

Highlights

A Time-varying Shockwave Speed Model for Trajectory Reconstruction using Lagrangian and Eulerian Observations

Yifan Zhang, Anastasios Kouvelas, Michail A. Makridis

- Integrate Lagrangian and Eulerian observations to reconstruct trajectories.
- Calibrate time-varying short-term shockwave speeds using the two types of data.
- Reconstruct trajectories for non-connected vehicles based on shockwave theory.
- Optimize trajectories by adding vehicle dynamics for better energy estimation.
- Evaluation on real-world datasets shows excellent performances from several aspects.

A Time-varying Shockwave Speed Model for Trajectory Reconstruction using Lagrangian and Eulerian Observations

Yifan Zhang^{a,*}, Anastasios Kouvelas^a and Michail A. Makridis^a

^a*Institute for Transport Planning and Systems, ETH Zurich, Zurich, 8093, Switzerland*

ARTICLE INFO

Keywords:

Shockwave Speed Calibration
Trajectory Reconstruction
Fixed Sensor
Connected Vehicle
Fundamental Diagram
Energy Consumption

ABSTRACT


Inference of detailed vehicle trajectories is crucial for applications such as traffic flow modeling, energy consumption estimation, and traffic flow optimization. Static sensors can provide only aggregated information, posing challenges in reconstructing individual vehicle trajectories. Shockwave theory is used to reproduce oscillations that occur between sensors. However, as the emerging of connected vehicles grows, probe data offers significant opportunities for more precise trajectory reconstruction. Existing methods rely on Eulerian observations (e.g., data from static sensors) and Lagrangian observations (e.g., data from probe vehicles) incorporating shockwave theory and car-following modeling. Despite these advancements, a prevalent issue lies in the static assignment of shockwave speed, which may not be able to reflect the traffic oscillations in a short time period caused by varying response times and vehicle dynamics. Moreover, energy consumption estimation is largely ignored. In response, this paper proposes a novel framework that integrates Eulerian and Lagrangian observations for trajectory reconstruction. The approach introduces a calibration algorithm for time-varying shockwave speed. The calibrated shockwave speed of the CV is then utilized for trajectory reconstruction of other non-connected vehicles based on shockwave theory. Additionally, vehicle and driver dynamics are introduced to optimize the trajectory and estimate energy consumption. The proposed method is evaluated using real-world datasets, demonstrating superior performance in terms of trajectory accuracy, reproducing traffic oscillations, and estimating energy consumption.

1. Introduction

Vehicle trajectory information with a high sampling frequency plays a key role in many domains, including but not limited to traffic flow modeling, energy consumption estimation, and traffic flow optimization (Daganzo, 1997; Zhang, Chen, Wang, Zheng and Wu, 2022; Li, Jiang, He, Chen and Zhou, 2020; Guo, Li and (Jeff) Ban, 2019; Fiori, Arcidiacono, Fontaras, Makridis, Mattas, Marzano, Thiel and Ciuffo, 2019; He, Makridis, Fontaras, Mattas, Xu and Ciuffo, 2020). Collecting such granular information requires either a dense network of roadside sensors or transmission of location data from every individual vehicles. However, both approaches are financially and laboriously burdensome and maybe even infeasible in practice. In light of these constraints, it is crucial but challenging to infer the complete set of trajectories given limited sensing data rather than observing all of them directly. Thus, accurate reconstruction of trajectories with a minimal amount of Lagrangian observations provides a clear and significant contribution toward mixed traffic conditions.

Some methods are tailored to reconstruct trajectories using fixed sensors, including loop detectors, video cameras, radars, etc., which are deployed at a fixed location on the roadside to detect passing vehicles. This type of sensing data typically provides the information at an aggregated level, such as vehicle counts, average speed, and instantaneous velocities of passing vehicles, which is known as Eulerian observation. The data can be adopted to analyze and optimize the traffic flow from macroscopic views (Kachroo, Ozbay and Hobeika, 2001; Wang, Daamen, Hoogendoorn and van Arem, 2011), estimate travel time (Coifman, 2002), and reconstruct trajectory (Coifman, 2002; van Lint and van der Zijpp, 2003; Lint, 2010). The shockwave theory is usually applied to this type of data to infer vehicles' trajectory, given that the traffic state along the shockwave is the same. Conventionally, the traffic state is assumed to propagate in the traffic flow at a fixed speed, referred to as the shockwave speed. Although the value of shockwave speed is constant on average over a long time period, the fact is that the traffic state may not propagate at the same speed in the short time

*Corresponding author

 yifan.zhang@ivt.baug.ethz.ch (Y. Zhang); kouvelas@ethz.ch (A. Kouvelas); michail.makridis@ivt.baug.ethz.ch (M.A. Makridis)

ORCID(s): 0000-0001-8882-4327 (Y. Zhang)

period due to the existence of varying response times and vehicle dynamics (Makridis, Leclercq, Ciuffo, Fontaras and Mattas, 2020), leading to sub-optimal reconstructed trajectories.

With the deployment of connected vehicles (CVs), mobile sensing data from CVs that report location information through vehicular networks, termed Lagrangian observation, largely contributes to trajectory reconstruction. The collected trajectory data includes temporal and spatial data of these CVs, providing a rich information to analyze human driving behaviors (Makridis, Anesiadou, Mattas, Fontaras and Ciuffo, 2023; Suarez, Makridis, Anesiadou, Komnos, Ciuffo and Fontaras, 2022; Wang, Wei and Chen, 2020), estimate fundamental diagram (FD) (Seo, Kawasaki, Kusakabe and Asakura, 2019; Li, Chen, Zhou, Xie and Laval, 2022), and reconstruct the trajectory of vehicles (Montanino and Punzo, 2015; Yao, Liu, Jiang, Tang and Ran; Makridis and Kouvelas, 2023). The trajectory of the CVs can be enhanced with a high accuracy using the collected data through interpolation curves (Sun, Wu, Chu, Xie, Liu and Li), particle filters (Wei, Wang and Chen, 2021), maximum likelihood estimation (Hao, Boriboonsomsin, Wu and Barth, 2014), etc. The trajectory of non-connected vehicles (NCVs) is inferred by essentially replicating the trajectory of CVs based on Newells model, when the shockwave/backward speed in the model is a constant value (Montanino and Punzo, 2015; Yao et al.). However, such a manner may not be able to characterize realistic traffic scenarios, given the inherent variability in vehicle and driver dynamics, as previously discussed.

To overcome the above challenges of separately using Lagrangian or Eulerian observations, more attention is attracted to integrating both for reconstructing trajectories. Specifically, most of the methods (Deng, Cao, Ren, Ma and Zhu, 2023; Chen, Yin, Qin, Tang, Wang and Sun, 2022a; Mehran, Kuwahara and Naznin, 2012) are developed by deriving some reference points or candidate trajectories using one type of data and then optimizing them using the other type of data. Although the performance has been largely improved by the above hybrid methods, the following issues still need to be addressed. (1) Prevalent models often adhere to a fixed shockwave speed (Chen et al., 2022a; Mehran et al., 2012), while in reality, it has been observed to vary (Sakhare, Li and Bullock, 2023; Lei, Wang, Lu and Sun, 2014). A constant value of shockwave speed implies that independently of the current traffic conditions or disturbances the propagation of oscillations follows a consistent pattern which is unrealistic (Makridis et al., 2020). (2) Efforts to optimize time-varying shockwave speeds, as seen in (Deng et al., 2023), encounter limitations when there are fewer than two vehicles affected by the same shockwave. (3) Reconstructed trajectories can capture traffic flow dynamics but they are still unreliable for energy consumption estimations where a realistic distribution of observed accelerations is essential (Apostolakis, Makridis, Kouvelas and Ampountolas, 2023). (4) Calibration of car-following models for simulation will not be feasible when instantaneous vehicle dynamics are not realistic (Ciuffo, Makridis, Toledo and Fontaras, 2018; He, Montanino, Mattas, Punzo and Ciuffo, 2022; Zheng, Makridis, Kouvelas, Jiang and Jia, 2023).

To this end, we design a new hybrid method based on the trajectories of a limited number of CVs and the instantaneous speed of vehicles collected by a loop detector. Specifically, to address the issues associated with fixed shockwave speed, we propose an innovative algorithm to calibrate the time-varying shockwave speed using Lagrangian and Eulerian observations. Given the calibrated shockwave speed, we develop an algorithm to reconstruct the reference points for NCVs. Moreover, to fill the gap in energy consumption estimation, we apply a vehicle movement model, that accounts for vehicle and driver dynamics, to optimize the reference points and generate the final trajectory of NCVs. In this paper, we mainly focus on highway scenarios without considering the disturbance of traffic signals and intersections. Our main contributions are summarized as follows.

- We propose a new algorithm to calibrate the time-varying shockwave speed using the fixed sensor data and mobile sensor data. Specifically, our calibration algorithm is designed based on shockwave theory-compliant trajectory reconstruction methods. The algorithm aims to determine time-varying shockwave speed values that minimize the error between the reconstructed trajectory and the ground truth. We design and implement a Monte Carlo sampling-based method to achieve the aforementioned goal.
- We then design a trajectory reconstruction method using the time-varying shockwave speed and the data collected by fixed sensors. Each CV is considered as a leading vehicle, and our objective is to reconstruct the trajectory of the following vehicle of the CV. We derive reference points for the following vehicle using the CV's calibrated results according to shockwave theory. These reference points are further optimized to reconstruct the trajectory, which is achieved by applying the microsimulation free-flow acceleration model (MFC) to fulfill the constraint of vehicle and driver dynamics.

- We evaluate and validate our method using real-world datasets. We compare with several baseline models in terms of time headway accuracy, speed accuracy, fuel consumption accuracy, CO₂ emission accuracy, and the capability of reproducing traffic oscillations. The experiment results show that our method achieves better performance compared with other baseline models, especially in the case of a low CV penetration rate.
- We further introduce the concept of a reconstructed spatial-temporal area to assess the contribution of CVs with different average speeds. Our findings reveal that a lower average speed can contribute more to reconstructing a larger spatial-temporal area. Moreover, we explore the relationship between the amount of required information for calibration and the average speed of the CV. The results show that the average speed does not affect the amount of required information.

The paper is organized as follows. In Section 2, we review the literature on trajectory reconstruction and shockwave speed estimation. We first clarify the background and formulate the problem in Section 3. We then present the basic shockwave theory-based method and our method for calibrating time-varying shockwave speed in Section 4. The reconstruction process, including deriving reference points and reconstructing trajectory, is elaborated on in Section 5. We then show the experiment results of our method compared with baseline models in Section 6. In Section 7, we further analyze the contribution of CVs in reconstructing trajectories and the amount of required information for calibrating time-varying shockwave speed. We finally conclude our paper and discuss the future work in Section 8.

2. Literature Review

Trajectory reconstruction has been studied for decades using different types of data and methods. The data from fixed sensors is largely used to reconstruct trajectories based on FIFO and shockwave theory. In recent years, the trajectory of some connected vehicles has also been widely applied to improve the reconstruction accuracy, which is regarded as the data from mobile sensors. In the rest of the section, we mainly review the literature about different trajectory reconstruction methods. Moreover, as most of the methods including ours are based on the shockwave theory, where the shockwave speed is one of the most important parameters, we also review the literature on shockwave speed estimation.

2.1. Trajectory Reconstruction

The method of trajectory reconstruction is mainly differentiated based on the type of data it leverages.

Fixed sensors like loop detectors are very common in modern cities but they provide flow and potentially speed measurements at specific locations. Given the arrival time and velocity of vehicles collected by a fixed loop detector, (Coifman, 2002) firstly proposed to reconstruct the trajectory using the shockwave theory. Specifically, the vehicle's trajectory is reconstructed by concatenating the collected velocity of the following vehicles under the assumption that the velocity dissipates along the road at a shockwave speed. (van Lint and van der Zijpp, 2003; Lint, 2010) proposed to use data from multiple loop detectors. (van Lint and van der Zijpp, 2003) Compared with (Coifman, 2002), (van Lint and van der Zijpp, 2003) developed a method to calculate the linear velocity instead of applying the collected velocities as constant values to reconstruct the trajectories. (Lint, 2010) proposed filtered inverse speed-based (FISB) to construct a speed map for the reconstructed spatial-temporal space. The free flow and congestion are separately considered to estimate the speed of individual vehicles and further reconstruct trajectories.

The second type is the data from mobile sensors. Usually, this type of data is reported by the individual CVs. One of the objectives of using this type of data is to reconstruct the trajectory of CVs since the frequency of the reported data from CVs may be too low to get enough information. Thus, many probabilistic methods are proposed to improve the resolution of the trajectory (Hao et al., 2014; Wan, Vahidi and Luckow, 2016; Wei et al., 2021). Moreover, trajectory reconstruction of NCVs is also one of the objectives. Usually, the historical trajectory of NCVs is reconstructed by replicating the trajectory of CVs using Newell-based models (Montanino and Punzo, 2015; Yao et al.). Besides, a huge amount of effort is made to reconstruct the future trajectory of the vehicle with the machine learning techniques (Liu, Mao, Fang, Zhu and Meng, 2021). In the rest of this paper, we mainly discuss reconstructing the fully sampled historical trajectory of NCVs.

The third one is to integrate the data from both fixed sensors and mobile sensors. More and more researchers focus on leveraging these two types of data to improve overall accuracy. Specifically, (Mehran et al., 2012) proposed a data fusion framework to reconstruct the trajectories at signalized intersections based on the kinematic wave theory and variational solution of (Daganzo, 2005a,b). The temporal-spatial space of the trajectory is represented as a cumulative

surface in a three-dimensional space. The data collected by the loop detector is the height of the cumulative surface at the specific temporal-spatial point. The trajectory of the CV is interpreted as a contour on the surface of cumulative curves. The other trajectories can be reproduced based on the estimation of this surface. Based on (Mehran et al., 2012), (Chen, Yin, Tang, Tian and Sun, 2022b) designed a new framework, which classifies the temporal-spatial space into four regions and applies a Particle Filter-based fusion method to estimate the trajectory of vehicles at signalized intersections. (Rey, Jin and Ritchie, 2019; Deng et al., 2023) proposed to estimate the trajectory of vehicles on multi-lane roads based on the data from multiple loop detectors and CVs. A new variable of a vehicle's order is defined to represent its cumulative flows at different times and locations. (Chen et al., 2022a) designed a new framework to integrate both micro and macro models. The macro models based on shockwave theory aim to generate the velocity contour map, which includes the reference velocity. The micro model produces several candidate trajectories according to the trajectory of CVs. The final trajectories are optimized using the velocity contour map and candidate trajectories. All of the above methods are implemented based on wave theory, where the shockwave speed is one of the most important parameters.

The main issue of the above methods is that the value of the shockwave speed is usually set as a constant value, which implies homogeneous vehicle and driver dynamics in the traffic flow. However, it is unrealistic according to the real-world observations (Makridis et al., 2020). Moreover, the reconstructed trajectory is derived through a mathematical formulation or connecting several reference points directly, which may not satisfy the vehicle dynamics and can not be used for estimating energy consumption.

2.2. Shockwave Speed Estimation

As mentioned above, shockwave speed estimation plays a key role in reconstructing individual vehicle trajectories. Shockwave speed is an important parameter in describing the traffic flow. Many researchers have proposed different methods to estimate the value of shockwave speed in actual traffic. (Seo et al., 2019) performs a linear regression between the flow and density for each pair of probe vehicles to estimate the shockwave speed. The final shockwave speed is the mean of all the regression results. (Mehran et al., 2012) estimates the shockwave speed using the maximum flow, jam density, and free flow speed, which is largely determined by the road infrastructure. (Anuar and Cetin, 2017; Deng et al., 2023) locate the inflection point in the congested state and calculate the shockwave speed for each shockwave. Although (Deng et al., 2023) estimate different shockwave speeds for different periods to reconstruct the trajectories, it requires the trajectory of at least two CVs in the same shockwave to estimate the shockwave speed, which largely limits the performance with a lower penetration rate of CVs.

In summary, the trajectory reconstruction based on shockwave theory using fixed sensor and mobile sensor data has been widely studied. However, the challenges, including estimating and applying dynamic shockwave speeds, still remain unsolved. Moreover, the evaluation of the reconstructed trajectory usually focuses on the accuracy of the trajectory and largely ignores the energy aspects.

3. Background and Problem Formulation

We mainly consider the two following types of data to reconstruct the trajectories. The first one contains the arrival time and instantaneous speed of vehicles, collected by a fixed sensor deployed at the start of a road segment. The second one includes the trajectory of CVs, which are distributed randomly in the traffic flow at a lower penetration rate. As shown in Fig. 1, the fixed sensor is deployed at the start point of the road, and the blue vehicles are CVs. The objective is to use the above data to reconstruct the trajectories of other vehicles, namely, the black ones in Fig. 1.

Denote t_i and v_i as the arrival time and instantaneous speed of i^{th} vehicle, respectively, which is collected by the fixed sensor at position x_0 . Denote $\Pi_i = \{x_i^\tau, \tau \in [t_i, T]\}$ as the trajectory of i^{th} vehicle, where x_i^τ is the longitudinal position of i^{th} vehicle at time τ and T is the end time of the trajectory. Denote I as the union of all vehicles and J as the union of all CVs, where $J \subset I$ and $|J| < |I|$. The problem of reconstructing the unknown trajectories is formulated as Eq. 1.

$$\Pi_i = f(\Pi_J, t_J, v_J), \text{ where } i \in I \ \& \ i \notin J. \quad (1)$$

We aim to derive the function $f(\cdot)$ to reconstruct $\{\Pi_i, i \in I \ \& \ i \notin J\}$ given Π_J, t_J , and v_J .

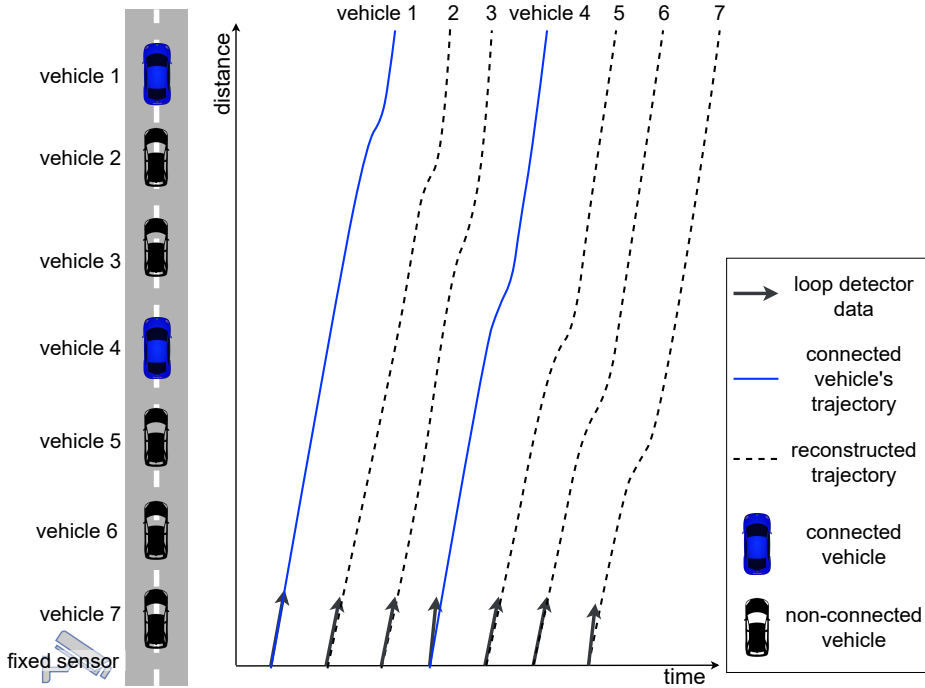


Figure 1: Illustration of the loop detector's information and connected vehicles' information.

4. Time-varying Shockwave Speed

The shockwave theory is the basis for reconstructing trajectories using different sources of data. Current methods usually assume a constant value for the parameter shockwave speed, which is not capable of handling dynamic and real traffic scenarios (Sakhare et al., 2023; Lei et al., 2014). To address this issue, we propose to calibrate the time-varying shockwave speed for trajectory reconstruction using the data from a double loop detector and CVs. In the rest of the section, we first introduce the preliminaries of shockwave theory and trajectory reconstruction. We then present our algorithm to calibrate the time-varying shockwave speed values on the basis of the shockwave theory.

4.1. Basics of Shockwave Theory

The traffic shockwave is caused by some disturbances and propagates backward to the following vehicles, which is firstly proposed by Lighthill, Whitham and Richards (LWR) (Lighthill and Whitham, 1955). LWR model describes the relation between flow and density, which vary with location and time as shown in Eq. 2.

$$\frac{\partial k(x, t)}{\partial t} + \frac{\partial q(x, t)}{\partial x} = 0, \quad (2)$$

where $k(x, t)$ is the density and $q(x, t)$ is the flow at location x and time t . Eq. 2 can be further rewritten as Eq. 3.

$$\frac{1}{w} \frac{\partial q(x, t)}{\partial t} + \frac{\partial q(x, t)}{\partial x} = 0, \quad (3)$$

where $w = \frac{\partial q}{\partial k}$ is the shockwave speed.

Following by LWR, Newell (Newell, 1993) proposed a triangular fundamental diagram (FD) as Eq. 4.

$$q(k) = \begin{cases} v_f k, & k \leq k_m \\ w(k_j - k), & k > k_m \end{cases} \quad (4)$$

where v_f is the free flow speed, k_m is the road capacity, and k_j is the jam density. The triangle FD is shown as Fig. 2(a). The traffic states propagate among vehicles at the same speed on the same leg of the triangle, namely, v_f for the free

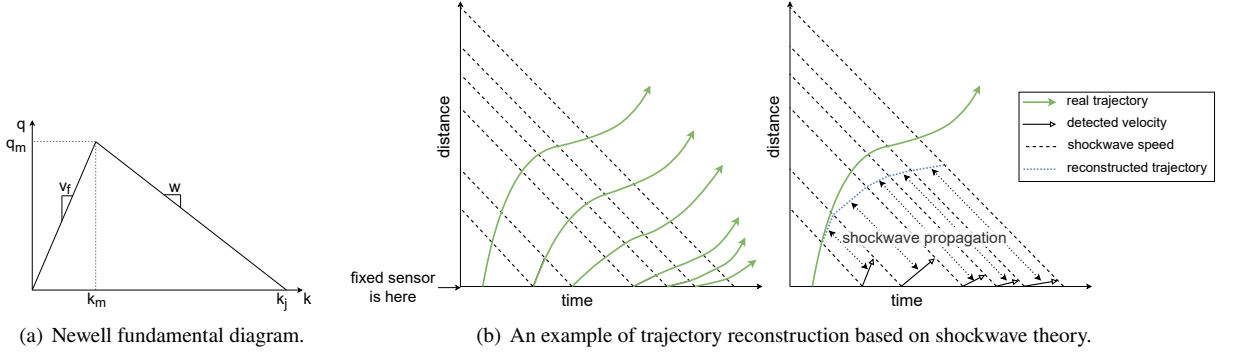


Figure 2: Trajectory reconstruction based on shockwave theory.

flow and w for the congested flow. According to Eq. 4 and Fig. 2(a), the traffic states are the same in one shockwave, which makes traffic states predictable given the partial traffic state in the shockwave. The most important parameter to derive the traffic states is the shockwave speed, indicating how fast the traffic state of the leading vehicle is passed to the following ones.

4.2. Trajectory Reconstruction Based on Shockwave Theory

According to (Coifman, 2002), the trajectory of vehicles can be reconstructed given the arrival time and instantaneous speed of their following vehicles detected by the loop detector. Based on the above assumptions that the traffic state along the shockwave is the same, the trajectory of a vehicle is reconstructed by concatenating the trajectory segment of its following vehicles, as shown in Fig. 2(b). Specifically, the trajectory segment of the following vehicle can be derived using the detected instantaneous speed of the vehicle. Then, these trajectory segments can be further concatenated following the shockwave propagation direction to construct the trajectory of the leading vehicle.

To simplify the reconstruction process, the reconstructed trajectory of each vehicle contains multiple line segments and each one is represented as a linear function in the temporal-spatial space, denoted as $g_i^{\tau k}(\tau)$, where i is the number of the vehicle and k is the k^{th} following vehicle of vehicle i . As shown in Eq. 5, the slope of $g_i^{\tau k}(\tau)$ is v_{i+k} , namely, the speed of k^{th} following vehicle given the subject leading vehicle i , and \hat{x}_i^{τ} is the reconstructed trajectory point of vehicle i at time τ . Moreover, the shockwave propagation line that represents the propagation of speed v_{i+k} is also modeled as a linear function $h_i^{\tau k}(\tau)$ with slope $-w$, as shown in Eq. 6. The notations are illustrated in Fig. 3 for easy understanding.

$$g_i^{\tau k}(\tau) = \begin{cases} v_i \tau + (x_0 - v_i t_i), & k = 0 \\ v_{i+k} \tau + (\hat{x}_i^{\tau k-1} - v_{i+k} \tau_{k-1}), & k > 0 \end{cases} \quad (5)$$

$$h_i^{\tau k}(\tau) = -w\tau + (x_0 + w t_{i+k+1}). \quad (6)$$

As mentioned above, the arrival time and instantaneous speed of vehicles at a fixed location x_0 are known. For obtaining the first trajectory segment of vehicle i , i.e., $k = 0$, we calculate $(\tau_0, \hat{x}_i^{\tau_0})$ which is the point of intersection of $g_i^{\tau_0}(\tau)$ and $h_i^{\tau_0}(\tau)$. Thus, the first trajectory segment of vehicle i is the line from point (t_i, x_0) to $(\tau_0, \hat{x}_i^{\tau_0})$. Given $(\tau_0, \hat{x}_i^{\tau_0})$, v_{i+1} , w , and (t_{i+2}, x_0) , we can obtain $g_i^{\tau_1}(\tau)$ and $h_i^{\tau_1}(\tau)$ so that the second trajectory segment is the line from $(\tau_0, \hat{x}_i^{\tau_0})$ to $(\tau_1, \hat{x}_i^{\tau_1})$. As such, $(\tau_k, \hat{x}_i^{\tau_k})$, $g_i^{\tau k}(\tau)$, $h_i^{\tau k}(\tau)$ can be further derived to complete the trajectory of vehicle i by repeating the above procedures. In the rest of the paper, the process of calculating the k^{th} trajectory segment is abbreviated as reconstruction step k .

4.3. Calibration Algorithm

Several studies in the literature discuss the estimation of shockwave speed (Seo et al., 2019; Deng et al., 2023). Some of them have proved that the value should not be fixed (Deng et al., 2023; Anuar and Cetin, 2017). However,

Algorithm 1 Time-varying Shockwave Speed Calibration

Input: Arrival time $\{t_i, i \in I\}$ and speed $\{v_i, i \in I\}$ collected by loop detector, trajectory of CVs $\{\Pi_j, j \in J\}$, upper bound and lower bound of shockwave speed $[w_{ub}, w_{lb}]$, number of samples N , accept thresholds ϵ , maximum iterations M

Output: Time-varying shockwave speed of CVs $\{W_j, j \in J\}$, where $W_j = \{w_j^{\tau_k}, k \in [0, 1, 2, \dots, k_{end}]\}$

```
1: for  $j \in J$  do
2:    $D = \{(t_i, v_i), \forall t_i > t_j\}$ 
3:   Sort  $D$  in ascending order of  $t_i$ 
4:    $k = 0$ 
5:   for  $(t_i, v_i) \in D$  do do
6:      $iter = 0$ 
7:     while True do
8:       Sample  $N$  samples from uniform distribution  $U(w_{ub}, w_{lb})$ 
9:       Using the samples  $\mathbf{w} = \{w_n, n \in [1, 2, \dots, N]\}_j^{\tau_k}$  to derive  $g_j^{\tau_k}(\tau)$  and  $\tilde{h}_j^{\tau_k}(\tau, w_n)$ 
10:      Calculate the point of intersection  $(\tau_k, \hat{x}_j^{\tau_k})_{w_n} = g_j^{\tau_k}(\tau) \otimes \tilde{h}_j^{\tau_k}(\tau, w_n)$ 
11:      Measure errors  $\mathbf{err} = \{||(\tau_k, \hat{x}_j^{\tau_k})_{w_n}, (\tau_k, x_j^{\tau_k})||, n \in [1, 2, \dots, N]\}$ 
12:       $w^* = \arg \min_w \mathbf{err}$ 
13:      if  $\min(\mathbf{err}) < \epsilon$  then
14:         $w_j^{\tau_k} = w^*$ 
15:        break
16:      else if  $(\hat{x}_j^{\tau_k})_{w^*} < x_j^{\tau_k}$  then
17:        Increase  $v_i$ 
18:      else if  $(\hat{x}_j^{\tau_k})_{w^*} > x_j^{\tau_k}$  then
19:        Decrease  $v_i$ 
20:      end if
21:       $iter = iter + 1$ 
22:      if  $iter > M$  then
23:         $w_j^{\tau_k} = w^*$ 
24:        break
25:      end if
26:    end while
27:     $k = k + 1$ 
28:  end for
29: end for
```

the shockwave speed in Coifman's method and its variants is set as a constant value, which may lead to sub-optimal results. This motivated this work to calibrate time-varying shockwave speed values.

Besides the loop detector data, the trajectory data reported by CVs is leveraged to achieve this goal. The objective of calibration is to find the optimal shockwave speed value $w_j^{\tau_k}$ for CV j at reconstruction step k , which minimizes the error between the ground truth trajectory and the reconstructed one, as shown in Fig. 3. The objective function is formulated as Eq. 7.

$$w_j^{\tau_k} = \arg \min_w ||g_j^{\tau_k}(\tau) \otimes \tilde{h}_j^{\tau_k}(\tau), (\tau_k, x_j^{\tau_k})||, \quad (7)$$

where the operation \otimes is to calculate the point of intersection, $||\cdot, \cdot||$ is the distance between two points, and $\tilde{h}_j^{\tau_k}(\tau)$ is the time-varying variant of Eq. 6. Specifically, the time-varying shockwave line function is illustrated in Eq. 8.

$$\tilde{h}_j^{\tau_k}(\tau) = -w_j^{\tau_k} \tau + (x_0 + w_j^{\tau_k} t_{j+k+1}). \quad (8)$$

The optimization problem in Eq. 7 can not be solved explicitly since the ground truth trajectory can not be represented using the enclosed mathematical formulation so that the point $(\tau_k, x_j^{\tau_k})$ can not be determined before

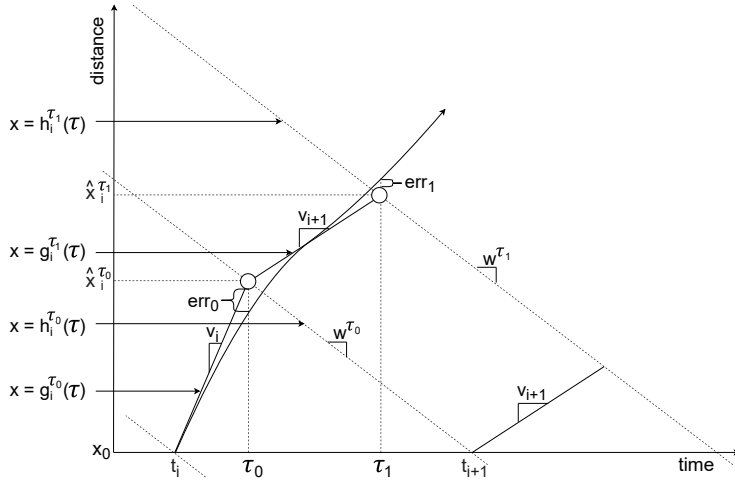


Figure 3: Process of shockwave speed calibration.

knowing $\tilde{h}_j^{\tau_k}(\tau)$ and $g_j^{\tau_k}(\tau)$. However, the slope $-w_j^{\tau_k}$ of $\tilde{h}_j^{\tau_k}(\tau)$ is the unknown parameter to be optimized. Namely, $(\tau_k, x_j^{\tau_k})$ is indeterminate leading that $w_j^{\tau_k}$ in Eq. 7 can not be solved explicitly. Thus, we propose to solve this problem through the Monte Carlo sampling-based method. Specifically, for each reconstruction step k , we generate N samples of $w_j^{\tau_k}$, denoted as $\{w_n, n \in [1, 2, \dots, N]\}_j^{\tau_k}$. For each sampled value w_n of $w_j^{\tau_k}$, we then calculate the point of intersection between the trajectory line $g_j^{\tau_k}(\tau)$ and the shockwave line $\tilde{h}_j^{\tau_k}(\tau, w_n)$. Given the point of intersection, we can derive the error between the point of intersection and the corresponding point on the ground truth trajectory, denoted as $\{err_n, n \in [1, 2, \dots, N]\}$ corresponding to $\{w_n, n \in [1, 2, \dots, N]\}_j^{\tau_k}$, respectively. Note that considering the reconstructed trajectory may have a longer travel time than the ground truth, we calculate the time difference under the same position between the ground truth point and the point of intersection. The error in the temporal dimension is independent of speed, which is better than the one in the spatial dimension for optimization. Given $\{err_n, n \in [1, 2, \dots, N]\}$, the minimal error is compared with a pre-defined threshold ϵ . The corresponding w_n is kept for the reconstruction step k if the above minimal error is smaller than ϵ .

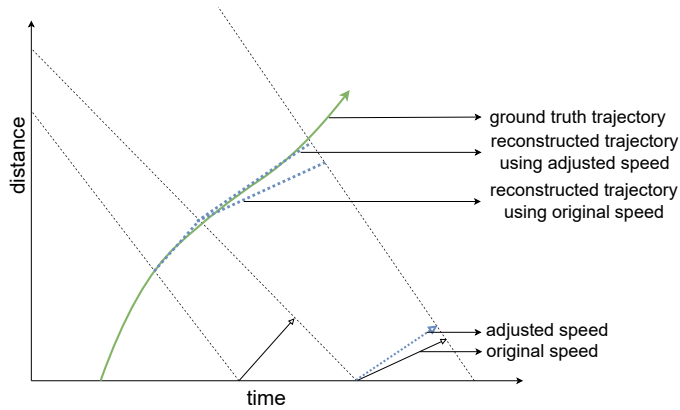


Figure 4: Adjust the speed collected by the loop detector.

However, as the value of $w_j^{\tau_k}$ is bounded by an upper and a lower bound, it could be possible that there is no such w_n that can achieve such a small error. To solve this problem, we optimize the value of both $w_j^{\tau_k}$ and v_{j+k} . Translating the speed of the following vehicle to reconstruct the trajectory of the leading vehicle can be modeled as a quantification of the impact of the leading vehicle on the following ones. Nonetheless, the stochasticity of such impacts can not be

ignored, leading to the aforementioned challenge of solving the optimal $w_j^{\tau_k}$. Thus, to well handle the stochasticity, we adjust the speed of the following vehicles to approximate the ground truth trajectory, as illustrated in Fig. 4. Specifically, we compare the point of intersection $(\tau_k, \hat{x}_j^{\tau_k})$ with minimal error and the point $(\tau_k, x_j^{\tau_k})$ on the ground truth trajectory, when the minimal error can not satisfy the threshold ϵ . The value of v_{j+k} is increased/decreased if the reconstructed point is at the rear/front position of the ground truth point, as shown in Fig. 4. The pseudo-code of the algorithm is illustrated in Algorithm 1.

5. Trajectory Reconstruction using Time-varying Shockwave Speed

Given the calibrated time-varying shockwave speed of the CV, we regard the CV as a leading vehicle to reconstruct the reference points for the NCVs which are the following vehicles of the CV. The reconstruction process is designed based on the method introduced in Section 4.2. However, simply connecting these reference points leads to a harsh trajectory, which can not fulfill the vehicle dynamics and thus can not be used to estimate energy consumption. Hence, we propose to design and apply a trajectory reconstruction method to filter the reference points and generate the final trajectory. In the rest of this section, we introduce the process of reconstructing the reference points and optimizing the reference points to reconstruct the final trajectory.

5.1. Reference Points Reconstruction

With the calibrated time-varying shockwave speed, it is simple to reconstruct the trajectory of CVs according to the process in Section 4.2. However, the trajectory reconstruction of other NCVs remains unsolved. To this end, we propose to reconstruct the reference points of these NCVs using the calibrated shockwave speed under the assumption that the shockwave speed does not change in a short time period.

Given the calibrated time-varying shockwave speed, which is denoted as $\{w_j^{\tau_k}, k \in [0, 1, \dots, k_{end}], j \in J\}$, the representation of the shockwave line in Eq. 6 is rewritten as Eq. 8. The reference points of CVs can be reconstructed according to the procedures in Section 4.2 using the calibrated shockwave speed, Eq. 5, and Eq. 8. Considering that the traffic conditions do not change dramatically in a short time period, we use the calibrated shockwave speed of the closest leading CV for the NCVs to reconstruct their reference points. Specifically, for j^{th} CV with calibrated time-varying shockwave speed, each shockwave speed value $w_j^{\tau_k}$ is paired with v_{j+k} the speed of k^{th} following vehicle, where the CV is considered as the leading vehicle. Thus, the reference points of k^{th} following vehicle of the CV can be reconstructed through calculating the points of intersection $\{\tilde{h}_{j+k}^{\tau_m}(\tau, w_j^{\tau_{k+m}}) \otimes g_{j+k}^{\tau_m}(\tau), m \in [0, 1, \dots, k_{end} - k]\}$, as shown in Fig. 5. The stochasticity of the impact of the leading vehicle on the following ones through shockwave propagation is reflected by sampling \tilde{v}_{j+k+m} from Gaussian distribution with mean v_{j+k+m} and pre-defined variance σ^2 to derive $g_{j+k}^{\tau_m}(\tau)$. After calculating the point of intersection $(\hat{x}_{j+k}^{\tau_m}, \tau_m^t)$, we need to validate it by comparing the position and time with the existing reference point $(\tau_{m'-1}, \hat{x}_{j+k}^{\tau_{m'-1}})$. This new reference point is added to reconstruct the trajectory if the position and time of this new point are greater than the existing point of the previous time step. The reconstruction process of reference points is presented in Algorithm 2.

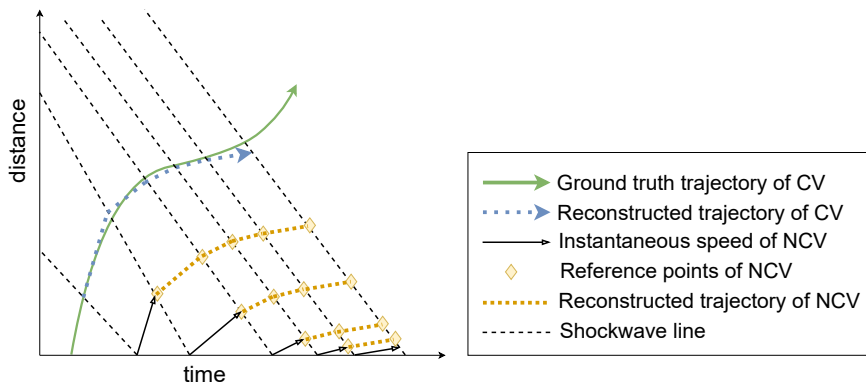


Figure 5: Trajectory reconstruction of other non-connected vehicles.

Algorithm 2 Reference Point Reconstruction for NCVs

Input: Arrival time $\{t_{j+k}, t_{j+k+1}, \dots, t_{j+k_{end}}\}$ and speed $\{v_{j+k}, v_{j+k+1}, \dots, v_{j+k_{end}}\}$ of the $(j+k)^{th}$ to $(j+k_{end})^{th}$ NCVs, the calibrated shockwave speed of j^{th} CV $\{w_j^{\tau_0}, w_j^{\tau_1}, \dots, w_j^{\tau_\kappa}, \dots, w_j^{\tau_{k_{end}}}\}$, j^{th} CV is the closest leading CV for its following 1^{st} to κ^{th} NCVs, the location of loop detector x_0 , standard deviation σ

Output: Reconstructed trajectory of $(j+k)^{th}$ NCV, where $k \leq \kappa$

```
1:  $\Pi_{j+k} = \{(t_{j+k}, x_0)\}$ 
2: for  $m = k : \kappa$  do
3:   Sample  $\tilde{v}_{j+m}$  from Gaussian distribution  $N(v_{j+m}, \sigma^2)$ 
4:   Derive  $\tilde{h}_{j+k}^{\tau_{m-k}}(\tau)$  using Eq. 8 and  $w_j^{\tau_m}$ 
5:   Derive  $g_{j+k}^{\tau_{m-k}}(\tau)$  using Eq. 5 and  $\tilde{v}_{j+m}$ 
6:   Calculate the point of intersection  $(\tau_{m-k}, \hat{x}_{j+k}^{\tau_{m-k}}) = \tilde{h}_{j+k}^{\tau_{m-k}}(\tau) \otimes g_{j+k}^{\tau_{m-k}}(\tau)$ 
7:   if  $\tau_{m-k} > \tau_{m-k-1}$  and  $\hat{x}_{j+k}^{\tau_{m-k}} > \hat{x}_{j+k}^{\tau_{m-k-1}}$  then
8:     Append  $(\tau_{m-k}, \hat{x}_{j+k}^{\tau_{m-k}})$  in  $\Pi_{j+k}$ 
9:   end if
10: end for
```

5.2. Trajectory Reconstruction

Although the above reference points can be connected directly to produce the final trajectory of the vehicle, the trajectory is not realistic since it does not consider vehicle or driver dynamics. Moreover, the energy consumption estimated using such a trajectory is not reliable either. To this end, we propose a further step to reconstruct the trajectory using the reference points. The reconstructed trajectory is constrained by the vehicle and driver dynamics for a more realistic trajectory and better energy estimation.

We apply a vehicle movement model to introduce vehicle and driver dynamics simultaneously. We choose the microsimulation free-flow acceleration model (MFC) (Makridis, Fontaras, Ciuffo and Mattas, 2019) as the vehicle movement model. MFC is developed to capture the vehicle acceleration dynamics accurately and consistently. The driver dynamics are also considered in MFC by introducing the driver style parameter. MFC takes the desired speed as input to simulate the acceleration behaviors given the driver style and vehicle's physical parameters, such as the mass of the vehicle, the maximum tractive force of the vehicle, etc. As such, we consider the speed of the reference points as the desired speed and input to the MFC to obtain the acceleration behaviors. Specifically, the desired speed is first enhanced by the interpolation method to increase the frequency. The interpolated desired speed is then input to the MFC to derive acceleration behaviors, which are further used to reconstruct the final trajectory and speed.

6. Experiments

To evaluate the performance of our proposed trajectory reconstruction method, we conduct the following experiments: (1) comparison with several SOTA baseline models in terms of trajectory reconstruction accuracy and energy consumption accuracy; (2) evaluation of the capability of capturing the traffic oscillation; (3) exploration the relationship between the accuracy of reconstructed trajectory and the penetration rate of CVs. The details are elaborated in the rest of this section.

6.1. Data Preprocessing

We use the HighD (Krajewski, Bock, Kloeker and Eckstein, 2018) and NGSIM (Federal Highway Administration, 2007, 2006) datasets to evaluate our method, which are collected on the highways of Germany and the US, respectively. They include the trajectory of vehicles of several minutes and in multiple lanes. We first separate each dataset into several sub-datasets according to lanes. Namely, the trajectory of vehicles on the same lane is formed as a new sub-dataset. Moreover, since our method is based on shockwave speed theory, we only use the sub-dataset that exists shockwave to evaluate the performance of our method. Only the car-following vehicles are kept without considering lane-change ones. The trajectories of some of the selected sub-dataset are shown in Fig. 6. We assume that a loop detector is deployed at the start position of the road segment, as shown in Fig. 1. We collect the velocity of all of the

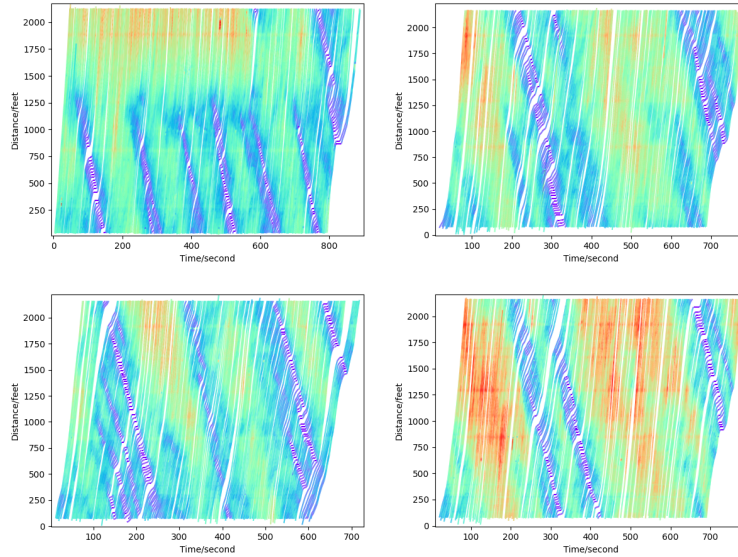


Figure 6: Examples of extracted trajectory data.

Metric \ Penetration rate		5%					
		Time headway (s)		Speed (m/s)		Fuel consumption (L/100km)	
Model		avg.	std.	avg.	std.	avg.	std.
Coifman		5.61	-	3.86	-	0.37	-
Macro-micro		3.82	0.27	2.86	0.19	0.32	0.06
Ours with polynomial		2.65	0.17	2.74	0.18	0.65	0.06
Ours with MFC		2.64	0.19	2.65	0.17	0.23	0.02
Metric \ Penetration rate		10%					
		Time headway (s)		Speed (m/s)		Fuel consumption (L/100km)	
Model		avg.	std.	avg.	std.	avg.	std.
Coifman		5.61	-	3.86	-	0.37	-
Macro-micro		2.27	0.13	2.70	0.06	0.28	0.03
Ours with polynomial		2.24	0.11	2.74	0.11	0.66	0.03
Ours with MFC		2.19	0.11	2.61	0.10	0.21	0.01
Metric \ Penetration rate		15%					
		Time headway (s)		Speed (m/s)		Fuel consumption (L/100km)	
Model		avg.	std.	avg.	std.	avg.	std.
Coifman		5.61	-	3.86	-	0.37	-
Macro-micro		1.72	0.12	2.15	0.07	0.24	0.02
Ours with polynomial		2.07	0.08	2.71	0.06	0.63	0.02
Ours with MFC		2.01	0.10	2.59	0.07	0.21	0.01

Table 1

Comparison of MAE of time headway, fuel consumption, and CO₂ among our model with different trajectory reconstruction methods and baseline models.

vehicles that pass the loop detector. Some of the vehicles are defined as CVs with known trajectories, which are the blue ones in Fig. 1.

6.2. Accuracy Comparison

To assess the performance of our method, we choose the following baseline models: (1) Coifman's method (Coifman, 2002) ; (2) macro-micro model (Chen et al., 2022a). We follow the values of parameters as mentioned in (Chen

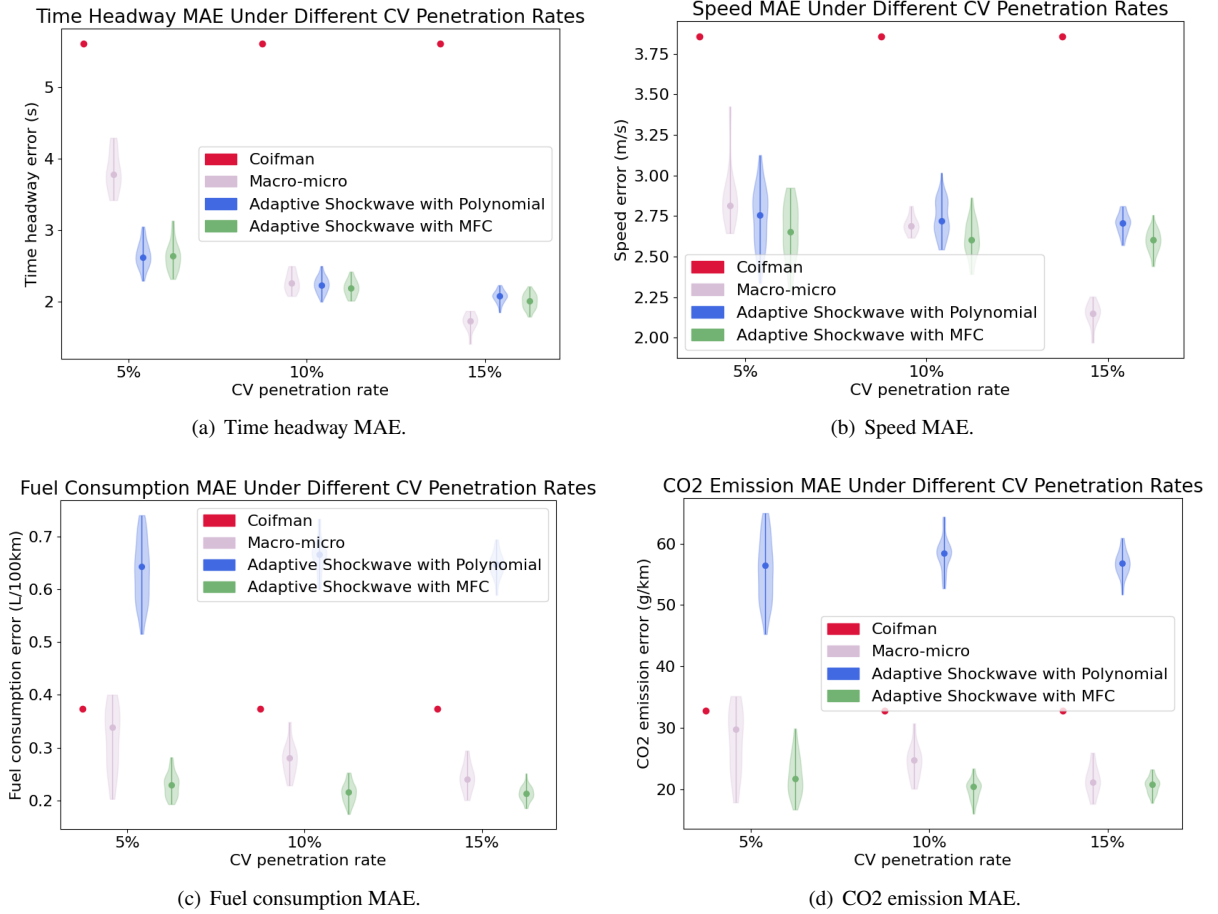


Figure 7: Accuracy comparison of time headway, fuel consumption, and CO₂ emission among baseline models and ours.

et al., 2022a) to reproduce the reconstruction results. Additionally, we also conduct an ablation study on the selection of trajectory reconstruction method. Besides the aforementioned trajectory reconstruction in 5.2, we apply a polynomial reconstruction method for comparison. Specifically, the reference points are reconstructed using a 5-degree polynomial optimized by least-square regression. We use mean absolute error (MAE) on time headway and speed to measure the accuracy of the reconstructed trajectory. To estimate the energy consumption and emission, we apply the simplified fuel consumption model (Makridis et al., 2019) and CO₂ emission model, which is publicly available¹ and is based on CO₂MPAS (European Commission, a) model, developed by the European Commission. We compare the total amount of energy consumed on the same travel length between our model and the ground truth trajectory.

Moreover, we set different penetration rates, 5%, 10%, and 15%, to explore its impact on the accuracy. To achieve 5% and 10% penetration rates, we randomly select one vehicle as the CV with a known trajectory every twenty or ten vehicles, respectively. For 15% penetration rate, we randomly select three vehicles as the CVs with known trajectory every twenty vehicles. For different penetration rates, we run 50 times and the CVs are selected randomly every time. As we have no information about the vehicle dynamics from the dataset, we randomly select an average vehicle model of European Class C (European Commission, b) to estimate fuel consumption and CO₂ emissions.

The comparison results of two baseline models and our model with different reconstruction methods are shown in Tab. 1 and violin plot as Fig. 7. **In terms of time headway accuracy**, ours with both polynomial and MFC reconstruction methods perform much better than the baseline models when the penetration rate is low, i.e., 5% in our experiment. When the penetration rate increases to 10%, our model still has better performance than the baseline models. Although

¹<https://pypi.org/project/co2mpas-driver/>

the macro-micro model achieves 14% higher performance than ours in average value while the penetration rate is 15%, our model has a more stable performance. *In terms of speed accuracy*, the performance is similar as the one of time headway. *In terms of fuel consumption and CO₂ emissions*, we only list the detailed value of fuel consumption since CO₂ has a linear relationship with fuel consumption. Our model with MFC reconstruction method achieves a higher accuracy than the baseline models under all of the penetration rates. Besides the average value, it also provides a more stable performance than others. Moreover, we analyze the reason for the higher error of the polynomial reconstruction method. The fuel consumption of the reconstructed trajectory is always lower than the one of the ground truth trajectory, which implies that the trajectory may be over-smooth to lose some randomness existing in the ground truth trajectory. The randomness may largely affect the performance in terms of energy consumption. In summary, our model has a better performance, especially when the CV penetration rate is low.

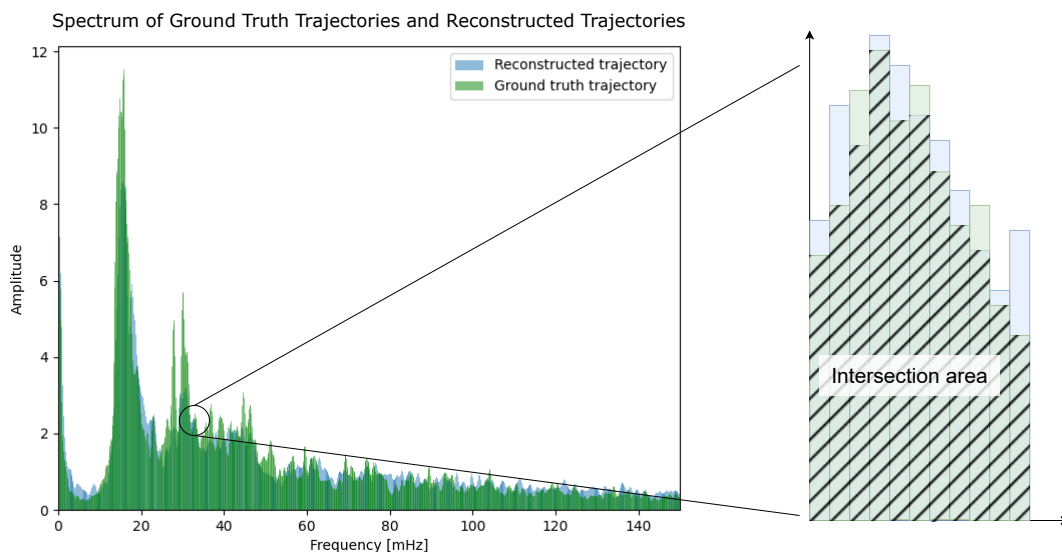


Figure 8: Evaluation metric for reproducing traffic oscillation.

6.3. Traffic Oscillation Reproduction

In this part, we evaluate the capability of the proposed methodology in reproducing observed traffic oscillations. To characterize the traffic oscillations, we use the frequency and amplitude of the velocity sequence since the velocity can well reflect empirical observations according to (Makridis et al., 2020). To this end, we apply the efficient Fast Fourier Transform (FFT) on the velocity sequence to get the frequency domain information. We use one of the subdatasets to illustrate the results. As shown in Fig. 9, Fig. 9.(b) is the spectrum of Fig. 9.(a), which is expressed using frequency and amplitude. Specifically, except for the dominant peak amplitude, the existence of other peak amplitudes implies multiple traffic oscillations (Makridis et al., 2020). We apply FFT to the reconstructed trajectory of Coifman’s method, macro-micro method, and ours with MFC reconstruction. The reconstructed trajectory and the corresponding Fourier spectrum are shown in Fig. 9.(c)-(h), respectively. Comparing Fig. 9.(b), (d), (f), and (h), our model has the most similar peak frequency and amplitude as the one of ground truth, which indicates that our reconstruction method can successfully reproduce the traffic oscillations.

Moreover, considering that the connected vehicles are selected randomly in the above evaluation experiments, the spectrum may be affected by the selection of connected vehicles. Thus, we introduce a new metric to quantify the capability of traffic oscillation reproduction. Specifically, we calculate the intersection area between the spectrum of ground trajectories and reconstructed trajectories. The method of calculating the intersection area is shown in Fig. 8. Then, we use the ratio of the intersection area on the whole area of the reconstructed spectrum to measure the quality of the reconstructed trajectories in terms of reproducing traffic oscillation. The ratio of Coifman’s method is **77.20%**. The mean and standard deviation value of the ratio achieved by macro-micro model is **63.90%** and **1.01%**, respectively. For

our model with MFC reconstruction method, we achieve **88.61%** and **1.41%**, in terms of mean and standard deviation value of the ratio. Thus, compared with other baseline models, our model is the most capable one to reproduce the traffic oscillations.

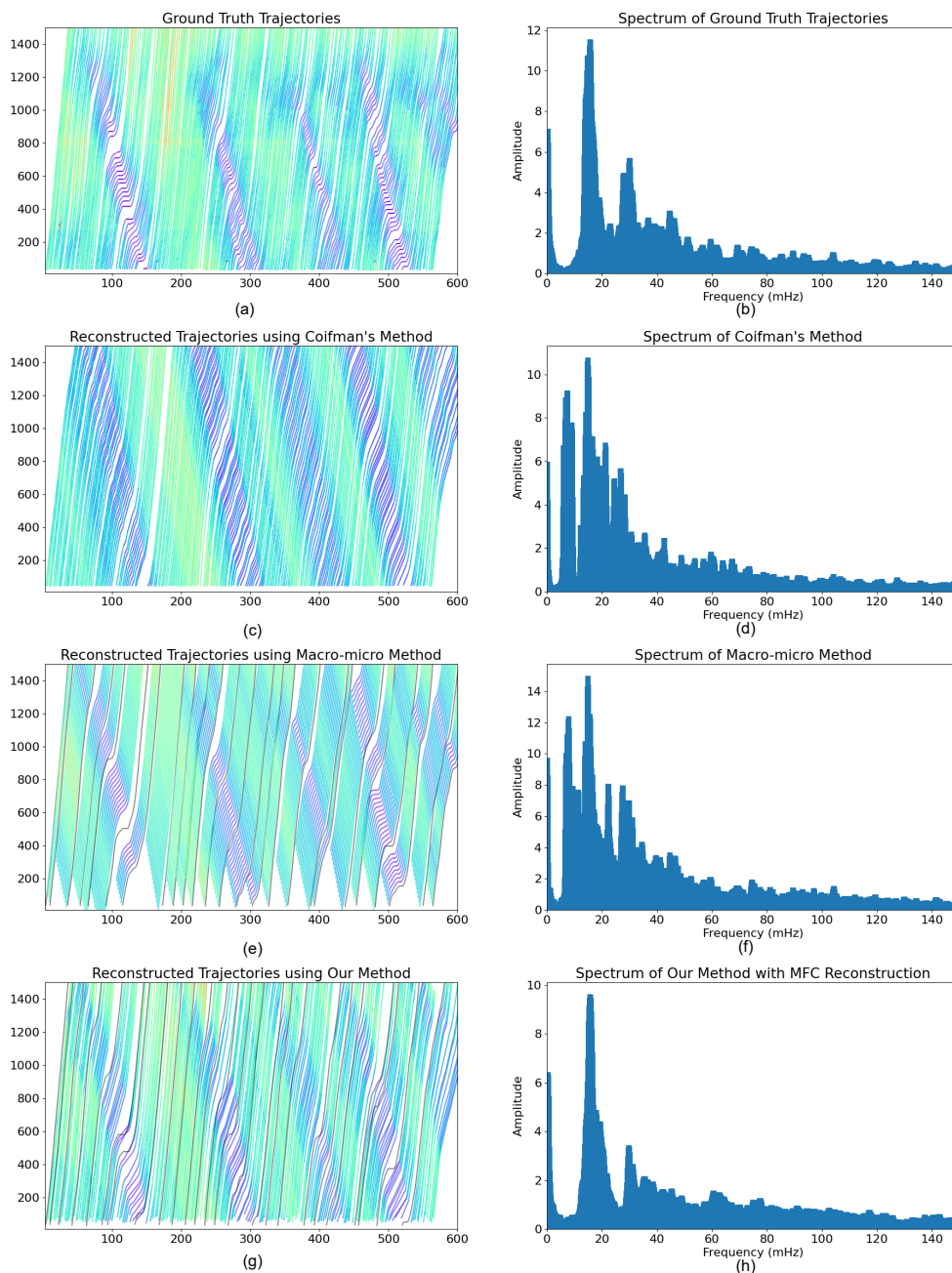


Figure 9: The Fourier Spectrum of ground truth data and reconstructed data. (a) Ground truth trajectories. (b) Fourier spectrum of ground truth. (c) Reconstructed trajectories by Coifman's method. (d) Fourier spectrum of reconstructed trajectories of Coifman's method. (e) Reconstructed trajectories by macro-micro method, where the CV penetration rate is 10% and the black line is the trajectory of CV. (f) Fourier spectrum of reconstructed trajectories of macro-micro method. (g) Reconstructed trajectories by our method with MFC reconstruction, where the CV penetration rate is 10% and the black line is the trajectory of CV. (h) Fourier spectrum of reconstructed trajectories of our method.

7. Discussion

As the above experimental results illustrate that the accuracy of the reconstructed trajectory is affected by the characteristics of the CVs' trajectory. To well measure the contribution of different CVs, we propose a new metric to evaluate the contribution of the CV to trajectory reconstruction, which is spatial-temporal area coverage. Further analysis is conducted to measure the amount of information required to calibrate different CVs. By evaluating these two proposed metrics, we can identify the most critical and informative CV.

7.1. Spatial-Temporal Area Reconstruction

Considering that the trajectory of the vehicles that are farther from the CV can not be fully recovered due to the limited number of calibrated shockwave speed values, we propose to reconstruct a spatial-temporal (ST) area instead of a set of complete trajectories. Given a set of calibrated shockwave speed values for a CV, the covered ST area of this CV is defined as the area enclosed by the reconstructed trajectory of the CV and the last shockwave line that can contribute to recovering the trajectory of other NCVs, as shown in Fig. 10(a). Note that the trajectory of an NCV is considered as reconstruction completion when the reconstructed segment is more than 80% of the whole journey. The larger of this area represents that knowing the trajectory of this CV can contribute more to trajectory reconstruction.

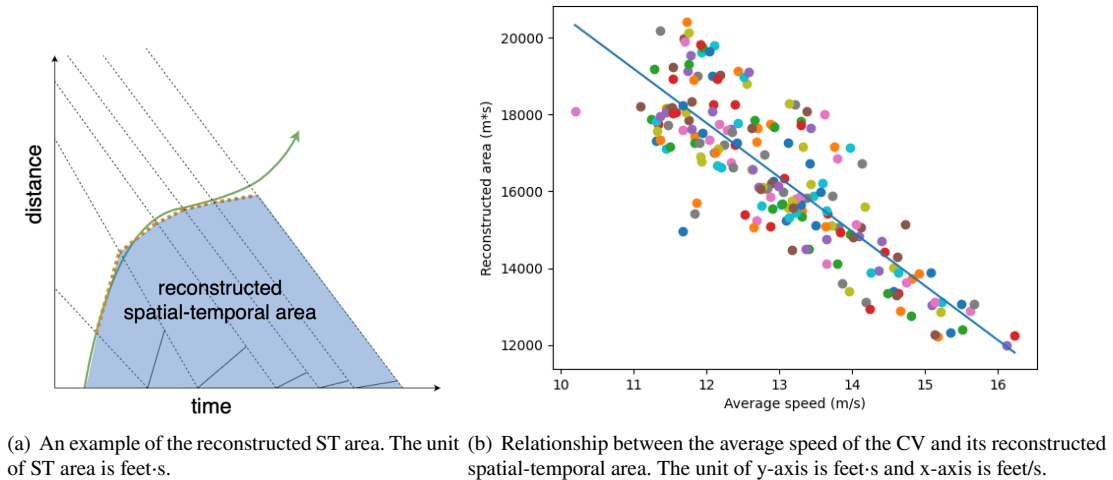


Figure 10: Analysis of the reconstructed spatial-temporal area of CVs.

Given the above definition, we explore the relationship between the average speed of the CV and its covered ST area. We calibrate the time-varying shockwave speed for all vehicles. The results are shown in Fig. 10(b), which indicates that the CV with a smaller average speed can cover more ST area and contribute more to reconstructing the trajectory of its following vehicles.

7.2. Information Amount Requirement for Calibration

As introduced in the previous sections and experiments, the collected data, including arrival time and speed of following vehicles, is required to calibrate the time-vary shockwave speed values of the leading CV. However, for different CVs, they need different amount of data for calibration. We define the number of following vehicles' speed that are used to calibrate the shockwave speed of the CV as the information amount required for calibrating the time-varying shockwave speed. As shown in Fig. 11, there is no significant relationship between the average speed of the CV and its required amount of information. Thus, the selection of a CV has little impact on the efficiency and effectiveness in terms of the data requirement.

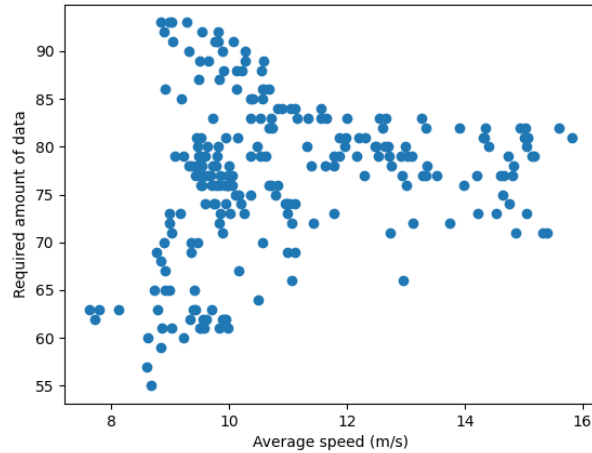


Figure 11: Relationship between the average speed of CVs and the required amount of data for shockwave calibration.

8. Conclusion and Future Work

In conclusion, our proposed framework presents a comprehensive solution to the challenges in reconstructing vehicle trajectories using a combination of fixed and connected vehicle sensor data. The integration of fixed sensor information and connected vehicle data enables a more accurate and robust trajectory reconstruction process. The introduced calibration algorithm for time-varying shockwave speed further enhances precision, addressing the limitations of existing methods. Moreover, the vehicle dynamics are considered while reconstructing the trajectory for better energy estimation. Through real-world evaluations, our method consistently outperforms baseline models, demonstrating its effectiveness across various traffic conditions, especially in scenarios with a low penetration rate of CVs. The concept of a reconstructed ST area and information amount requirement provide valuable insights into the contribution and effectiveness of CVs with different average speeds, enriching our understanding of the reconstruction process. This work not only advances the field of trajectory reconstruction but also lays the foundation for more accurate modeling and optimization in traffic flow management and related domains.

In future work, we plan to apply our method to traffic scenarios with signals and intersections to measure the queue length of intersections. Moreover, we will explore and build the relationship between the time-varying shockwave speed and traffic data collected by a loop detector using deep learning methods. As such, we can reconstruct the trajectories without knowing any CV trajectory, which can significantly enlarge the application scenarios and mitigate the reliance on the existence of CVs.

References

- Anuar, K., Cetin, M., 2017. Estimating freeway traffic volume using shockwaves and probe vehicle trajectory data. *Transportation Research Procedia* 22, 183–192. URL: <https://www.sciencedirect.com/science/article/pii/S2352146517301606>, doi:<https://doi.org/10.1016/j.trpro.2017.03.025>. "19th EURO Working Group on Transportation Meeting, EWGT2016, 5-7 September 2016, Istanbul, Turkey".
- Apostolakis, T., Makridis, M.A., Kouvelas, A., Ampountolas, K., 2023. Energy-based assessment and driving behavior of acc systems and humans inside platoons. *IEEE Transactions on Intelligent Transportation Systems* 24, 12726–12735. doi:10.1109/TITS.2023.3285296.
- Chen, X., Yin, J., Qin, G., Tang, K., Wang, Y., Sun, J., 2022a. Integrated macro-micro modelling for individual vehicle trajectory reconstruction using fixed and mobile sensor data. *Transportation Research Part C: Emerging Technologies* 145, 103929.
- Chen, X., Yin, J., Tang, K., Tian, Y., Sun, J., 2022b. Vehicle trajectory reconstruction at signalized intersections under connected and automated vehicle environment. *IEEE Transactions on Intelligent Transportation Systems* 23, 17986–18000. doi:10.1109/TITS.2022.3150577.
- Ciuffo, B., Makridis, M., Toledo, T., Fontaras, G., 2018. Capability of current car-following models to reproduce vehicle free-flow acceleration dynamics. *IEEE Transactions on Intelligent Transportation Systems* 19, 3594–3603. doi:10.1109/TITS.2018.2866271.
- Coifman, B., 2002. Estimating travel times and vehicle trajectories on freeways using dual loop detectors. *Transportation Research Part A: Policy and Practice* 36, 351–364. URL: <https://www.sciencedirect.com/science/article/pii/S0965856401000076>, doi:[https://doi.org/10.1016/S0965-8564\(01\)00007-6](https://doi.org/10.1016/S0965-8564(01)00007-6).

- Daganzo, C.F., 1997. Fundamentals of transportation and traffic operations. Emerald Group Publishing Limited.
- Daganzo, C.F., 2005a. A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B: Methodological* 39, 187–196. URL: <https://www.sciencedirect.com/science/article/pii/S0191261504000487>, doi:<https://doi.org/10.1016/j.trb.2004.04.003>.
- Daganzo, C.F., 2005b. A variational formulation of kinematic waves: Solution methods. *Transportation Research Part B: Methodological* 39, 934–950. URL: <https://www.sciencedirect.com/science/article/pii/S0191261505000068>, doi:<https://doi.org/10.1016/j.trb.2004.05.003>.
- Deng, Y., Cao, Q., Ren, G., Ma, J., Zhu, S., 2023. Vehicle trajectory reconstruction incorporating probe and fixed sensor data. *Journal of Transportation Engineering, Part A: Systems* 149, 04023088. URL: <https://ascelibrary.org/doi/abs/10.1061/JTEPBS.TEENG-7788>, doi:10.1061/JTEPBS.TEENG-7788, arXiv:<https://ascelibrary.org/doi/pdf/10.1061/JTEPBS.TEENG-7788>.
- European Commission, a. CO₂MPAS Vehicle simulator predicting NEDC CO₂ emissions from WLTP. <https://co2mpas.readthedocs.io/en/stable/>.
- European Commission, b. EU classification of vehicle types. <https://alternative-fuels-observatory.ec.europa.eu/general-information/vehicle-types>.
- Federal Highway Administration, 2006. US highway 80 dataset.
- Federal Highway Administration, 2007. US highway 101 dataset.
- Fiori, C., Arcidiacono, V., Fontaras, G., Makridis, M., Mattas, K., Marzano, V., Thiel, C., Ciuffo, B., 2019. The effect of electrified mobility on the relationship between traffic conditions and energy consumption. *Transportation Research Part D: Transport and Environment* 67, 275–290. URL: <https://www.sciencedirect.com/science/article/pii/S136192091830662X>, doi:<https://doi.org/10.1016/j.trd.2018.11.018>.
- Guo, Q., Li, L., (Jeff) Ban, X., 2019. Urban traffic signal control with connected and automated vehicles: A survey. *Transportation Research Part C: Emerging Technologies* 101, 313–334. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X18311641>, doi:<https://doi.org/10.1016/j.trc.2019.01.026>.
- Hao, P., Boriboonsomsin, K., Wu, G., Barth, M., 2014. Probabilistic model for estimating vehicle trajectories using sparse mobile sensor data, in: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1363–1368. doi:10.1109/ITSC.2014.6957877.
- He, Y., Makridis, M., Fontaras, G., Mattas, K., Xu, H., Ciuffo, B., 2020. The energy impact of adaptive cruise control in real-world highway multiple-car-following scenarios. *European Transport Research Review* 12, 1–11.
- He, Y., Montanino, M., Mattas, K., Punzo, V., Ciuffo, B., 2022. Physics-augmented models to simulate commercial adaptive cruise control (acc) systems. *Transportation Research Part C: Emerging Technologies* 139, 103692. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X22001334>, doi:<https://doi.org/10.1016/j.trc.2022.103692>.
- Kachroo, P., Ozbay, K., Hobeika, A., 2001. Real-time travel time estimation using macroscopic traffic flow models, in: ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat.No.01TH8585), pp. 132–137. doi:10.1109/ITSC.2001.948643.
- 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), pp. 2118–2125. doi:10.1109/ITSC.2018.8569552.
- Lei, F., Wang, Y., Lu, G., Sun, J., 2014. A travel time reliability model of urban expressways with varying levels of service. *Transportation Research Part C: Emerging Technologies* 48, 453–467. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X14002897>, doi:<https://doi.org/10.1016/j.trc.2014.09.019>.
- Li, L., Jiang, R., He, Z., Chen, X.M., Zhou, X., 2020. Trajectory data-based traffic flow studies: A revisit. *Transportation Research Part C: Emerging Technologies* 114, 225–240. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X19312987>, doi:<https://doi.org/10.1016/j.trc.2020.02.016>.
- Li, T., Chen, D., Zhou, H., Xie, Y., Laval, J., 2022. Fundamental diagrams of commercial adaptive cruise control: Worldwide experimental evidence. *Transportation Research Part C: Emerging Technologies* 134, 103458. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X21004460>, doi:<https://doi.org/10.1016/j.trc.2021.103458>.
- Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves ii. a theory of traffic flow on long crowded roads. *Proceedings of the royal society of london. series a. mathematical and physical sciences* 229, 317–345.
- van Lint, J.W.C., van der Zijpp, N.J., 2003. Improving a travel-time estimation algorithm by using dual loop detectors. *Transportation Research Record* 1855, 41–48. URL: <https://doi.org/10.3141/1855-05>, doi:10.3141/1855-05, arXiv:<https://doi.org/10.3141/1855-05>.
- Lint, J.W.C.H.V., 2010. Empirical evaluation of new robust travel time estimation algorithms. *Transportation Research Record* 2160, 50–59. URL: <https://doi.org/10.3141/2160-06>, doi:10.3141/2160-06, arXiv:<https://doi.org/10.3141/2160-06>.
- Liu, J., Mao, X., Fang, Y., Zhu, D., Meng, M.Q.H., 2021. A survey on deep-learning approaches for vehicle trajectory prediction in autonomous driving, in: 2021 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 978–985. doi:10.1109/ROBIO54168.2021.9739407.
- Makridis, M., Fontaras, G., Ciuffo, B., Mattas, K., 2019. Mfc free-flow model: Introducing vehicle dynamics in microsimulation. *Transportation Research Record* 2673, 762–777. URL: <https://doi.org/10.1177/0361198119838515>, doi:10.1177/0361198119838515, arXiv:<https://doi.org/10.1177/0361198119838515>.
- Makridis, M., Leclercq, L., Ciuffo, B., Fontaras, G., Mattas, K., 2020. Formalizing the heterogeneity of the vehicle-driver system to reproduce traffic oscillations. *Transportation Research Part C: Emerging Technologies* 120, 102803. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X20307129>, doi:<https://doi.org/10.1016/j.trc.2020.102803>.
- Makridis, M.A., Anesiadou, A., Mattas, K., Fontaras, G., Ciuffo, B., 2023. Characterising driver heterogeneity within stochastic traffic simulation. *Transportmetrica B: Transport Dynamics* 11, 725–743.

- Makridis, M.A., Kouvelas, A., 2023. Adaptive physics-informed trajectory reconstruction exploiting driver behavior and car dynamics. *Scientific Reports* 13, 1121. URL: <https://doi.org/10.1038/s41598-023-28202-1>, doi:10.1038/s41598-023-28202-1.
- Mehran, B., Kuwahara, M., Naznin, F., 2012. Implementing kinematic wave theory to reconstruct vehicle trajectories from fixed and probe sensor data. *Transportation research part C: emerging technologies* 20, 144–163.
- Montanino, M., Punzo, V., 2015. Trajectory data reconstruction and simulation-based validation against macroscopic traffic patterns. *Transportation Research Part B: Methodological* 80, 82–106. URL: <https://www.sciencedirect.com/science/article/pii/S0191261515001393>, doi:<https://doi.org/10.1016/j.trb.2015.06.010>.
- Newell, G., 1993. A simplified theory of kinematic waves in highway traffic, part i: General theory. *Transportation Research Part B: Methodological* 27, 281–287. URL: <https://www.sciencedirect.com/science/article/pii/S019126159390038C>, doi:[https://doi.org/10.1016/0191-2615\(93\)90038-C](https://doi.org/10.1016/0191-2615(93)90038-C).
- Rey, A., Jin, W.L., Ritchie, S.G., 2019. An extension of newell’s simplified kinematic wave model to account for first-in-first-out violation: With an application to vehicle trajectory estimation. *Transportation Research Part C: Emerging Technologies* 109, 79–94. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X18310209>, doi:<https://doi.org/10.1016/j.trc.2019.10.005>.
- Sakhare, R.S., Li, H., Bullock, D.M., 2023. Methodology for the identification of shock wave type and speed in a traffic stream using connected vehicle data. *Future Transportation* 3, 1147–1174. URL: <https://www.mdpi.com/2673-7590/3/4/63>, doi:10.3390/futuretransp3040063.
- Seo, T., Kawasaki, Y., Kusakabe, T., Asakura, Y., 2019. Fundamental diagram estimation by using trajectories of probe vehicles. *Transportation Research Part B: Methodological* 122, 40–56. URL: <https://www.sciencedirect.com/science/article/pii/S0191261518303527>, doi:<https://doi.org/10.1016/j.trb.2019.02.005>.
- Suarez, J., Makridis, M., Anesiadou, A., Komnos, D., Ciuffo, B., Fontaras, G., 2022. Benchmarking the driver acceleration impact on vehicle energy consumption and co2 emissions. *Transportation Research Part D: Transport and Environment* 107, 103282. URL: <https://www.sciencedirect.com/science/article/pii/S1361920922001109>, doi:<https://doi.org/10.1016/j.trd.2022.103282>.
- Sun, C., Wu, C., Chu, D., Xie, L., Liu, L., Li, H., . Vehicle Trajectory Restoration Based on Vondrak Filtering and Cubic Spline Interpolation. pp. 235–248. URL: <https://ascelibrary.org/doi/abs/10.1061/9780784479896.023>, doi:10.1061/9780784479896.023, arXiv:<https://ascelibrary.org/doi/pdf/10.1061/9780784479896.023>.
- Wan, N., Vahidi, A., Luckow, A., 2016. Reconstructing maximum likelihood trajectory of probe vehicles between sparse updates. *Transportation Research Part C: Emerging Technologies* 65, 16–30. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X16000280>, doi:<https://doi.org/10.1016/j.trc.2016.01.010>.
- Wang, M., Daamen, W., Hoogendoorn, S., van Arem, B., 2011. Estimating acceleration, fuel consumption, and emissions from macroscopic traffic flow data. *Transportation Research Record* 2260, 123–132. URL: <https://doi.org/10.3141/2260-14>, doi:10.3141/2260-14, arXiv:<https://doi.org/10.3141/2260-14>.
- Wang, Y., Wei, L., Chen, P., 2020. Trajectory reconstruction for freeway traffic mixed with human-driven vehicles and connected and automated vehicles. *Transportation Research Part C: Emerging Technologies* 111, 135–155. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X19307818>, doi:<https://doi.org/10.1016/j.trc.2019.12.002>.
- Wei, L., Wang, Y., Chen, P., 2021. A particle filter-based approach for vehicle trajectory reconstruction using sparse probe data. *IEEE Transactions on Intelligent Transportation Systems* 22, 2878–2890. doi:10.1109/TITS.2020.2976671.
- Yao, Z., Liu, M., Jiang, Y., Tang, Y., Ran, B., . Trajectory reconstruction for mixed traffic flow with regular, connected, and connected automated vehicles on freeway. *IET Intelligent Transport Systems* n/a. URL: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/itr2.12294>, doi:<https://doi.org/10.1049/itr2.12294>, arXiv:<https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/itr2.12294>.
- Zhang, Y., Chen, X., Wang, J., Zheng, Z., Wu, K., 2022. A generative car-following model conditioned on driving styles. *Transportation Research Part C: Emerging Technologies* 145, 103926. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X22003394>, doi:<https://doi.org/10.1016/j.trc.2022.103926>.
- Zheng, S.T., Makridis, M.A., Kouvelas, A., Jiang, R., Jia, B., 2023. A multi-objective calibration framework for capturing the behavioral patterns of autonomously-driven vehicles. *Transportation Research Part C: Emerging Technologies* 152, 104151. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X23001407>, doi:<https://doi.org/10.1016/j.trc.2023.104151>.