based on selected trajectory.

Generally, the objective function of CF model calibration is a multivariate, nonlinear, and non-convex optimization problem, which aims to minimize the difference between the field measurement and simulated result. It can be formulated as follows (Sharma et al., 2019):

$$\min f(MoP^{obs}, MoP^{sim}) \quad (14)$$

where MoP^{obs} and MoP^{sim} represent the observed and simulated MoP, respectively. $MoP^{sim} = F(\theta_{micro})$ and θ_{micro} are vectors of the CF model parameters (the number of parameters varies according to the CF model adopted), subject to $LB \leq \theta_{micro} \leq UB$. LB and UB denote the vectors of the lower and upper bounds for the parameters to be calibrated in θ_{micro} , respectively. $F(\cdot)$ is the CF model that displays microscopic car-following behavior, and $f(\cdot)$ is the GoF.

Based on the investigated CF model, the following delay differential equations were adopted to obtain the acceleration and speed profiles of the object vehicle i :

$$\dot{v}_i(t|\theta_{micro}) = \frac{dv_i(t|\theta_{micro})}{dt} = F(\Delta x_{i-1,i}(t-\tau), v_i(t-\tau), \Delta v_{i-1,i}(t-\tau)|\theta_{micro}) \quad (15)$$

$$\dot{x}_i(t|\theta_{micro}) = \frac{dx_i(t|\theta_{micro})}{dt} = v_i(t|\theta_{micro}) \quad (16)$$

where $v_i(t|\theta_{micro})$ represents the velocity of the i^{th} vehicle at time t . $x_i(t|\theta_{micro})$ is the position of the object i vehicle at time t and τ is the reaction time. $\Delta x_{i-1,i}(t)$ and $\Delta v_{i-1,i}(t)$ denote the spacing headway and speed difference between the leading vehicle $i-1$ and its follower object vehicle i , respectively.

Calibration methods can be classified into two categories as local and global fit approaches (Treiber and Kesting, 2013b). In contrast to local-fit methods, global-fit approaches are often used in which the movements of the following vehicle are simulated sequentially based on the initial state, including the follower's velocity, spacing, and speed difference between the leader and the following vehicle. The discrete-time model used in the global fit is updated with a fixed time interval and is suitable for traffic simulation (Li et al., 2016), which can be formulated as follows:

$$v_i(t+T|\theta_{micro}) = v_i(t|\theta_{micro}) + \dot{v}_i(t|\theta_{micro}) \cdot T \quad (17)$$

$$x_i(t+T|\theta_{micro}) = x_i(t|\theta_{micro}) + v_i(t|\theta_{micro}) \cdot T + \frac{1}{2} \cdot \dot{v}_i(t|\theta_{micro}) \cdot T^2 \quad (18)$$

where T represents the fixed update time interval.

It is well-recognized that the calibration results for a specific CF model are quite different when compared with the simulated results (time series of the object vehicle's speed or spacing) from Eqs. (15)–(18) through the operation of multiple MoPs and GoFs.

$$\min_{\theta_{micro} \in \Psi_2} f_2 = \sqrt{\frac{1}{N} \cdot \sum_{j=1}^N [x_i(jT|\theta_{micro}) - \hat{x}_i(jT)]^2} \quad (19)$$

$$\min_{\theta_{micro} \in \Psi_2} f_3 = \sqrt{\frac{1}{N} \cdot \sum_{j=1}^N [v_i(jT|\theta_{micro}) - \hat{v}_i(jT)]^2} \quad (20)$$

where Ψ_2 is the feasible domain of θ_{micro} . $x_i(jT|\theta_{micro})$ and $\hat{x}_i(jT)$ represent the simulated and observed positions at a time $t = jT$, respectively. Similarly, $v_i(jT|\theta_{micro})$ and $\hat{v}_i(jT)$ denote the simulated and observed speeds of the vehicle i at time $t = jT$, respectively.

3.3. Macro-micro integrated optimization (Module III)

After completing the above macro- and micro-level objective functions, the next step is the calibration process and optimal parameter selection. By combining Eqs. (13), (19) and (20), we can derive the objective function of the integrated macro-micro traffic flow calibration problem in Eq. (21). To ensure parameter consistency, we denote $\theta = \theta_{macro} \cup \theta_{micro}$ and each objective function shares the same search domain Ψ .

$$\min_{\theta \in \Psi} \begin{cases} f_1 = \sqrt{\frac{1}{P} \cdot \sum_{i=1}^P (v_i - \hat{v}(k_i|\theta_{macro}))^2} \\ f_2 = \sqrt{\frac{1}{N} \cdot \sum_{j=1}^N [\hat{x}_i(jT) - x_i(jT|\theta_{micro})]^2} \\ f_3 = \sqrt{\frac{1}{N} \cdot \sum_{j=1}^N [\hat{v}_i(jT) - v_i(jT|\theta_{micro})]^2} \end{cases} \quad (21)$$

It is well established that multi-objective optimization algorithms trade off conflicting objective functions and present multiple non-dominated solutions, which are called Pareto optimal fronts. To address the above integration optimization problem, we propose the IMOWSO multi-objective algorithm, further details can be found in Section 4. Module III goes through the IMOWSO algorithm and obtains the PF $\vec{\theta} = [\theta_1, \theta_2, \dots, \theta_n]$, which contains the n candidate solutions. In practical applications, although the solutions in the

optimal set may appear equivalent, they are not necessarily equivalent for the decision maker. Therefore, it is necessary to select the solution with the highest comprehensive satisfaction. This issue also arises when calibrating macro–micro traffic models. In such cases, it is important to identify a solution that exhibits superior fitting performance on both the microscopic CF model and its corresponding macroscopic representation, particularly when there are no explicit weights assigned between multiple objectives. Consequently, the grey relational analysis (Wang and Rangaiah, 2017) was adopted to obtain the optimal solution θ^* , the principle of which is to acquire the similarity measure between each solution and the best reference solution, that is, a vector consisting of the best values for each objective function. In this section, we consider the PF of n candidate solutions and m objective functions as an example.

$$F_{ij}(\theta_i) = \frac{f_{ij}(\theta_i) - \min_{k \in n} f_{kj}(\theta_k)}{\max_{k \in n} f_{kj}(\theta_k) - \min_{k \in n} f_{kj}(\theta_k)} \quad (22)$$

$$F_j^+ = \max_{k \in n} F_{ij}(\theta_k) \quad (23)$$

$$\Delta I_{ij}(\theta_i) = |F_j^+ - F_{ij}(\theta_i)| \quad (24)$$

where f_{ij} denotes the value of the i^{th} candidate solution in the j^{th} objective function.

The following detailed procedure was used to calculate the gray relational coefficient (GRC). The larger the GRC value corresponding to the solution, the higher its aggregated satisfaction in each objective function. It can be formulated as follows:

$$GRC(\theta_i) = \frac{1}{n} \sum_{j=1}^m \frac{\Delta \min + \Delta \max}{\Delta I_{ij}(\theta_i) + \Delta \max} \quad (25)$$

$$\theta^* = \arg \max_{\theta_k \in \theta} GRC(\theta_k) \quad (26)$$

where $\Delta \max = \max_{k \in n, j \in m} \Delta I_{kj}(\theta_k)$ and $\Delta \min = \min_{k \in n, j \in m} \Delta I_{kj}(\theta_k)$.

4. Improved solution algorithm for calibration

Note that the macro- and microscopic traffic flow models were theoretically consistent, and achieving self-consistent calibration of analytical macro–micro traffic flow models based on the same trajectory data, which have multiple objective functions from both the macro- and micro-aspects, remains challenging. Multi-objective optimization aims to achieve a trade-off between conflicting objectives concurrently and obtain Pareto-efficient solutions, the objective points of which are located on the PF, as opposed to single-objective optimization. To overcome the difficulty of integrated optimization, we propose an IMOWSO algorithm that guarantees

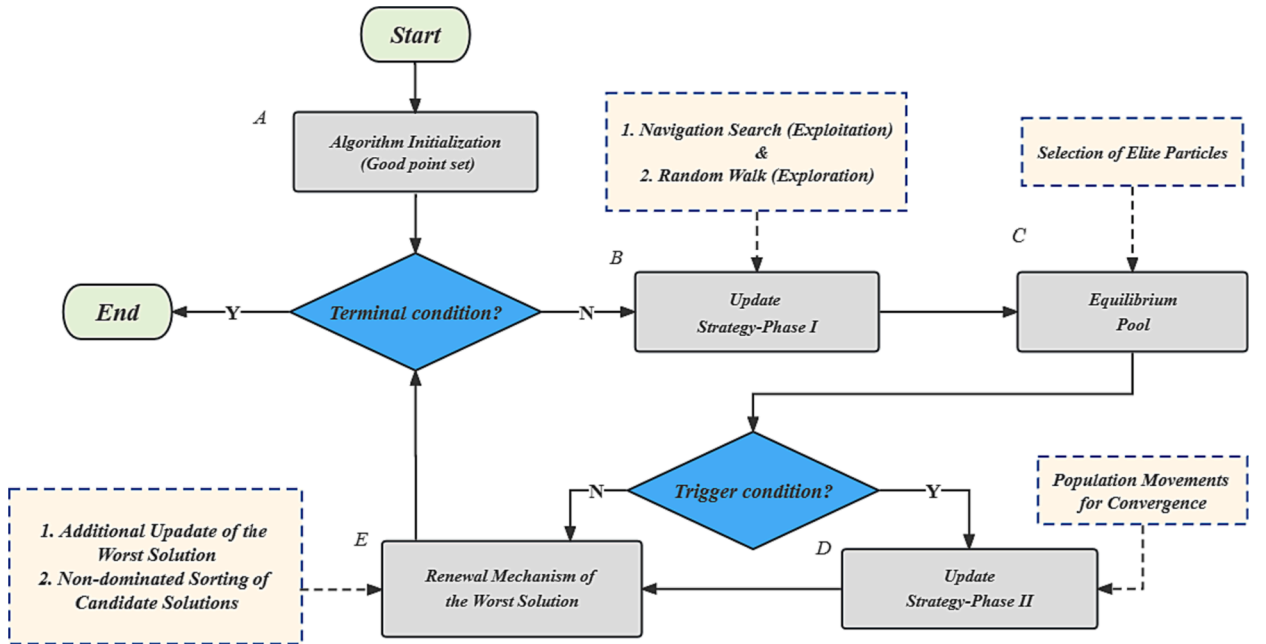


Fig. 5. Schematic representation of IMOWSO algorithm. (The grey boxes indicate the procedures of the IMOWSO algorithm, and the yellow dotted squares represent additional explanations of each particular step).

the calibration results of microscopic traffic flow and macroscopic FD models simultaneously. The selected benchmark, WSO, is a state-of-the-art metaheuristic algorithm for tackling global optimization problems that involve continuous search domains. The core idea of the algorithm stems from imitating the white shark tracking prey, where ‘prey’ refers to the global optimum. It is noteworthy that the WSO adequately balances exploration and exploitation when searching for an optimal solution. Current WSO is limited in its ability to handle multiresolution problems and eliminate the dilemma of falling into a local optimum. In this study, we improved the WSO to apply it to solving multi-objective optimization problems. A schematic representation of the IMOWSO is shown in Fig. 5. Each principal step is marked with the numbers A–E. A common methodology for transforming the WSO algorithm into a multi-objective WSO (MOWSO) is incorporating a non-dominated ranking technique such as fast non-dominated sorting in NSGA-II (Deb et al., 2002). Based on the MOWSO algorithm, a modified initialization strategy, equilibrium pool principle, and update mechanism of the worst solution were introduced in the IMOWSO.

The notations used in this section are summarized as follows.

Notation	Explanation
n	Population size
d	Number of decision variables
w_k	Set of candidate solutions at iteration k
R	Initial set of generated coefficients
$S(\cdot)$	Good Point Set
v_k^i	Speed for the i^{th} solution at iteration k
w_{gbest_k}	Global best solution by iteration k
Idx	Index vector
μ	Control coefficients
δ_{min}	Lower control factor
δ_{max}	Upper control factor
r	Predetermined frequency
K	Maximum iteration number
mv	Movement impact factor
\vec{C}_{EP}^k	Equilibrium pool at iteration k
\tilde{w}_i^k	Top solution on the i^{th} objective function at iteration k
\tilde{w}_{avg}^k	Average solution in equilibrium pool at iteration k
\tilde{w}_k	Elite candidate from equilibrium pool at iteration k
\bar{D}_w	Distance between current and global best candidate solution
w_k^i	Desired update position of i^{th} solution at iteration k
w_{best}^k	Candidate solution having highest degree of fit coefficient at iteration k
w_{worse}^k	Candidate solution having lowest degree of fit coefficient at iteration k
P^*	Pareto-optimal front
P	PF obtained from the optimization algorithm
$Leb(\cdot)$	Lebesgue Measure
Ref	Set of reference points

4.1. Algorithm structure

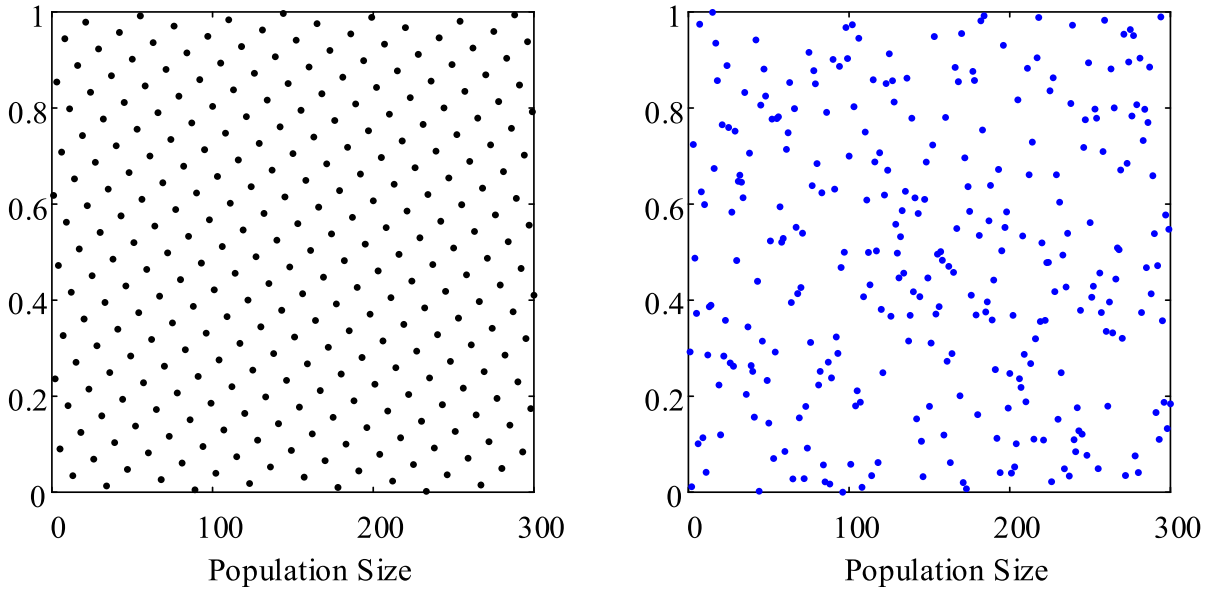
A. Initialization of IMOWSO.

Similar to other population-based optimization algorithms, IMOWSO involves particle initialization in the search region. Consequently, when dealing with problems that include multiple decision variables, we can represent the initial population explicitly using a matrix:

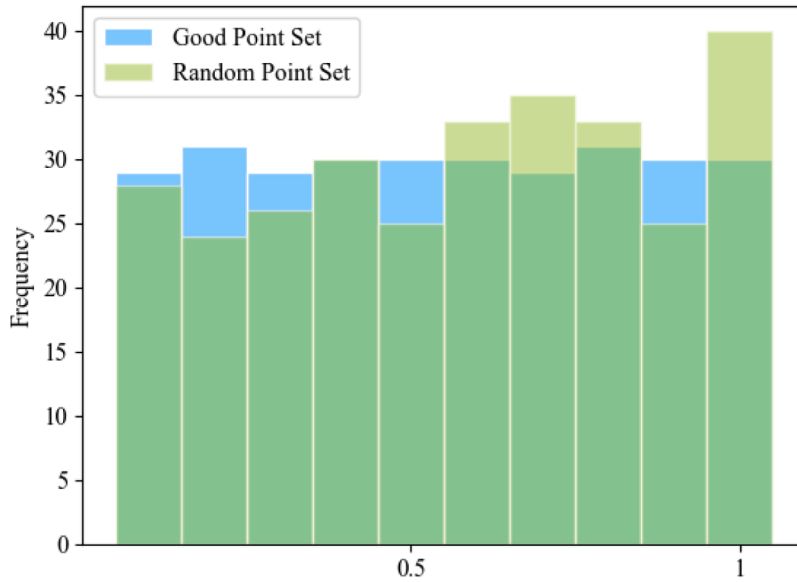
$$w = \begin{bmatrix} w_1^1 & w_2^1 & \dots & \dots & w_d^1 \\ w_1^2 & w_2^2 & \dots & \dots & w_d^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_1^n & w_2^n & \dots & \dots & w_d^n \end{bmatrix} \quad (27)$$

where n denotes the population size, d represents the number of decision variables, w is the location of white sharks in the search domain, in other words, vector of parameter values. Each row in the matrix w stands for a candidate solution that contains d variables.

Generally, random sampling with a uniform distribution is used as a standard initialization practice. This is because initializing the parameter set with a uniform distribution can significantly increase the population diversity and improve the exploration of objective function values. Therefore, we utilized the initialization process based on the Good Point Set principle proposed by Hua and Wang (1978). This principle involves generating uniform initial populations in a high-dimensional space. As can be seen from Fig. 6, given the same population size ($n = 300$) and range of decision variables (from 0 to 1), the initial population generated by the good-point set-based method exhibits a more even distribution than that generated by the random method. This improved distribution leads to better diversity and uniformity in population traversal, thereby preventing the algorithm from getting stuck in suboptimal solutions. Mathematically, the good-point set can be expressed as:



(a)



(b)

Fig. 6. (a) 1D initial population distribution with population size $n = 300$ of Good Point Set (left) and Random Point Set (right); (b) Frequency distribution histogram.

$$R = \{2\cos(2\pi t/p), 1 \leq t \leq d\} \quad (28)$$

$$S(i) = \{(\{r_1 \cdot i\}, \{r_2 \cdot i\}, \dots, \{r_d \cdot i\}), 1 \leq i \leq n, r_i \in R\} \quad (29)$$

where R is the initial set of generated coefficients and p is the minimum prime number that satisfies the condition that $(p-3)/2 \geq d$, $S(i)$ stands for the i^{th} , and $\{\cdot\}$ represent the decimal part.

Furthermore, by mapping the Good Point Set S onto the research area for initialization, we obtained the following initial population:

$$w_i = (UB - LB) \cdot S(i) + LB \quad (30)$$

where w_i denotes the i^{th} candidate solution, UB and LB represent the upper/lower bounds for the decision variables. Details on the good-point set are provided by [Hua and Wang \(1978\)](#).

B. Update strategy -Phase I.

After initialization, it is necessary to continuously iterate and improve the population to approach the PF. It is worth noting that the relevant predetermined coefficients and constants in IMOWSO are references from Braik et al. ([Braik et al., 2022](#)). In the given k^{th} iteration, there are two stages of population update. The update strategy in the first stage involves optimizing the speed of each individual and its corresponding movement. More details of the procedure are presented in **Appendix A1**.

C. Equilibrium pool.

In this procedure, we introduce the concept of an equilibrium pool, as proposed by [Faramarzi et al. \(2020\)](#) based on the WSO algorithm. The equilibrium pool consists of elite individuals selected to guide the positional movement of the overall population and enhance population diversity. The number of elite candidates selected depends on the number of objective functions involved in a particular optimization problem. For example, if an optimization problem involves m objective functions, m particles with the best performance on each of the single objective functions are selected into the equilibrium pool. In addition, an average particle based on the m candidates mentioned above is also included. Thus, the equilibrium pool, composed of $(m + 1)$ candidates in iteration k is defined as follows:

$$\vec{C}_{EP}^k = \{\tilde{w}_1^k, \tilde{w}_2^k, \dots, \tilde{w}_m^k, \tilde{w}_{avg}^k\} \quad (31)$$

where \tilde{w}_i^k stands for the candidate solution that is the top performer on the i^{th} objective function ($1 \leq i \leq m$) at iteration k , and \tilde{w}_{avg}^k represents the average candidate. It is worth noting that the probabilities of the individual particles fetched from the equilibrium pool were equal.

D. Update strategy -Phase II.

Unlike the update strategy in the first stage, the update in Phase II is not necessarily executed in every iteration and has some degree of randomness. The strategy for this phase was inspired by the schooling behavior of white sharks, which move towards the white shark closest to the best predator, that is, the global optimal solution. More information of the step is presented in **Appendix A2**.

E. Renewal mechanism of the worst solution

In each iteration of the algorithm, it is crucial to conduct additional updates for the candidate solution in the population based on the shuffled frog-leaping algorithm (SFLA) developed by [Eusuff et al. \(2006\)](#). This further accelerates the convergence rate. An information entropy-based technique for the order of preference by similarity to the ideal solution (TOPSIS) ([Hwang and Yoon, 1981](#)) was adopted to obtain the best candidate solution w_{best} and worst candidate solution w_{worst} (More details are in **Appendix A3**). Furthermore, we incorporated an update mechanism for the worst candidate solution to improve the quality and convergence speed of the population. The steps involved are as follows:

$$\bar{w}^k = w_{worst}^k + \text{rand} \cdot (w_{best}^k - w_{worst}^k) \quad (32)$$

where \bar{w}^k is the updated candidate solution, w_{best} is the candidate solution with the highest degree of fit coefficient at iteration k , and w_{worst} is the candidate solution with the lowest degree of fit coefficient at iteration k . In principle, for the updated candidate solution \bar{w} , if \bar{w} is superior to w_{worst} , substitute \bar{w} for w_{worst} ; alternatively, replace w_{worst} with w_{best} . Subsequently, particles that exceeded the upper and lower boundaries were assessed and corrected. Next, we employed fast non-dominated sorting in NSGA-II ([Deb et al., 2002](#)) to re-rank the populations using the proposed algorithm. The detailed pseudocode for the IMOWSO algorithm is presented in **Appendix A4**.

4.2. Illustrate examples

The IMOWSO algorithm was tested using typical benchmark functions.³ To provide an objective evaluation of its performance, several representative multi-objective optimization algorithms were included in the assessment. These algorithms were Multiple Objective Particle Swarm Optimization (MOPSO) ([Coello and Lechuga, 2002](#)), NSGA-II ([Deb et al., 2002](#)), a Multi-objective Evolutionary Algorithm based on decomposition (MOEA/D) ([Zhang and Li, 2007](#)) and comparison group- MOWSO without improvement measures, which refers to the equilibrium pool strategy and renewal mechanism of the worst solution. In order to further investigate the validity of the Good Point Set, we set up another simplified comparison group – IMOWSO*, represents the absence of the Good Point Set.

The inverse generational distance (IGD) metric ([Zhang et al., 2008](#)) was used to evaluate the performance of the algorithms. This provides a simultaneous assessment of the convergence and diversity of the population at a low computational cost. Mathematically, the IGD metric in Eq. (33):

³ The typical benchmark functions include Walking Fish Group (WFG) problems ([Huband et al., 2006](#)), Unconstrained Functions (UF) ([Zhang et al., 2008](#)) and Composition Functions (CF) ([Liang et al., 2005](#)).

$$IGD(P, P^*) = \frac{1}{|P^*|} \sum_{i=1}^{|P^*|} \min_{j=1}^{|P|} \|p_i^* - p_j\|_2 \quad (33)$$

where P^* denotes the Pareto-optimal front and P represents the set of final non-dominated points obtained from the specific algorithm.

Additionally, the evaluation criterion for measuring the performance of a solution set is the hypervolume (HV) (Zitzler and Thiele, 1999). HV is a one-dimensional metric that conforms to the concept of Pareto dominance, and it provides a comprehensive measure of both the convergent properties and the variety of a given solution set. The larger the value of HV, the superior the solution set is, which can be formulated as follows:

$$HV(P, R) = Leb\left(\bigcup_{i=1}^{|P|} [p_i, \text{Ref}]\right) \quad (34)$$

where $Leb(\cdot)$ denotes the Lebesgue Measure, $[p_i, \text{Ref}]$ represents the HV composed of the reference point set Ref and i^{th} solutions in the solution set. The reference point set Ref was kept consistent across all algorithm trials meaning that it comprised the highest values obtained by each algorithm in every dimension.

Tables 1 & 2 present the optimization results (IGD and HV) and average computation time of several algorithms over 30 independent runs respectively on the WFG test problems, which include differentiability or non-differentiability, single or multiple peaks, and convex or non-convex PF. The maximum number of iterations K , the population size n and the repository size n' are configured as 200, 500 and 200, respectively. Appendices B & C provide the hardware configuration, parameter settings of the aforementioned global search algorithms and optimization results for UF problems (Table C1-C2) and CF problems (Table C3-C4). To better evaluate the performance of IMOWSO, we compared it with other algorithms (including IMOWSO*) in terms of the best and worst IGD and HV values. Based on the IGD metric, the IMOWSO algorithm ranked #1 in the best and worst cases out of 30 independent runs for 72 % of the complex multidimensional benchmark. As for the HV metric, the IMOWSO algorithm ranked #1 in the best case and worst case out of 30 independent runs for 66 % of the complex multidimensional benchmark. With regard to the computation time, the proposed IMOWSO in general presents the promising efficiency, maintaining the top 2 on more than 80 % of the complicated benchmark (after excluding the simplified comparison group – MOWSO and IMOWSO*). These outcomes form the basis for the self-consistent calibration framework.

5. Case study

5.1. Empirical trajectory experiment

A. Data description and parameter setting.

In this section, we validate the performance of proposed physics-informed calibration framework using a high resolution vehicle trajectory dataset (0.1 s) captured by video cameras, namely the Zen Traffic Data (ZTD) (Ztd, 2018). Large-scale detector data using image sensing technology were collected on Hanshin Expressway Route 11 (Ikeda Route), Osaka City, Japan. Fig. 7(a) illustrates the heavily trafficked segment near the intersection, consisting of two lanes in each direction, and complex traffic conditions, including “s-curves” and sagging. The expressway segment is 2 km long, and the traffic flow is monitored by several vehicle detectors, covering

Table 1

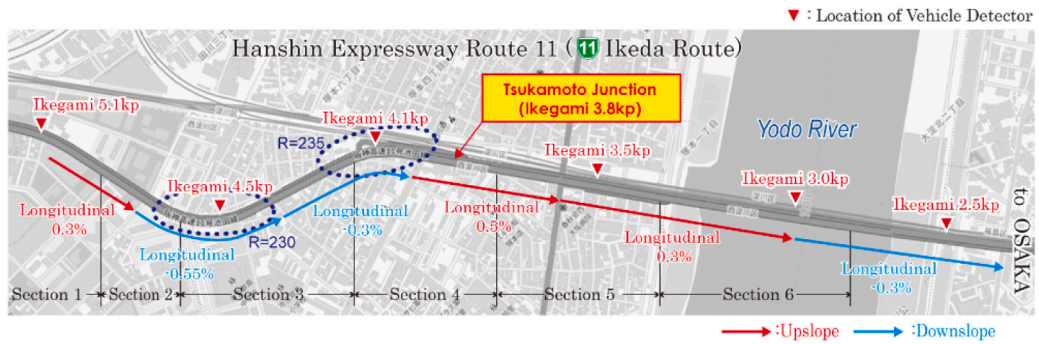
IGD values (Best and Worst cases) and average computation time (ACT) obtained for several optimization algorithms on WFG problems over 30 independent runs.

Benchmark Functions		MOPSO	NSGA-II	MOEA/D	MOWSO	IMOWSO*	IMOWSO
WFG-1	Best case	4.9768×10^{-2}	2.2712×10^{-1}	2.5219×10^{-1}	4.5846×10^{-2}	4.1844×10^{-2}	4.0109×10^{-2}
	Worst case	1.3429×10^{-1}	2.4379×10^{-1}	2.7774×10^{-1}	5.2783×10^{-2}	4.2139×10^{-2}	4.0831×10^{-2}
	ACT (s)	18.5397	20.1789	11.7856	19.7513	19.9863	20.0297
WFG-2	Best case	5.5487×10^{-2}	8.4713×10^{-2}	5.6521×10^{-2}	4.5213×10^{-3}	3.9941×10^{-2}	3.8181×10^{-2}
	Worst case	7.0554×10^{-2}	8.6662×10^{-2}	7.9736×10^{-2}	4.8876×10^{-3}	4.1681×10^{-2}	4.0677×10^{-2}
	ACT (s)	15.4569	17.5376	10.8793	16.4897	16.8763	16.9802
WFG-3	Best case	7.3833×10^{-2}	2.6067×10^{-1}	7.9861×10^{-2}	7.5891×10^{-2}	7.2911×10^{-2}	7.4484×10^{-2}
	Worst case	7.7981×10^{-2}	3.6992×10^{-1}	1.0461×10^{-1}	7.8772×10^{-2}	7.6012×10^{-2}	7.5757×10^{-2}
	ACT (s)	18.9578	21.0457	12.1635	19.8637	20.1467	20.2019
WFG-4	Best case	7.2798×10^{-2}	9.7223×10^{-2}	2.4216×10^{-1}	4.7123×10^{-2}	4.6921×10^{-2}	4.6121×10^{-2}
	Worst case	9.1627×10^{-2}	1.3619×10^{-1}	3.0073×10^{-1}	5.3997×10^{-2}	5.4812×10^{-2}	5.4789×10^{-2}
	ACT (s)	22.8714	18.6543	9.4587	19.1581	19.6831	19.9726
WFG-5	Best case	5.4619×10^{-2}	7.8523×10^{-2}	3.4681×10^{-2}	3.8851×10^{-2}	4.3266×10^{-2}	4.1251×10^{-2}
	Worst case	6.2721×10^{-2}	1.0871×10^{-1}	5.1719×10^{-2}	4.6123×10^{-2}	4.7814×10^{-2}	4.3611×10^{-2}
	ACT (s)	29.9643	16.1327	10.7551	15.7786	15.9963	16.0571
WFG-6	Best case	7.9122×10^{-2}	2.1203×10^{-1}	6.1266×10^{-2}	1.3172×10^{-1}	6.7127×10^{-2}	5.7229×10^{-2}
	Worst case	1.0124×10^{-1}	2.3771×10^{-1}	6.7849×10^{-2}	1.3713×10^{-1}	9.0401×10^{-2}	1.1701×10^{-1}
	ACT (s)	37.2817	25.9171	15.3658	26.8715	27.0179	27.1581
WFG-7	Best case	3.4216×10^{-2}	6.3547×10^{-2}	4.1681×10^{-2}	3.2012×10^{-2}	5.7229×10^{-2}	3.1149×10^{-2}
	Worst case	3.4771×10^{-2}	8.6954×10^{-1}	7.7319×10^{-2}	3.2482×10^{-2}	1.1701×10^{-1}	3.2131×10^{-2}
	ACT (s)	30.7684	27.8413	16.6511	25.0315	25.3617	25.4179

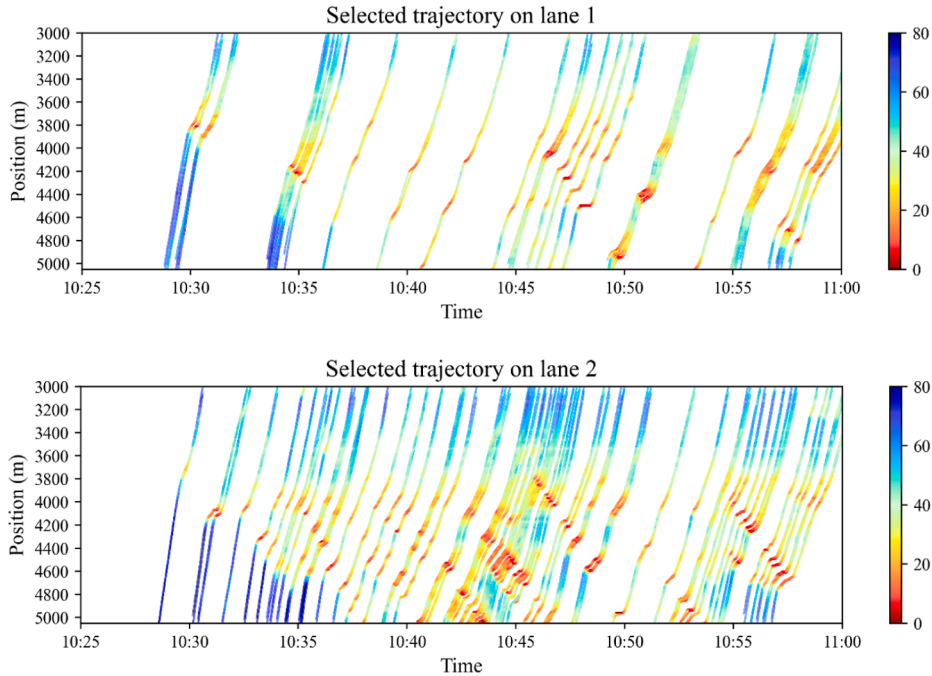
Table 2

HV values (Best and worst cases) obtained for several optimization algorithms on WFG problems over 30 independent runs.

Benchmark Functions		MOPSO	NSGA-II	MOEA/D	MOWSO	IMOWSO*	IMOWSO
WFG-1	Best case	6.1813×10^{-1}	6.0852×10^{-1}	5.7631×10^{-1}	6.2161×10^{-1}	6.2693×10^{-1}	6.2699×10^{-1}
	Worst case	6.0898×10^{-1}	5.9427×10^{-1}	5.2348×10^{-1}	6.1272×10^{-1}	6.2419×10^{-1}	6.2474×10^{-1}
WFG-2	Best case	5.5402×10^{-1}	5.3701×10^{-1}	5.5561×10^{-1}	5.6245×10^{-1}	5.6317×10^{-1}	5.6439×10^{-1}
	Worst case	5.4831×10^{-1}	5.2861×10^{-1}	5.4546×10^{-1}	5.5681×10^{-1}	5.6121×10^{-1}	5.6226×10^{-1}
WFG-3	Best case	3.0549×10^{-1}	2.2326×10^{-1}	2.9978×10^{-1}	3.0628×10^{-1}	3.0691×10^{-1}	3.0782×10^{-1}
	Worst case	3.0513×10^{-1}	2.0547×10^{-1}	2.8733×10^{-1}	3.0539×10^{-1}	3.0511×10^{-1}	3.0555×10^{-1}
WFG-4	Best case	3.1392×10^{-1}	2.9452×10^{-1}	2.3171×10^{-1}	3.2372×10^{-1}	3.2712×10^{-2}	3.2988×10^{-2}
	Worst case	3.0139×10^{-1}	2.9255×10^{-1}	2.0217×10^{-1}	3.2162×10^{-1}	3.2719×10^{-2}	3.2714×10^{-2}
WFG-5	Best case	3.2347×10^{-1}	3.1082×10^{-1}	3.3511×10^{-1}	3.1677×10^{-1}	3.2463×10^{-1}	3.2492×10^{-1}
	Worst case	3.2169×10^{-1}	3.0118×10^{-1}	3.3178×10^{-1}	2.9612×10^{-1}	3.1214×10^{-1}	2.9114×10^{-1}
WFG-6	Best case	3.1827×10^{-1}	2.8659×10^{-1}	3.1917×10^{-1}	3.1639×10^{-1}	3.2465×10^{-1}	3.2465×10^{-1}
	Worst case	3.1362×10^{-1}	2.7891×10^{-1}	2.7557×10^{-1}	2.9673×10^{-1}	3.2723×10^{-1}	2.9123×10^{-1}
WFG-7	Best case	3.3053×10^{-1}	3.1793×10^{-1}	3.2761×10^{-1}	3.2959×10^{-1}	3.3103×10^{-1}	3.3103×10^{-1}
	Worst case	3.2983×10^{-1}	3.1026×10^{-1}	3.0746×10^{-1}	3.2902×10^{-1}	3.3086×10^{-2}	3.3086×10^{-2}



(a) Aerial orthophoto of the investigated segment. Source from (Ztd, 2018)



(b) Selected microscopic trajectory pairs for calibration

Fig. 7. Demonstration of ZTD trajectory dataset.

approximately 100 % of the flow. In contrast to other commonly used microtrajectory datasets, the ZTD data incorporates various traffic flow states, including free-flow states, congestion states, and stop-and-go waves, which significantly facilitate the validation of the feasibility of the proposed integrated macro–micro calibration framework. Based on the filtering principle of the leading-following pairs described in the integrated framework, we set the predefined minimum following length T_{\min} and maximum headway s_{\max} as 20 s and 40 m (Hammit et al., 2018, He et al., 2018) to ensure that the following process is long enough to capture vehicle behavior characteristics. As shown in Fig. 7(b), we randomly selected 100 leading-following pairs that conformed to the above criteria in lanes 1 and 2 and experienced at least one shockwave. Furthermore, we assumed that FD was equivalent for the entire study segment. We utilize the above-mentioned ZTD trajectory for the calibration of macro–micro traffic flow models based on the self-consistent framework. Generally, the scale of empirical experiment depends on the spatial and temporal extent of the selected trajectory data and comprises two aspects. For macroscopic level, we utilized the two-lane ZTD with a spatial–temporal range of about $2\text{km} \times 0.5\text{h}$, which is uniformly divided into 4879 spatial–temporal regions. To analyze the macro-level data, we conducted calibration and validation processes. Specifically, 50 % of the observations were randomly selected for calibration, while the remaining data were reserved for validating the performance of the estimation. For microscopic level, we adopted the selected trajectory data of 50 leading-following pairs for parameter calibration to eliminate overfitting and used the remaining 50 pairs for validation (Saifuzzaman et al., 2015, Sharma et al., 2019).

In the steady-state traffic detection approach based on deep learning, we use 10 % ($4879 \times 10\% \approx 488$ observations) of the collected space–time regions as training and validation data to ensure that the proposed detection model is sufficiently trained. Specifically, 80 % of this subset is designated for training, while the remaining 20 % is reserved for validation. Considering the limited number of space–time regions in training dataset, we employ the hybrid of quantitative and qualitative approaches to resolve the training dataset selection problem. Here, we adopt the partial criteria proposed by Wang et al. (2019) for the simplified identification of steady state in training dataset.

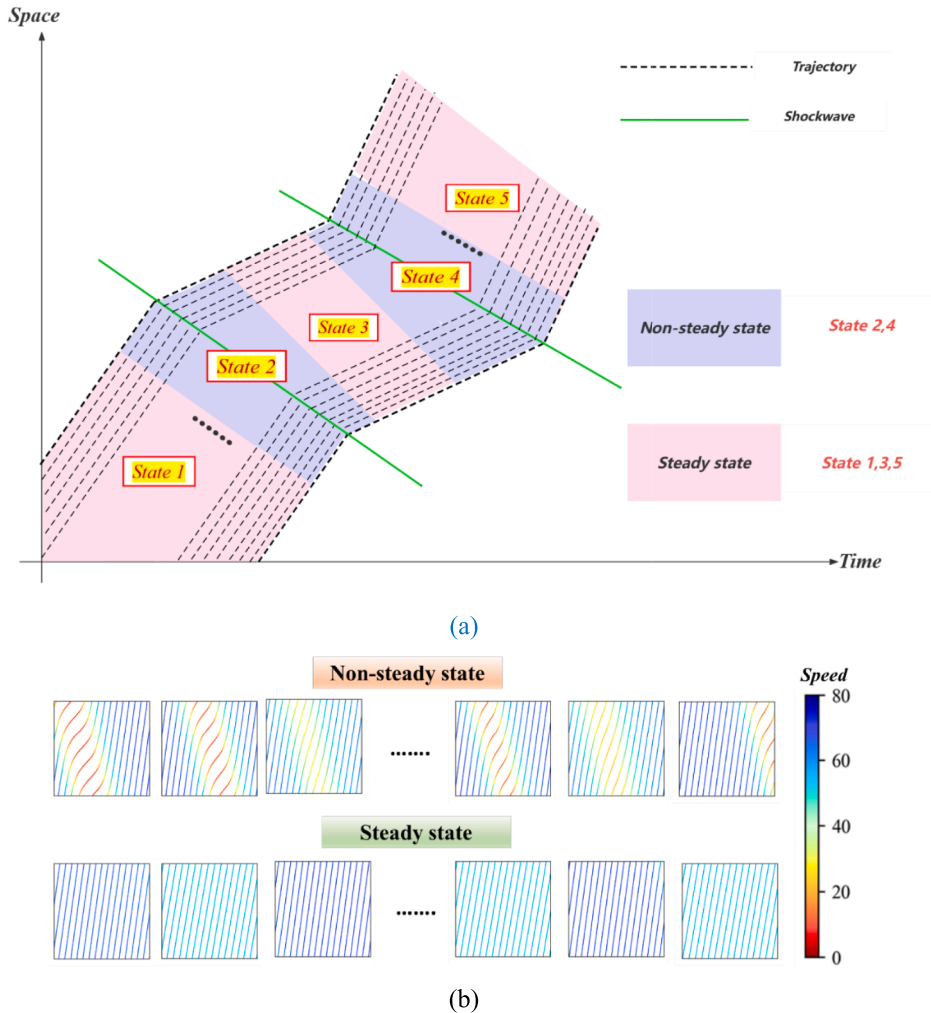


Fig. 8. (a) Illustration of steady state and non-steady state. (b) Examples of steady state and non-steady state data in training dataset.

- Standard deviation (STD) of the consecutive speed records (every 0.1 s) of the following vehicle should be less than 0.5 m/s;
- STD of the consecutive spacing records (every 0.1 s) should be less than 0.5 m.

Furthermore, we qualitatively select non-steady data from the remaining ones that do not fulfill the above-mentioned conditions. The specific selection demonstration is based on the Fig. 8a, where the state 2 & 4 are non-steady (Typically accompanied by shockwave disturbances). The non-steady and steady-state data adopted in the training set occupies half of the data respectively to acquire the information adequately. Fig. 8b illustrates the selection of partial samples after the aforementioned integration method. The input images of the space-time regions (As seen in Fig. 8b) were resized to 224×224 . The length of the feature vector for each extracted is set to 100, and the dimension of each space-time region ω is set to $50\text{m} \times 30\text{s}$. In addition, the learning rate was fixed uniformly at 1×10^{-3} .

In this study, continuous-time OVM and IDM were selected as examples. According to the literature (Li et al., 2016) and combined with the ZTD trajectory data characteristics, the recommended bounds of the IDM parameters are defined as $\theta_{\text{IDM},\min} \leq \theta_{\text{IDM}} \leq \theta_{\text{IDM},\max}$, where $\theta_{\text{IDM},\min} = [1\text{m/s}^2, 1\text{m/s}^2, 10\text{m/s}, 0.1\text{s}, 1\text{m}]^T$ and $\theta_{\text{IDM},\max} = [5\text{m/s}^2, 4\text{m/s}^2, 30\text{m/s}, 3\text{s}, 10\text{m}]^T$. The recommended bounds of the OVM parameters are set using $\theta_{\text{OVM},\min} = [0.5\text{s}, 10\text{m/s}, 5\text{m}, 1]^T$ and $\theta_{\text{OVM},\max} = [1.5\text{s}, 30\text{m/s}, 20\text{m}, 2.5]^T$. The average vehicle length L is set to 4 m.

Several previously mentioned typical multi-objective optimization algorithms were incorporated into the evaluation to measure the effectiveness of the proposed IMOWSO algorithm, where the maximum number of iterations K , population size n , and repository size n' were configured as 50, 200, and 100, respectively. Appendix B presents the hardware configuration and specific configuration parameters for each optimization algorithm.

B. Investigations on the calibration results.

To evaluate the performance of the macro-micro self-consistent calibration framework, we first filter the macroscopic traffic data by selecting traffic quantities in the steady-state flow state. To address the problem, we use the proposed steady-state detection approach. The confusion matrix for the binary classification task is used to evaluate the performance of the detection methodology and to compare it with 3 benchmark models, namely, *ResNet*, *EfficientNet* and *Transformer*. The confusion matrix consists of True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). TP means the proportion of non-steady state predicted as non-

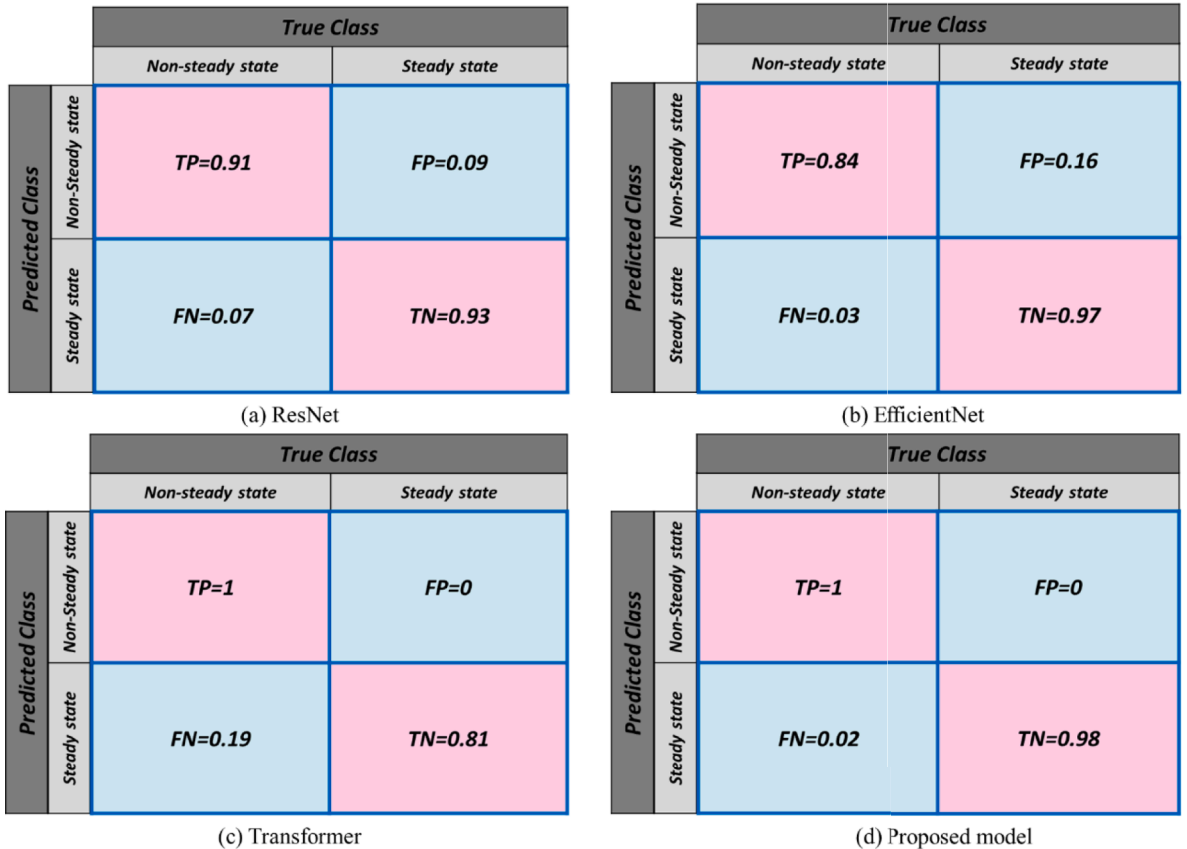


Fig. 9. Confusion matrices for traffic state classification.

steady and TN represents the ratio of steady-state estimated as steady, while the FP stands for the percentage of non-steady state predicted as steady and vice versa for FN. The detection results are presented in Fig. 9(a)-(d) by comparing the existing benchmark models based on the confusion matrix. The proposed detection methodology achieves 100 % precision with outstanding performance in FN and TN classification in the validation set. Combined with the Transformer's metric of only 81 % on TN, the new detection model efficiently filters steady-state traffic. This model outperforms several traditional deep-learning models in both TP and TN metrics. The macroscopic dataset originally contained 4,879 observations. As shown in Fig. 10, after the above-described detection procedures, the number of selected steady-state observations is 2,246. To avoid overfitting, we randomly selected 50 % (namely, $2,246 \times 50\% = 1,123$) observations for calibration, and the remaining observations ($2,246 \times 50\% = 1,123$) were used for the validation.

In this experiment, the speed (v) and spacing headway (s) in the microscopic trajectory and traffic density (\hat{k}) in the macroscopic FD were selected as the MoPs for the calibration of the IDM. The RMSE is typically utilized in statistical analysis to measure the deviation between the estimated and ground truth values and was selected as the GoF. Tables 3 & 4 display the calibration and validation results of the self-consistent framework with various algorithms integrated into the IDM and OVM. As shown in Table 3, IMOWSO ranked #1 for predicted speed, and ranked #2 for the estimated spacing headway and traffic density. Similar results were obtained during the validation. According to the outcomes of the OVM calibration, IMOWSO ranked #1 in the estimated spacing headway and traffic speed while ranked #3 in the calibration results of the predicted speed. This ranking was also supported during the validation phase of OVM presented in Table 4.

The current separation between macro- and microscopic traffic model calibration processes can lead to inconsistency of identical parameters in the two dimensions. To demonstrate and verify this, we analyzed the microscopic trajectories of vehicles with ID 774 and 1665. We validated our findings by implementing the optimal parameters calibrated from the microscopic trajectory data into their corresponding macroscopic FD and analogously by introducing the optimal parameters of identical physical significance calibrated in the macroscopic FD into the consistent microscopic CF model. Fig. 11(a) and (c) depict the results of this analysis, where the red lines represent real-world headway, the solid blue line represents the predicted headway derived from the self-consistent calibration framework, and the green dashed line (Mode 2) is generated from the parameters calibrated by the macroscopic FD with the single objective function ($f_3 - RMSE(\hat{v})$). Note that owing to the lack of microscopic vehicle dynamics parameters and driver features calibration, we introduced typical reference values on highways for IDM (i.e., $a_{\max} = 1\text{m/s}^2$ and $b = 1.5\text{m/s}^2$) and OVM (i.e., $\tau = 0.65\text{s}$) (Treiber and Kesting, 2013b). The solid blue lines (Mode 1) in Fig. 11(b) and (d) represent the 2D relationship curves of speed and density derived from the parameters calibrated by the proposed framework, whereas the green dashed line is produced by the parameters from a single microscopic calibration ($f_2 - RMSE(s)$) based on the trajectory data. It is evident that the current independent calibration procedures for macro and micro models lead to a lack of self-consistency in the respective results. This also demonstrates the relatively robust physical interpretation of the proposed physics-informed framework. Considering that there are no practical parameter values in the real-world trajectory data for comparison, it is infeasible to determine if a global optimum has been achieved. There is a certain degree of possibility that our present analysis is based on a comparison of local optima. Therefore, to further investigate the robustness of the proposed calibration framework, we produced synthetic data by predetermining the parameter values in Section 5.2.

5.2. Synthetic trajectory experiment

A. Experiment design.

Given that real-world trajectory data lacks explicit parameter values for reference, it is difficult to determine whether global optimal data has been found. To address this, we used numerical simulations based on the operating principles of the OVM to further investigate the performance of the proposed calibration framework. We used OVM as a demonstration, but this experimental design

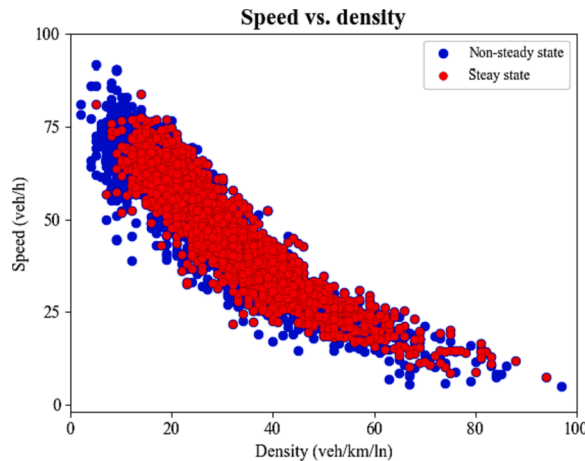


Fig. 10. Speed-density plane (non-steady vs. steady observations).

Table 3

Comparison of *RMSEs* for estimated speed (v), spacing headway (s) and traffic density (\hat{k}) obtained from several algorithms in multi-objective functions (IDM).

Algorithms		$f_1 - RMSE(v)$	$f_2 - RMSE(s)$	$f_3 - RMSE(\hat{v})$
MOPSO	Calibration	0.263	0.587	6.812
	Validation	0.627	1.217	6.915
NSGA-II	Calibration	0.274	0.851	7.631
	Validation	0.563	1.657	6.449
MOEA/D	Calibration	0.278	0.614	5.072
	Validation	0.589	0.963	3.832
MOWSO	Calibration	0.262	0.732	9.453
	Validation	0.604	1.453	6.611
IMOWSO	Calibration	0.251	0.513	6.076
	Validation	0.589	1.076	3.812

Table 4

Comparison of *RMSEs* for estimated speed (v), spacing headway (s) and traffic speed (\hat{v}) obtained from several algorithms in each objective function (OVM).

Algorithms		$f_1 - RMSE(v)$	$f_2 - RMSE(s)$	$f_3 - RMSE(\hat{v})$
MOPSO	Calibration	0.262	0.977	8.189
	Validation	0.587	1.836	8.271
NSGA-II	Calibration	0.284	0.993	5.391
	Validation	0.543	1.878	6.817
MOEA/D	Calibration	0.383	2.314	10.367
	Validation	0.646	2.178	12.452
MOWSO	Calibration	0.254	0.982	4.978
	Validation	0.597	1.769	6.016
IMOWSO	Calibration	0.242	0.957	3.917
	Validation	0.561	1.259	3.566

can be applied to other analytical CF models.

Fig. 12 shows the simulation environment, which consisted of a looped single lane with 80 vehicles connected at the beginning and end. The total simulation time was 1,200 s, and the time interval was Δt set to 0.01 s. In the initial stage, the vehicles have an equilibrium spacing headway (20m) and equilibrium speed (12m/s). At 20 s, a small perturbation was applied to the leading vehicle, which was decelerated by -1.5 m/s^2 for 2 s, and then repeated every 50 s. The simulation set-up included free flow, congestion, and stop-and-go waves (as shown in Fig. 13). To generate the synthetic trajectory, we used predefined values of the relevant parameters, which are presented in Table 5 (Treiber and Kesting, 2013b). The recommended bounds of the OVM with $\theta_{\text{OVM},\min} = [0.5s, 10\text{m/s}, 5\text{m}, 1]^T$ and $\theta_{\text{OVM},\max} = [1.5s, 30\text{m/s}, 20\text{m}, 2.5]^T$ are given. The parameters for several multi-objective optimization algorithms were configured in the same manner as in Section 5.1. In general, the scale of synthetic trajectory experiment contains two dimensions. For macroscopic level, we utilized the generated trajectory with a spatial-temporal range of about $1.6\text{km} \times 0.3\text{h}$, which is uniformly divided into 1280 space-time regions (ω is set to $50\text{m} \times 30\text{s}$). For microscopic level, we adopted the trajectory data of 40 leading-following pairs for parameter calibration to eliminate overfitting and used the remaining 40 pairs for validation.

B. Investigations of the test results.

We used similar self-consistent calibration procedures as in Section 5.1. First, steady-state traffic detection was performed to select 874 observations from the original dataset, as shown in Fig. 14(a)-(b). The new detection model demonstrated remarkably well in the steady-state classification task, with 100 % accuracy in terms of TF and TN based on the validation set. To further investigate the framework's robustness, 50 % (i.e., $874 \times 50\% = 437$) of the observations were randomly selected for calibration of the corresponding steady-state macro model, while 50 % were reserved for validation.

Table 6 lists various optimization algorithms embedded in this physics-informed calibration framework. Comparing their outcomes enables us to evaluate their performances and determine whether a global optimal solution can be obtained. The results obtained from the self-consistent framework closely approached the global optimal solution at a microscopic level. The analysis shows that IMOWSO ranked #1 in both the estimated spacing headway and speed and #2 in the predicted traffic speed. Table 7 shows that the optimal solution $\theta^* = [15.217, 0.653, 8.071, 1.536]^T$ derived from IMOWSO accurately characterizes the variation in microscopic driving behavior, and has a relatively strong physical interpretation of the fitting performance of the macroscopic FD. No existing algorithms can guarantee that a global optimal solution is discovered within a limited time. However, in general, we can conclude that the proposed framework can address nonlinear and nonconvex calibration problems with multiple parameters and can avoid falling into local optima to a certain extent because of its balance between exploration and exploitation.

Previous studies have examined the effect of the trajectory data sampling rate on the calibration results (Treiber and Kesting, 2013a). We further analyzed the impact of the sampling interval on the self-consistent calibration framework in combination with available synthetic data. Based on the original sampling interval, we reset the sampling interval to 0.05 s, 0.1 s, 0.5 s, and 1 s,

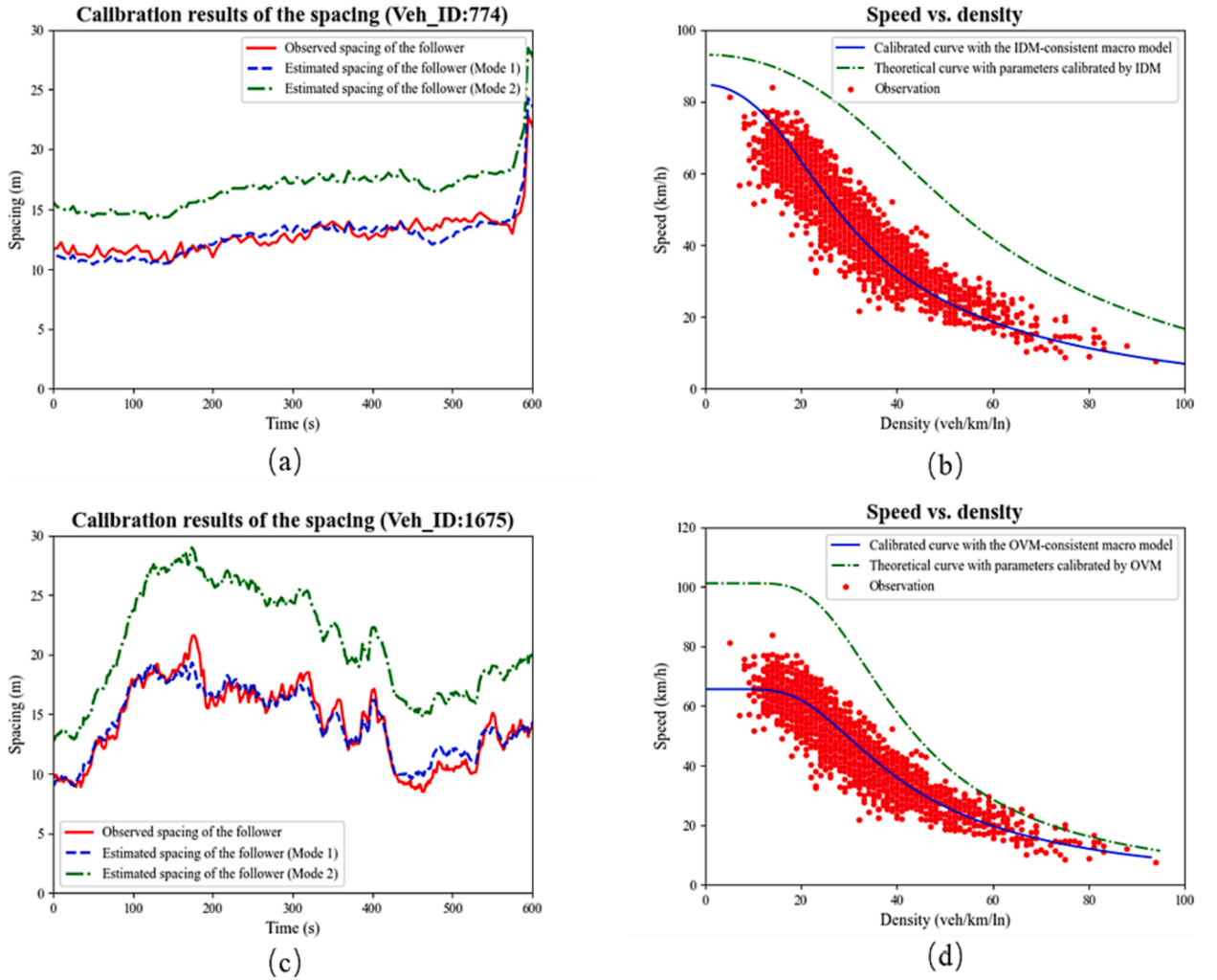


Fig. 11. Macro and micro results for IDM (a & b) and OVM (c & d). (It is necessary to note that Mode1 represents the calibration results corresponding to macro and micro (solid blue line in a & b) and Mode 2 demonstrates the separate calibration outcomes (the green dashed line in c & d).).

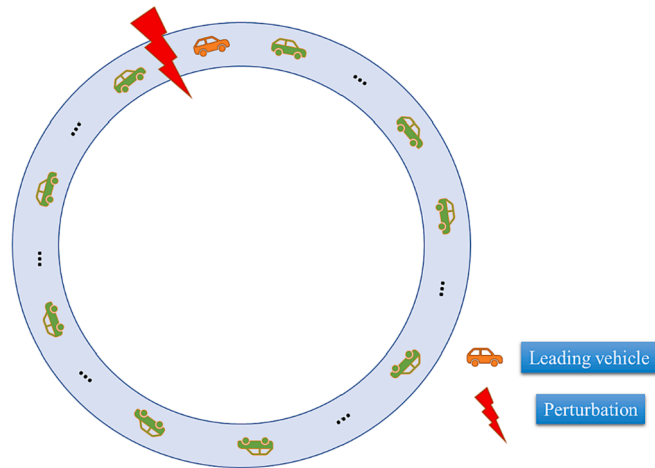


Fig. 12. Demonstration of simulation environment settings.

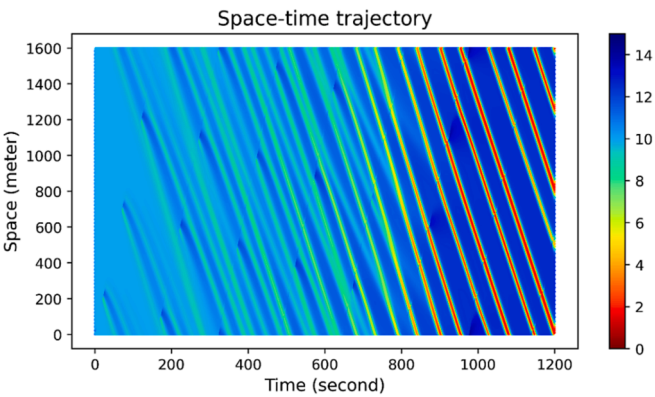
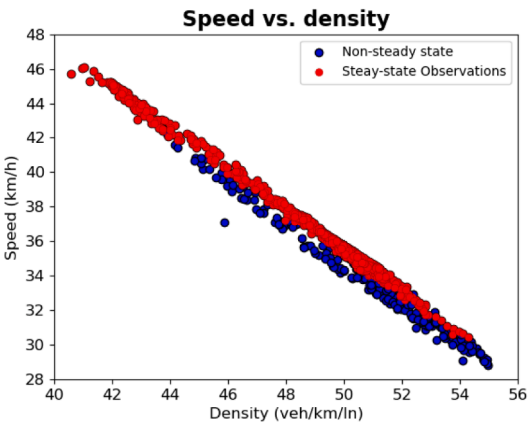


Fig. 13. Synthetic vehicle trajectories based on OVM.

Table 5
CF parameters associated with OVM.

Parameter (unit)	OVM
Adaptation time τ (s)	0.65
Desired speed v_0 (m/s)	15
Transition width Δs (m)	8
Form factor β	1.5



(a) Speed-density plane (non-steady vs. steady observations)

		True Class	
		Non-steady state	Steady state
Predicted Class	Non-Steady state	TP=1	FP=0
	Steady state	FN=0	TN=1

(b) Confusion matrix for traffic state classification (Proposed model)

Fig. 14. Steady-state detection results.

Table 6Calibration results and errors (%) of OVM with different intervals Δt .

Δt	0.01 s		0.05 s		0.1 s		0.5 s		1 s	
	Results	Errors	Results	Errors	Results	Errors	Results	Errors	Results	Errors
v_0 (m/s)	15.217	1.446 %	14.789	1.406 %	15.479	3.193 %	14.513	4.871 %	13.271	11.533 %
τ (s)	0.653	0.462 %	0.619	4.769 %	0.621	4.461 %	0.607	6.615 %	0.519	20.153 %
Δs (m)	8.071	0.0875 %	8.237	2.963 %	8.587	7.338 %	7.412	7.351 %	10.817	35.213 %
β	1.536	2.413 %	1.631	8.873 %	1.671	11.412 %	1.689	12.667 %	2.017	34.667 %

Table 7Comparison of *RMSEs* for estimated speed (v), spacing headway (s) and traffic speed (\hat{v}) obtained from several algorithms in each objective function (sampling interval $\Delta t = 0.01$ s).

Algorithms		$f_1 - RMSE(v)$	$f_2 - RMSE(s)$	$f_3 - RMSE(\hat{v})$
MOPSO	Calibration	1.187×10^{-3}	0.129	1.489
	Validation	2.067×10^{-3}	0.203	5.285
NSGA-II	Calibration	8.351×10^{-4}	0.071	1.318
	Validation	1.817×10^{-3}	0.189	1.071
MOEA/D	Calibration	1.947×10^{-3}	0.281	4.457
	Validation	4.163×10^{-3}	0.411	9.451
MOWSO	Calibration	1.977×10^{-3}	0.194	3.817
	Validation	4.318×10^{-3}	0.317	7.123
IMOWSO	Calibration	7.972×10^{-4}	0.097	1.782
	Validation	1.783×10^{-3}	0.169	1.801

respectively, meaning that the number of trajectory data becomes 20 %, 10 %, 2 %, and 1 % of the original dataset, while the total simulation time remains intact.

Based on the predetermined reference values of the parameters in Table 5, we proceed to discuss the calibration results presented in Table 6. In addition, *MAPE* was measured as an indicator of the error in the calibration results, which is a relative measure that adopts absolute values to prevent positive and negative errors from offsetting each other. The relative error between the calibration results of each parameter and the specified reference value is sustained within 15 % between 0.01 s and 0.5 s. In contrast, the *MAPE* of all three parameters exceeds 20 % when the sampling interval is $\Delta t = 1$ s, and even that of form factor β is approaching 50 %. Similarly, we compared the built-in optimization algorithms with different sampling evaluation rates Δt (Table 7–11). It is obvious that IMOWSO has superior performance at the sample rate of 0.01 s–0.5 s and was ranked in the top 3 for each objective function. However, at $\Delta t = 1$ s, each objective function value is orders of magnitude different from previous results. Based on the above test results, we conclude that the time interval $\Delta t = 0.5$ s is adequate for calibrating macro/micro traffic models in this experimental analysis, with a reduction of 98 % from the previous simulation data. However, when the sampling width Δt was extended to 1 s, there was a large discrepancy between the calibrated parameters and the ultimate estimation results.

6. Conclusion

In this study, we present a self-consistent methodology for calibrating macro- and microscopic traffic flow models, inspired by the fact that existing calibration approaches produce significantly diverse outcomes for identical parameters at both macro and micro levels. Previous research has focused on the fitting performance of microscopic trajectories, overlooking the physical significance of the steady-state macroscopic form of the CF model. Consequently, parameters calibrated using only the objective function of

Table 8Comparison of *RMSEs* for estimated speed (v), spacing headway (s) and traffic speed (\hat{v}) obtained from several algorithms in each objective function (sampling interval $\Delta t = 0.05$ s).

Algorithms		$f_1 - RMSE(v)$	$f_2 - RMSE(s)$	$f_3 - RMSE(\hat{v})$
MOPSO	Calibration	6.817×10^{-3}	0.126	2.651
	Validation	1.449×10^{-2}	0.257	3.089
NSGA-II	Calibration	8.128×10^{-3}	0.168	1.763
	Validation	1.317×10^{-2}	0.232	1.864
MOEA/D	Calibration	1.255×10^{-2}	0.241	6.113
	Validation	1.697×10^{-2}	0.363	6.678
MOWSO	Calibration	1.271×10^{-2}	0.203	4.413
	Validation	1.851×10^{-2}	0.367	3.983
IMOWSO	Calibration	5.428×10^{-3}	0.109	2.403
	Validation	1.191×10^{-2}	0.261	2.781

Table 9

Comparison of *RMSEs* for estimated speed (v), spacing headway (s) and traffic speed (\hat{v}) obtained from several algorithms in each objective function (sampling interval $\Delta t = 0.1s$).

Algorithms		$f_1 - RMSE(v)$	$f_2 - RMSE(s)$	$f_3 - RMSE(\hat{v})$
MOPSO	Calibration	7.109×10^{-3}	7.136×10^{-2}	2.612
	Validation	1.478×10^{-2}	0.237	2.831
NSGA-II	Calibration	7.861×10^{-3}	7.941×10^{-2}	5.258
	Validation	1.671×10^{-2}	0.169	6.069
MOEA/D	Calibration	2.234×10^{-2}	0.208	10.529
	Validation	4.593×10^{-2}	0.509	12.187
MOWSO	Calibration	1.509×10^{-2}	0.151	3.168
	Validation	3.572×10^{-2}	0.397	3.781
IMOWSO	Calibration	5.972×10^{-3}	5.991×10^{-2}	2.012
	Validation	1.478×10^{-3}	0.129	2.347

Table 10

Comparison of *RMSEs* for estimated speed (v), spacing headway (s) and traffic speed (\hat{v}) obtained from several algorithms in each objective function (sampling interval $\Delta t = 0.5s$).

Algorithms		$f_1 - RMSE(v)$	$f_2 - RMSE(s)$	$f_3 - RMSE(\hat{v})$
MOPSO	Calibration	0.173	0.361	6.667
	Validation	0.269	0.576	7.038
NSGA-II	Calibration	0.161	0.318	4.513
	Validation	0.277	0.497	4.863
MOEA/D	Calibration	7.289×10^{-2}	0.131	2.518
	Validation	9.785×10^{-2}	0.227	2.587
MOWSO	Calibration	0.129	0.309	3.139
	Validation	0.231	0.469	3.893
IMOWSO	Calibration	4.512×10^{-2}	0.091	3.075
	Validation	7.389×10^{-2}	0.218	3.717

Table 11

Comparison of *RMSEs* for estimated speed (v), spacing headway (s) and traffic speed (\hat{v}) obtained from several algorithms in each objective function (sampling interval $\Delta t = 1s$).

Algorithms		$f_1 - RMSE(v)$	$f_2 - RMSE(s)$	$f_3 - RMSE(\hat{v})$
MOPSO	Calibration	0.219	0.207	3.123
	Validation	0.261	0.237	3.845
NSGA-II	Calibration	0.247	0.216	4.493
	Validation	0.252	0.237	4.417
MOEA/D	Calibration	0.287	0.491	3.837
	Validation	0.336	0.681	3.814
MOWSO	Calibration	0.681	0.761	9.207
	Validation	0.769	0.814	9.511
IMOWSO	Calibration	0.513	0.612	8.683
	Validation	0.663	0.734	9.541

microscopic vehicle movements showed inferior behavior in the validation of the macroscopic FD models. To effectively resolve the calibration problem, we convert it into a multi-objective optimization problem by introducing macroscopic objective functions and using the IMOWSO to solve it. Compared with previous traditional algorithms, IMOWSO performs better on 72 % of the benchmark functions that were ranked #1 in terms of the IGD and HV. To improve the accuracy and robustness of steady-state traffic model calibration, a supervised steady-state detection approach based on an attention mechanism is presented to efficiently remove nonstationary scatters from macroscopic observations (i.e., traffic density, traffic flow, and space-mean speed). The real-world and synthetic trajectory data experimental results validate the proposed self-consistent calibration framework, which maximizes the descriptive power of macro and micro traffic flow models. Based on the same trajectory dataset, the proposed framework bridges the macro and micro traffic flow models and accommodates both macro and micro level accuracy with the augmentation of domain knowledge (characteristics of macroscopic and microscopic traffic stream) during the calibration process, breaking the inconsistency of the existing methodological system and substantially enhancing physical interpretability. Additionally, the scalability and transferability of this physics-informed framework is another purpose of this paper, which is universal for diverse traffic scenarios and not restricted to a specific CF or FD model.

The physics-informed calibration process of traffic flow models typically assumes that parameters are constant, which fails to account for the diversity among drivers with differing driving preferences and reaction times, as well as the heterogeneity of traffic

flow, such as different vehicle types (autonomous or human-driven) with varying power capabilities and road conditions. Incorporating these factors should be a focus of future research in traffic flow modeling and calibration. Instead of fixed values, representing the calibration results of traffic flow model parameters in a defined spatiotemporal range for a particular road section in the form of a distribution may be more appropriate and intuitive.

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 Key Project (No. 52131203) and the Youth Program (No. 52102375) of the National Natural Science Foundation of China, the Youth Program (No. BK20210247) of the Natural Science Foundation of Jiangsu Province, China, and the High-Level Personnel Project of Jiangsu Province, China (JSSCBS20220099).

Appendix A1

The update strategy (Phase I) involves the optimization of each individual's speed and movement, which are as follows:

$$v_{k+1}^i = \mu \left[v_k^i + \delta_1 (w_{\text{gbest}_k} - w_k^i) \times c_1 + \delta_2 (w_{\text{best}}^{Idx_k^i} - w_k^i) \times c_2 \right] \quad (\text{A1})$$

$$Idx = \lfloor n \times \text{rand}(1, n) \rfloor + 1 \quad (\text{A2})$$

where k defines the current iteration number and v_{k+1}^i represents the optimization speed for the i^{th} candidate solution in the $(k+1)^{\text{th}}$ iteration. Specifically, $v_1^i = 0$ was the first step. w_k^i is the current position of i^{th} candidate solutions. w_{gbest_k} is the best global candidate solution obtained from the population by the k^{th} iteration. $w_{\text{best}}^{Idx_k^i}$ represents the best position of an Idx_k^i individual up to the current iteration, where Idx is the index vector randomly generated in each iteration. The μ represents a constriction factor, which needs to be predefined. $\text{rand}(1, n)$ is a vector of random numbers generated by a uniform distribution in the range $[0, 1]$. c_1 and c_2 denote random numbers within the interval 0–1. δ_1 and δ_2 are control coefficients that vary with the number of iterations and can be formulated as:

$$\begin{cases} \delta_1 = \delta_{\max} + (\delta_{\max} - \delta_{\min}) \times e^{-(4k/K)^2} \\ \delta_2 = \delta_{\min} + (\delta_{\max} - \delta_{\min}) \times e^{-(4k/K)^2} \end{cases} \quad (\text{A3})$$

where δ_{\min} and δ_{\max} represent the lower and upper control factors, k and K defines the current and maximum numbers of iterations, respectively.

After obtaining the updated individual velocity, the next step was to perform a preliminary update to the position of each candidate solution. This update strategy simulates the process of a white shark-seeking prey and balances exploration and exploitation research behaviors to some extent. This can be formulated as follows:

$$w_{k+1}^i = \begin{cases} w_k^i \cdot \neg \oplus w_0 + UB \cdot a + LB \cdot b; & \text{rand} < mv \\ w_k^i + v_k^i / r; & \text{rand} \geq mv \end{cases} \quad (\text{A4})$$

$$\begin{cases} a = \text{sgn}(w_k^i - UB) > 0 \\ b = \text{sgn}(w_k^i - LB) < 0 \end{cases} \quad (\text{A5})$$

where \neg is a negation operator, represents a random number within the range of 0–1, and r is a predefined parameter that can be analogous to the frequency of the wave motion. Additionally, a and b are binary vectors, w_0 is a logical vector and mv represents the movement impact factor.

$$w_0 = \oplus(a, b) \quad (\text{A6})$$

$$mv = \frac{1}{(a_0 + e^{(K/2-k)/a_1})} \quad (\text{A7})$$

where \oplus is a bitwise operation (XOR), a_0 and a_1 are two predetermined constants. A smaller mv value results in a local search and vice versa for a global search.

Appendix A2

The detailed procedures of Update Strategy for phase II:

$$w_{k+1}^i = w_{gbest_k} + r_1 \vec{D}_w \text{sgn}(r_2 - 0.5) r_3 < s_s \quad (\text{A8})$$

$$\vec{D}_w = |\text{rand} \times (w_{gbest_k} - w_k^i)| \quad (\text{A9})$$

where w_{k+1}^i represents the target update position concerning the global optimal solution, $\text{sgn}(r_2 - 0.5)$ determines the search direction (0 or 1), and r_1 , r_2 , and r_3 are random numbers within the range of 0 to 1. \vec{D}_w denotes the distance between the current and the global best candidate solutions. The parameter s_s is defined as follows:

$$s_s = |1 - e^{(-a_2 \times k/K)}| \quad (\text{A10})$$

where a_2 is a positive constant.

All particles should be updated accordingly, except for the first optimal preserved candidate solution to maximize the imitation of fish behavior. Combined with the equilibrium pool, we define this expression as follows:

$$w_{k+1}^i = \frac{w_k^i + w_{k+1}^i + \tilde{w}}{3 \times \text{rand}} \quad (\text{A11})$$

where \tilde{w} denotes the elite candidate randomly selected from the equilibrium pool $\overrightarrow{C_{EP}^k}$.

Appendix A3

The main steps of TOPSIS are as follows:

$$\mu_{ij} = \frac{f_{j-\max} - f_{ij}}{f_{ij} - f_{j-\min}} \quad (\text{A12})$$

where f_{ij} and μ_{ij} represent the j^{th} objective function value and j^{th} normalized value of i^{th} candidate, $f_{j-\max}$ and $f_{j-\min}$ denotes the maximum and minimum values of the j^{th} objective function.

After obtaining the normalized values μ_{ij} , we sort the solution candidates according to the information entropy-based technique, the details are as follows:

$$W_j = -\frac{1}{\ln(n)} \sum_{i=1}^n \left(\frac{\mu_{ij}}{\sum_{i=1}^n \mu_{ij}} \cdot \ln \frac{\mu_{ij}}{\sum_{i=1}^n \mu_{ij}} \right) \quad (\text{A13})$$

$$\tau_j = \frac{1 - W_j}{n - \sum_{j=1}^m W_j} \quad (\text{A14})$$

$$\begin{cases} S_{i,+} = \sqrt{\sum_{j=1}^m (\tau_j \mu_{ij} - \tau_j \mu_{i+})^2} \\ S_{i,-} = \sqrt{\sum_{j=1}^m (\tau_j \mu_{ij} - \tau_j \mu_{i-1})^2} \end{cases} \quad (\text{A15})$$

$$C_i = \frac{S_{i,-}}{S_{i,-} + S_{i,+}} \quad (\text{A16})$$

where W_j is the information entropy of the j^{th} objective function. τ_j is the weight of the j^{th} objective function, $S_{i,+}$ and $S_{i,-}$ are the positive and negative ideal distances, respectively. μ_{i+} and μ_{i-} are the maximum and minimum normalized values of the i^{th} candidate for each objective function, respectively. C_i is a fit coefficient where a larger value indicates a higher level of satisfaction with a candidate solution.

Appendix A4

Algorithm 1 An Improved Multi-objective WSO Algorithm

```

1: Input: Set configuration values for predetermined parameters: maximum iteration number  $K$ ,
   population size  $n$ , initial and subordinate speed  $\delta_{\min}$  and  $\delta_{\max}$ , frequency of wave motion  $r$ , low and
   upper bound of decision variables  $UB$  and  $LB$ .
2: Initialization: Set the current iteration number  $k = 1$ . Generate and evaluate the initial population
   of IMOWSO  $S_k$  based on Good Point Set strategy. Initialize the speed of the initial population.
3: while the current iteration number  $k \leq K$  do
4:   Update the relevant parameters according to Eqs. (A1)-(A11)
5:   for  $i \in S_k$  do
6:     Update the velocity and position of the solution  $i$  based on Update Strategy-Phase I
7:   end for
8:   Evaluate and select the elite individuals on each objective function to establish the equilibrium pool
    $\vec{C}_{EP}^k$ 
9:   for  $i \in S_k$  do
10:    Update the position of the solution  $i$  according to Update Strategy-Phase II
11:   end for
12:   Implement Renewal Mechanism of the Worst Solution to obtain and optimize the worst solution
    $w_{\text{worse}}$ 
13:   Adjust the candidate solutions that beyond the  $UB$  or  $LB$ 
14:   Update the new positions of candidate solutions in  $S_k$  based on non-dominated sorting
15:   The iteration number  $k = k + 1$ 
16: end while
17: return  $S_K$ 

```

Appendix B

Hardware configuration:

The IMOWSO algorithm was coded in MATLAB (R2017b) language and executed on a personal computer (Intel Core i7-10750H CPU@2.6 GHz and 16 GB-Ram).

Specific parameter settings:

- MOPSO. The inertia weight was 0.4, and both the personal and global learning coefficients were 2. The number of grids per dimension is 20. Leader and deletion selection pressures were 2. The mutation rate is 0.2.
- NSGA-II. The crossover fraction was 0.9, the migration fraction was 0.2, and the distribution indices for both the simulated binary crossover and polynomial migration were set to 20.
- MOEA/D. The crossover fraction is 0.9, neighbor size is 15.
- MOWSO. The predefined frequency r is 0.9, the lower control factor δ_{\min} and upper control factor δ_{\max} are 0.5 and 1, constriction factor is 0.7. Recommend values for constants a_0 , a_1 and a_2 are 6.25, 100 and 5×10^{-3} , respectively.
- IMOWSO. The same parameter settings as MOWSO.

Appendix C

Table C1

IGD values (Best and worst cases) and average computation time (ACT) obtained for several optimization algorithms on UF problems over 30 independent runs.

Benchmark Functions		MOPSO	NSGA-II	MOEA/D	MOWSO	IMOWSO*	IMOWSO
UF-1	Best case	4.9693×10^{-2}	5.5732×10^{-2}	9.8564×10^{-3}	4.8931×10^{-2}	3.9223×10^{-2}	3.9106×10^{-2}
	Worst case	7.2037×10^{-2}	5.5975×10^{-2}	4.6491×10^{-2}	6.8778×10^{-2}	4.4573×10^{-2}	4.3271×10^{-2}
	ACT (s)	32.3183	25.3817	17.9148	24.2171	24.9684	25.1439
UF-2	Best case	2.8273×10^{-2}	2.0916×10^{-2}	3.6271×10^{-2}	2.6389×10^{-2}	2.0397×10^{-2}	2.0309×10^{-2}
	Worst case	3.0427×10^{-2}	2.3189×10^{-2}	4.6781×10^{-2}	3.0295×10^{-2}	2.6452×10^{-2}	2.6311×10^{-2}
	ACT (s)	37.6285	27.3966	19.4894	26.2439	26.7133	26.7321
UF-3	Best case	3.8271×10^{-1}	2.7389×10^{-1}	7.3816×10^{-1}	3.5371×10^{-1}	3.3589×10^{-1}	3.3412×10^{-1}
	Worst case	3.3916×10^{-1}	3.5112×10^{-1}	8.7813×10^{-1}	3.6269×10^{-1}	3.4813×10^{-1}	3.4881×10^{-1}
	ACT (s)	39.3147	29.9118	19.9479	28.4657	28.7856	29.0117
UF-4	Best case	4.4871×10^{-2}	7.2012×10^{-2}	5.8137×10^{-2}	4.6873×10^{-2}	4.2107×10^{-2}	4.2501×10^{-2}
	Worst case	5.3269×10^{-2}	7.2183×10^{-2}	7.1193×10^{-2}	4.8293×10^{-2}	4.7932×10^{-2}	4.7819×10^{-2}
	ACT (s)	31.8413	25.4878	16.4582	25.1097	25.4827	25.5098
UF-5	Best case	5.7884×10^{-1}	4.1559×10^{-1}	4.9427×10^{-1}	6.0619×10^{-1}	2.8796×10^{-1}	2.8723×10^{-1}
	Worst case	6.6361×10^{-1}	6.6793×10^{-1}	5.7814×10^{-1}	6.4293×10^{-1}	4.0139×10^{-1}	3.3902×10^{-1}

(continued on next page)

Table C1 (continued)

Benchmark Functions		MOPSO	NSGA-II	MOEA/D	MOWSO	IMOWSO*	IMOWSO
UF-6	ACT (s)	28.7941	19.7562	12.4697	18.5241	18.8947	18.9746
	Best case	4.5689×10^{-2}	1.6533×10^{-1}	2.0719×10^{-1}	6.9127×10^{-2}	3.3712×10^{-2}	3.2926×10^{-2}
	Worst case	5.8777×10^{-2}	2.1874×10^{-1}	4.3629×10^{-1}	8.1897×10^{-2}	4.6712×10^{-2}	4.4354×10^{-2}
	ACT (s)	23.7621	15.2714	8.9416	14.7628	14.9561	15.0471

Table C2

HV values (Best and Worst cases) obtained for several optimization algorithms on UF problems over 30 independent runs.

Benchmark Functions		MOPSO	NSGA-II	MOEA/D	MOWSO	IMOWSO*	IMOWSO
UF-1	Best case	6.5481×10^{-1}	6.3919×10^{-1}	5.7171×10^{-1}	6.5512×10^{-1}	6.6913×10^{-1}	6.7081×10^{-1}
	Worst case	6.1678×10^{-1}	6.2617×10^{-1}	5.6948×10^{-1}	6.2489×10^{-1}	6.5639×10^{-1}	6.5779×10^{-1}
UF-2	Best case	6.8731×10^{-1}	6.9632×10^{-1}	6.7034×10^{-1}	6.8732×10^{-1}	6.8773×10^{-1}	6.9725×10^{-1}
	Worst case	6.8577×10^{-1}	6.9123×10^{-1}	6.6517×10^{-1}	6.8219×10^{-1}	6.9741×10^{-1}	6.9002×10^{-1}
UF-3	Best case	1.7386×10^{-1}	3.3116×10^{-1}	6.4523×10^{-2}	2.2109×10^{-1}	2.9132×10^{-1}	2.9687×10^{-1}
	Worst case	1.5742×10^{-1}	2.5493×10^{-1}	1.6201×10^{-2}	2.1758×10^{-1}	2.7378×10^{-1}	2.7412×10^{-1}
UF-4	Best case	3.8544×10^{-1}	3.3617×10^{-1}	3.5652×10^{-1}	3.8477×10^{-1}	3.8441×10^{-1}	3.8529×10^{-1}
	Worst case	3.7819×10^{-1}	3.3409×10^{-1}	3.4137×10^{-1}	3.8249×10^{-1}	3.8027×10^{-2}	3.7912×10^{-2}
UF-5	Best case	8.3932×10^{-2}	1.7119×10^{-1}	1.3171×10^{-1}	1.56819×10^{-2}	1.2973×10^{-1}	1.2973×10^{-1}
	Worst case	7.9742×10^{-2}	1.0382×10^{-2}	8.1745×10^{-2}	1.2593×10^{-2}	1.2617×10^{-1}	1.2941×10^{-1}
UF-6	Best case	5.2177×10^{-1}	4.1671×10^{-1}	3.6211×10^{-1}	4.8193×10^{-1}	5.4019×10^{-1}	5.4084×10^{-1}
	Worst case	5.0661×10^{-1}	3.5448×10^{-1}	2.4593×10^{-1}	4.7091×10^{-1}	5.1778×10^{-1}	5.1634×10^{-1}

Table C3

IGD values (Best and worst cases) and average computation time (ACT) obtained for several optimization algorithms on CF problems over 30 independent runs.

Benchmark Functions		MOPSO	NSGA-II	MOEA/D	MOWSO	IMOWSO*	IMOWSO
CF-1	Best case	1.3379×10^{-2}	1.6249×10^{-2}	9.6582×10^{-3}	4.7119×10^{-3}	4.4193×10^{-3}	4.4173×10^{-3}
	Worst case	1.5213×10^{-2}	1.7593×10^{-2}	1.4473×10^{-2}	5.2465×10^{-3}	5.0173×10^{-3}	5.0091×10^{-3}
	ACT (s)	17.3179	15.3428	9.7562	14.1567	14.5697	14.6014
CF-2	Best case	5.0618×10^{-2}	4.7932×10^{-2}	8.9118×10^{-2}	5.6629×10^{-2}	4.1877×10^{-2}	4.1801×10^{-2}
	Worst case	6.2096×10^{-2}	8.2378×10^{-2}	1.0743×10^{-1}	5.7349×10^{-2}	4.4473×10^{-2}	4.4378×10^{-2}
	ACT (s)	22.4867	15.7132	9.2374	14.4173	14.5973	14.6108
CF-3	Best case	1.0778×10^{-1}	8.8893×10^{-2}	1.2227×10^{-1}	1.1077×10^{-1}	8.8771×10^{-2}	8.8212×10^{-2}
	Worst case	1.3136×10^{-1}	9.0837×10^{-2}	1.3693×10^{-1}	1.1816×10^{-1}	9.0793×10^{-2}	9.0742×10^{-2}
	ACT (s)	20.6873	28.7963	13.4587	24.6987	24.8893	24.9078
CF-4	Best case	4.3891×10^{-1}	1.4512×10^{-1}	3.3127×10^{-1}	6.2609×10^{-1}	1.8107×10^{-1}	1.8393×10^{-1}
	Worst case	5.7436×10^{-1}	2.5509×10^{-1}	4.6579×10^{-1}	6.8792×10^{-1}	3.1789×10^{-1}	4.1422×10^{-1}
	ACT (s)	30.6756	27.9873	15.6742	25.4763	25.6784	25.7102
CF-5	Best case	3.3593×10^{-2}	2.6784×10^{-2}	3.9496×10^{-2}	2.7227×10^{-2}	2.3188×10^{-2}	2.3109×10^{-2}
	Worst case	4.3114×10^{-2}	2.9501×10^{-2}	1.2678×10^{-1}	3.3266×10^{-2}	3.0293×10^{-2}	3.0224×10^{-2}
	ACT (s)	29.1462	18.7456	8.7413	17.6387	17.9673	18.0143
CF-6	Best case	7.6887×10^{-1}	1.9409×10^{-1}	3.6461×10^{-1}	2.7981×10^{-1}	2.5677×10^{-1}	2.5684×10^{-1}
	Worst case	8.3493×10^{-1}	3.7773×10^{-1}	5.4113×10^{-1}	5.2336×10^{-1}	2.7579×10^{-1}	2.7527×10^{-1}
	ACT (s)	37.2871	25.0387	16.7855	24.8793	25.1897	25.2017

Table C4

HV values (Best and worst cases) obtained for several optimization algorithms on CF problems over 30 independent runs.

Benchmark Functions		MOPSO	NSGA-II	MOEA/D	MOWSO	IMOWSO*	IMOWSO
CF-1	Best case	5.6357×10^{-1}	5.6193×10^{-1}	5.6759×10^{-1}	5.7881×10^{-1}	5.7857×10^{-1}	5.7846×10^{-1}
	Worst case	5.5219×10^{-1}	5.5819×10^{-1}	5.6631×10^{-1}	5.7812×10^{-1}	5.7809×10^{-1}	5.7841×10^{-1}
CF-2	Best case	6.3871×10^{-1}	6.5492×10^{-1}	5.8519×10^{-1}	6.4851×10^{-1}	6.5827×10^{-1}	6.5841×10^{-1}
	Worst case	6.3813×10^{-1}	6.2971×10^{-1}	5.6711×10^{-1}	6.4593×10^{-1}	6.5721×10^{-1}	6.5736×10^{-1}
CF-3	Best case	3.7891×10^{-1}	4.3155×10^{-1}	4.0723×10^{-1}	4.0816×10^{-1}	4.1813×10^{-1}	4.1918×10^{-1}
	Worst case	3.5102×10^{-1}	4.0713×10^{-1}	3.5609×10^{-1}	3.6712×10^{-1}	4.0896×10^{-1}	4.1147×10^{-1}
CF-4	Best case	9.8749×10^{-2}	2.0873×10^{-1}	2.2396×10^{-1}	1.8903×10^{-2}	2.7308×10^{-1}	2.7415×10^{-1}
	Worst case	5.4136×10^{-2}	1.8455×10^{-1}	1.5007×10^{-1}	1.2713×10^{-2}	8.5611×10^{-2}	8.5696×10^{-2}
CF-5	Best case	6.6273×10^{-1}	6.7116×10^{-1}	6.5374×10^{-1}	6.9187×10^{-1}	7.0239×10^{-1}	6.6697×10^{-1}
	Worst case	6.4149×10^{-1}	6.6748×10^{-2}	6.2639×10^{-1}	6.8642×10^{-1}	6.6239×10^{-1}	6.6463×10^{-1}
CF-6	Best case	8.9713×10^{-4}	3.2789×10^{-1}	4.2931×10^{-1}	3.3496×10^{-1}	3.2235×10^{-1}	3.2296×10^{-1}
	Worst case	2.8789×10^{-4}	2.7603×10^{-1}	8.9382×10^{-2}	1.3432×10^{-1}	3.1237×10^{-1}	3.1944×10^{-1}

References

- Alammar, J., 2018. The illustrated transformer. <https://jalammar.github.io/illustrated-transformer/> (accessed February 8, 2023).
- Alhatriqi, A., Gu, Z., Saberi, M., 2022. Calibration of the intelligent driver model (IDM) with adaptive parameters for mixed autonomy traffic using experimental trajectory data. *Transp. B* 10 (1), 421–440.
- Bando, M., Hasebe, K., Nakayama, A., Shibata, A., Sugiyama, Y., 1995. Dynamical model of traffic congestion and numerical simulation. *Phys. Rev. E* 51 (2), 1035–1042.
- Bourel, E., Lesort, J.B., 2003. Mixing Microscopic and Macroscopic Representations of Traffic Flow: Hybrid Model Based on Lighthill-Whitham-Richards Theory. *Transp. Res. Rec.* 1852, 193–200.
- Brackstone, M., McDonald, M., 1999. Car-following: a historical review. *Transp. Res. Part C Emerg. Technol.* 2 (4), 181–196.
- Braik, M., Hammouri, A., Atwan, J., Al-Betar, M.A., Awadallah, M.A., 2022. White shark optimizer: A novel bio-inspired meta-heuristic algorithm for global optimization problems. *Knowledge-Based Syst.* 243, 108457.
- Bramich, D.M., Menéndez, M., Ambühl, L., 2022. Fitting empirical fundamental diagrams of road traffic: A comprehensive review and comparison of models using an extensive data set. *IEEE Trans. Intell. Transp. Syst.* 23 (9), 14104–14127.
- Bramich, D., Menéndez, M., Ambühl, L., 2023. FitFun: A modelling framework for successfully capturing the functional form and noise of observed traffic flow-density-speed relationships. *Transp. Res. Part C Emerg. Technol.* 151, 104068.
- Brockfeld, E., Kühne, R.D., Wagner, P., 2004. Calibration and validation of microscopic traffic flow models. *Transp. Res. Rec.* 1876, 62–70.
- Cassidy, M.J., 1998. Bivariate relations in nearly stationary highway traffic. *Transp. Res. Part B Methodol.* 32 (1), 49–59.
- Castillo, J.M.D., Benítez, F.G., 1995. On the functional form of the speed-density relationship-II: Empirical investigation. *Transp. Res. Part B Methodol.* 29 (5), 391–406.
- Chandler, R.E., Herman, R., Montroll, E.W., 1958. Traffic dynamics: Studies in car following. *Oper. Res.* 6 (2), 165–184.
- Chen, Y., Bai, Y., Zhang, W., Mei, T., 2019. Destruction and construction learning for fine-grained image recognition. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, IEEE.
- Chen, X., Sun, J., Ma, Z., Sun, J., Zheng, Z., 2020. Investigating the long- and short-term driving characteristics and incorporating them into car-following models. *Transp. Res. Part C Emerg. Technol.* 117, 102698.
- Cheng, Q., Liu, Z., Lin, Y., Zhou, X., 2021. An s-shaped three-parameter (S3) traffic stream model with consistent car following relationship. *Transp. Res. Part B Methodol.* 153, 246–271.
- Cheng, Q., Lin, Y., Zhou, X., Liu, Z., 2024. Analytical formulation for explaining the variations in traffic states: A fundamental diagram modeling perspective with stochastic parameters. *Eur. J. Oper. Res.* 312 (1), 182–197.
- Cheng, Z., Wang, X., Chen, X., Trepanier, M., Sun, L., 2022. Bayesian calibration of traffic flow fundamental diagrams using Gaussian processes. *IEEE Open J. Intell. Transp. Syst.* 3, 763–771.
- Chong, L., Abbas, M.M., Medina Flintsch, A., Higgs, B., 2013. A rule-based neural network approach to model driver naturalistic behavior in traffic. *Transp. Res. Part C Emerg. Technol.* 32, 207–223.
- Ciuffo, B., Punzo, V., 2010. Verification of Traffic Micro-Simulation Model Calibration Procedures: Analysis of Goodness-of-Fit Measures. *Transportation Research Record*.
- Coello, C.A., Lechuga, M.S., 2002. MOPSO: A Proposal for Multiple Objective Particle Swarm Optimization CEC'02 (Cat. No.02TH8600).
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6 (2), 182–197.
- Edie, L.C., 1963. Discussion of traffic stream measurements and definitions. Port of New York Authority, New York, NY.
- Eusuff, M., Lansey, K., Pasha, F., 2006. Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization. *Eng. Optim.* 38 (2), 129–154.
- Faramarzi, A., Heidarinejad, M., Stephens, B., Mirjalili, S., 2020. Equilibrium optimizer: A novel optimization algorithm. *Knowledge-Based Syst.* 191, 105190.
- Federal Highway Administration, 2010. Next generation simulation (NGSIM). accessed February 8, 2023. <http://ops.fhwa.dot.gov/trafficanalysisstools/ngsim.htm>.
- Gazis, D.C., Herman, R., Rothery, R.W., 1961. Nonlinear follow-the-leader models of traffic flow. *Oper. Res.* 9 (4), 545–567.
- Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transp. Res. Part B Methodol.* 42 (9), 759–770.
- Gipps, P.G., 1981. A behavioural car-following model for computer simulation. *Transp. Res. Part B Methodol.* 15 (2), 105–111.
- Gong, Y., Gu, S., Guan, L., 2023. Fine-Grained Classification Network for Fish Based on the Attention Mechanism and EfficientNet. *Association for Computing Machinery*.
- Greenberg, H., 1959. An analysis of traffic flow. *Oper. Res.* 7 (1), 79–85.
- Greenshields, B.D., Bibbins, J., Channing, W., Miller, H., 1935. A study of traffic capacity. Highway Research Board proceedings, Washington, DC, National Research Council.
- Gu, Z., Saberi, M., Sarvi, M., Liu, Z., 2018. A big data approach for clustering and calibration of link fundamental diagrams for large-scale network simulation applications. *Transp. Res. Part C Emerg. Technol.* 94, 151–171.
- Gu, Z., Wang, Z., Liu, Z., Saberi, M., 2022. Network traffic instability with automated driving and cooperative merging. *Transp. Res. Part C Emerg. Technol.* 138, 103626.
- Gunter, G., Gloudemans, D., Stern, R.E., McQuade, S., Bhadani, R., Bunting, M., Delle Monache, M.L., Lysecky, R., Seibold, B., Sprinkle, J., Piccoli, B., Work, D.B., 2021. Are commercially implemented adaptive cruise control systems string stable? *IEEE Trans. Intell. Transp. Syst.* 22 (11), 6992–7003.
- Hamdar, S.H., Mahmassani, H.S., Treiber, M., 2015. From behavioral psychology to acceleration modeling: Calibration, validation, and exploration of drivers' cognitive and safety parameters in a risk-taking environment. *Transp. Res. Part B Methodol.* 78, 32–53.
- Hammit, B.E., Ghasemzadeh, A., James, R.M., Ahmed, M.M., Young, R.K., 2018. Evaluation of weather-related freeway car-following behavior using the SHRP2 naturalistic driving study database. *Transp. Res. Part F Traffic Psychol. Behav.* 59, 244–259.
- Hao, H., Ma, W., Xu, H., 2016. A fuzzy logic-based multi-agent car-following model. *Transp. Res. Part C Emerg. Technol.* 69, 477–496.
- He, J., Chen, J.-N., Liu, S., Kortylewski, A., Yang, C., Bai, Y., Wang, C., 2022. TransFG: A transformer architecture for fine-grained recognition. *Proc. AAAI Conf. Artif. Intell.* 36 (1), 852–860.
- He, X., Xu, D., Zhao, H., Moze, M., Aioun, F., Guillemard, F., 2018. A human-like trajectory planning method by learning from naturalistic driving data. In: 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, IEEE.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image Recognition. *IEEE Computer Society, Las Vegas, NV*.
- Helly, W., 1959. Simulation of bottlenecks in single-lane traffic flow. In: Warren, M.I., Motors, G. (Eds.), *Proceedings of the Symposium on Theory of Traffic Flow*.
- Herman, R., Montroll, E.W., Potts, R.B., Rothery, R.W., 1959. Traffic dynamics: Analysis of stability in car following. *Oper. Res.* 7 (1), 86–106.
- Hoogendoorn, S., Hoogendoorn, R., 2010. Calibration of microscopic traffic-flow models using multiple data sources. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 368 (1928), 4497–4517.
- Hua, L.G., Wang, Y., 1978. Application of number theory in approximate analysis. Science Press, Beijing.
- Huang, Y.X., Jiang, R., Zhang, H.M., Hu, M.B., Tian, J.F., Jia, B., Gao, Z.Y., 2018. Experimental study and modeling of car-following behavior under high speed situation. *Transp. Res. Part C Emerg. Technol.* 97, 194–215.
- Huband, S., Hingston, P., Barone, L., While, L., 2006. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Trans. Evol. Comput.* 10 (5), 477–506.
- Hwang, C.-L., Yoon, K., 1981. Methods for multiple attribute decision making. In: Hwang, C.-L., Yoon, K. (Eds.), *Multiple Attribute Decision Making: Methods and Applications a State-of-the-Art Survey*. Springer, Berlin Heidelberg, Berlin, pp. 58–191.
- Jiang, R., Wu, Q., Zhu, Z., 2001. Full velocity difference model for a car-following theory. *Phys. Rev. E* 64 (1), 017101.

- Jin, P.J., Yang, D., Ran, B., 2014. Reducing the error accumulation in car-following models calibrated with vehicle trajectory data. *IEEE Trans. Intell. Transp. Syst.* 15 (1), 148–157.
- Joueiai, M., Van Lint, H., Hoogendoorn, S., 2013. Generic Solutions for Consistency Problems in Multi-Scale Traffic Flow Models - Analysis and Preliminary Results. The Hague, IEEE.
- Kaur, P., Sikka, K., Wang, W., Belongie, S., Divakaran, A., 2019. Foodx-251: a dataset for fine-grained food classification. *arXiv preprint, arXiv:1907.06167*.
- Kim, J., Mahmassani, H.S., 2011. Correlated parameters in driving behavior models: Car-following example and implications for traffic microsimulation. *Transp. Res. Rec.* 2249, 62–77.
- Kometani, E., Sasaki, T., 1959. A safety index for traffic with linear spacing. *Oper. Res.* 7 (6), 704–720.
- Krajewski, R., Bock, J., Kloeker, L., Eckstein, L., 2018. The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, IEEE.
- Kurtc, V., Treiber, M., 2020. Simulating bicycle traffic by the intelligent-driver model-Reproducing the traffic-wave characteristics observed in a bicycle-following experiment. *J. Traffic Transp. Eng.* 7 (1), 19–29.
- Leclercq, L., 2007. Hybrid approaches to the solutions of the “Lighthill-Whitham-Richards” model. *Transp. Res. Part B Methodol.* 41 (7), 701–709.
- Li, L., Chen, X.M., Zhang, L., 2016. A global optimization algorithm for trajectory data based car-following model calibration. *Transp. Res. Part C Emerg. Technol.* 68, 311–332.
- Li, T., Ngoduy, D., Hui, F., Zhao, X., 2020. A car-following model to assess the impact of V2V messages on traffic dynamics. *Transp. B Transp. Dyn.* 8 (1), 150–165.
- Liang, J.J., Suganthan, P.N., Deb, K., 2005. Novel composition test functions for numerical global optimization. In: Proceedings 2005 IEEE Swarm Intelligence Symposium, 2005. SIS 2005., Pasadena, CA, IEEE.
- Lin, T.Y., RoyChowdhury, A., Maji, S., 2015. Bilinear CNN Models for Fine-Grained Visual Recognition. 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, IEEE.
- Liu, H., Miao, X., Mertz, C., Xu, C., Kong, H., 2021. CrackFormer: Transformer network for fine-grained crack detection. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, IEEE.
- Liu, Z., Lyu, C., Huo, J., Wang, S., Chen, J., 2022. Gaussian process regression for transportation system estimation and prediction problems: The deformation and a Hat Kernel. *IEEE Trans. Intell. Transp. Syst.* 23 (11), 22331–22342.
- Liu, Z., Lyu, C., Wang, Z., Wang, S., Liu, P., Meng, Q., 2023. A gaussian-process-based data-driven traffic flow model and its application in road capacity analysis. *IEEE Trans. Intell. Transp. Syst.* 24 (2), 1544–1563.
- Ma, J., Sun, J., Li, K., Zhang, L., 2011. A study on multi-resolution scheme of macroscopic-microscopic traffic simulation model. *IEEE Conference on Intelligent Transportation Systems*.
- Mahmassani, H.S., Williams, J.C., Herman, R., 1984. Investigation of network-level traffic flow relationships: some simulation results. *Transp. Res. Rec.* 971, 121–130.
- Makridis, M., Mattas, K., Anesiadou, A., Ciuffo, B., 2021. OpenACC. An open database of car-following experiments to study the properties of commercial ACC systems. *Transp. Res. Part C Emerg. Technol.* 125, 103047.
- Messina, N., Amato, G., Esuli, A., Falchi, F., Gennaro, C., Marchand-Maillet, S., 2021. Fine-grained visual textual alignment for cross-modal retrieval using transformer encoders. *ACM Trans. Multimed. Comput. Commun. Appl.* 17 (4), 128.
- Michaels, R., 1963. Perceptual factors in car-following. *Proc. of 2nd ISTTT*.
- Milanés, V., Shladover, S.E., 2014. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transp. Res. Part C Emerg. Technol.* 48, 285–300.
- Montanino, M., Montell, J., Punzo, V., 2021. From homogeneous to heterogeneous traffic flows: Lp String stability under uncertain model parameters. *Transp. Res. Part B Methodol.* 146, 136–154.
- Ni, D., Leonard, J.D., Jia, C., Wang, J., 2016. Vehicle longitudinal control and traffic stream modeling. *Transp. Sci.* 50 (3), 1016–1031.
- Niu, Z., Zhong, G., Yu, H., 2021. A review on the attention mechanism of deep learning. *Neurocomputing* 452, 48–62.
- Ossen, S., Hoogendoorn, S.P., 2008. Validity of trajectory-based calibration approach of car-following models in presence of measurement errors. *Transp. Res. Rec.* 2088, 117–125.
- Pan, Y.A., Guo, J., Chen, Y., Cheng, Q., Li, W., Liu, Y., 2024. A fundamental diagram based hybrid framework for traffic flow estimation and prediction by combining a Markovian model with deep learning. *Expert Syst. Appl.* 238, 122219.
- Papathanasopoulou, V., Antoniou, C., 2015. Towards data-driven car-following models. *Transp. Res. Part C Emerg. Technol.* 55, 496–509.
- Pei, X., Pan, Y., Wang, H., Wong, S.C., Choi, K., 2016. Empirical evidence and stability analysis of the linear car-following model with gamma-distributed memory effect. *Phys. A Stat. Mech. Its Appl.* 449, 311–323.
- Punzo, V., Ciuffo, B., Montanino, M., 2012. Can results of car-following model calibration based on trajectory data be trusted? *Transp. Res. Rec.* 2315, 11–24.
- Punzo, V., Simonelli, F., 2005. Analysis and comparison of microscopic traffic flow models with real traffic microscopic data. *Transp. Res. Rec.* 1934 (1), 53–63.
- Punzo, V., Zheng, Z., Montanino, M., 2021. About calibration of car-following dynamics of automated and human-driven vehicles: Methodology, guidelines and codes. *Transp. Res. Part C Emerg. Technol.* 128, 103165.
- Qu, X., Wang, S., Zhang, J., 2015. On the fundamental diagram for freeway traffic: A novel calibration approach for single-regime models. *Transp. Res. Part B Methodol.* 73, 91–102.
- Qu, X., Tang, P., Zou, Z., Cheng, Y., Dong, J., Zhou, P., Xu, Z., 2020. Fine-Grained Iterative Attention Network for Temporal Language Localization in Videos. Association for Computing Machinery.
- Rakha, H., Wang, W., 2009. Procedure for calibrating Gipps car-following model. *Transp. Res. Rec.* 2124, 113–124.
- Saifuzzaman, M., Zheng, Z., 2014. Incorporating human-factors in car-following models: A review of recent developments and research needs. *Transp. Res. Part C Emerg. Technol.* 48, 379–403.
- Saifuzzaman, M., Zheng, Z., Mazharul Haque, M., Washington, S., 2015. Revisiting the task-capability interface model for incorporating human factors into car-following models. *Transp. Res. Part B Methodol.* 82, 1–19.
- Sarwinda, D., Paradisa, R.H., Bustamam, A., Anggia, P., 2021. Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer. *Procedia Comput. Sci.* 179, 423–431.
- Seo, T., Kawasaki, Y., Kusakabe, T., Asakura, Y., 2019. Fundamental diagram estimation by using trajectories of probe vehicles. *Transp. Res. Part B Methodol.* 122, 40–56.
- Seu, 2019. SEU vehicle trajectory data. accessed February 8, 2023. <http://seuttraffic.com/>.
- Sharma, A., Zheng, Z., Bhaskar, A., 2019. Is more always better? The impact of vehicular trajectory completeness on car-following model calibration and validation. *Transp. Res. Part B Methodol.* 120, 49–75.
- Tian, J., Zhang, H.M., Treiber, M., Jiang, R., Gao, Z.Y., Jia, B., 2019. On the role of speed adaptation and spacing indifference in traffic instability: Evidence from car-following experiments and its stochastic model. *Transp. Res. Part B Methodol.* 129, 334–350.
- Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. e* 62 (2), 1805–1824.
- Treiber, M., Kesting, A., 2013a. Microscopic calibration and validation of car-following models – A systematic approach. *Procedia - Soc. Behav. Sci.* 80, 922–939.
- Treiber, M., Kesting, A., 2013b. Traffic flow dynamics: Data, models and simulation. Springer, Berlin.
- Underwood, R.T., 1961. Speed, volume, and density relationships. <https://trid.trb.org/view/115231> (accessed February 8, 2023).
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention is all you need. In: 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, NIPS.
- Wang, Z., Rangaiah, G.P., 2017. Application and analysis of methods for selecting an optimal solution from the pareto-optimal front obtained by multiobjective optimization. *Ind. Eng. Chem. Res.* 56 (2), 560–574.
- Wang, Z., Shi, Y., Tong, W., Gu, Z., Cheng, Q., 2023. Car-following models for human-driven vehicles and autonomous vehicles: A systematic review. *J. Transp. Eng. Part A Syst.* 149 (8), 04023075.

- Wang, H., Wang, W., Chen, J., Xu, C., Li, Y., 2019. Can we trust the speed–spacing relationship estimated by car-following model from non-stationary trajectory data? *Transp. A Transp. Sci.* 15 (2), 263–284.
- Wiedemann, R., 1974. Simulation des straßenverkehrsflusses. Master's thesis, Schriftenreihe des Instituts für Verkehrswesen.
- Würth, A., Binois, M., Goatin, P., Göttlich, S., 2022. Data-driven uncertainty quantification in macroscopic traffic flow models. *Adv. Comput. Math.* 48 (6), 75.
- Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K., 2017. Aggregated residual transformations for deep neural networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, IEEE.
- Xu, T., Laval, J., 2020. Statistical inference for two-regime stochastic car-following models. *Transp. Res. Part B Methodol.* 134, 210–228.
- Yan, Q., Sun, Z., Gan, Q., Jin, W.L., 2018. Automatic identification of near-stationary traffic states based on the PELT changepoint detection. *Transp. Res. Part B Methodol.* 108, 39–54.
- Yang, L., Yu, H., Cheng, Y., Mei, S., Duan, Y., Li, D., Chen, Y., 2021. A dual attention network based on efficientNet-B2 for short-term fish school feeding behavior analysis in aquaculture. *Comput. Electron. Agric.* 187, 106316.
- Yuan, Y., Zhang, Z., Yang, X.T., Zhe, S., 2021. Macroscopic traffic flow modeling with physics regularized Gaussian process: A new insight into machine learning applications in transportation. *Transp. Res. Part B Methodol.* 146, 88–110.
- Zhang, Q., Zhou, A., Zhao, S., Suganthan, P.N., Liu, W., Tiwari, S., 2008. Multiobjective optimization test instances for the CEC 2009 special session and competition. https://www.al-roomi.org/multimedia/CEC_Database/CEC2009/MultiObjectiveEA/CEC2009_MultiObjectiveEA_TechnicalReport.pdf (accessed February 8, 2023).
- Zhang, Q., Li, H., 2007. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.* 11 (6), 712–731.
- Zhang, J., Qu, X., Wang, S., 2018. Reproducible generation of experimental data sample for calibrating traffic flow fundamental diagram. *Transp. Res. Part A Policy Pract.* 111, 41–52.
- Zheng, H., Fu, J., Mei, T., Luo, J., 2017. Learning multi-attention convolutional neural network for fine-grained image recognition. In: 2017 IEEE International Conference on Computer Vision (ICCV), Venice, IEEE.
- Zheng, L., Jin, P.J., Huang, H., Gao, M., Ran, B., 2016. A vehicle type-dependent visual imaging model for analysing the heterogeneous car-following dynamics. *Transp. B Transp. Dyn.* 4 (1), 68–85.
- Zhong, R.X., Fu, K.Y., Sumalee, A., Ngoduy, D., Lam, W.H.K., 2016. A cross-entropy method and probabilistic sensitivity analysis framework for calibrating microscopic traffic models. *Transp. Res. Part C Emerg. Technol.* 63, 147–169.
- Zitzler, E., Thiele, L., 1999. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.* 3 (4), 257–271.
- Ztd., 2018. Zen traffic data. accessed February 8, 2023. <https://zen-traffic-data.net/english/>.