# AUTONOMOUS VEHICLES AS SENSORS: TRAFFIC STATE ESTIMATION

**Yunfei Zhang, Corresponding Author**
Chair of Traffic Engineering and Control, Technical University of Munich (TUM)
Arcisstrasse 21, 80333 Munich, Germany
yunfei.zhang@tum.de
ORCID:0000-0003-1902-1816

**Jeremias Gerner**
AImotion Bavaria, Technische Hochschule Ingolstadt
Ingolstadt, Germany
jeremias.gerner@thi.de
ORCID:0000-0002-1611-5086

**Mario Ilic**
Chair of Traffic Engineering and Control, Technical University of Munich (TUM)
Arcisstrasse 21, 80333 Munich, Germany
mario.ilic@tum.de
ORCID: 0000-0003-2457-698X

**Stefanie Schmidtner**
AImotion Bavaria, Technische Hochschule Ingolstadt
Ingolstadt, Germany
stefanie.schmidtner@thi.de
ORCID:0000-0003-2300-2159

**Klaus Bogenberger**
Chair of Traffic Engineering and Control, Technical University of Munich (TUM)
Arcisstrasse 21, 80333 Munich, Germany
klaus.bogenberger@tum.de
ORCID:0000-0003-3868-9571

Word Count: 5284 words + 3 table(s) $\times$ 250 = 6034 words

Submission Date: August 1, 2024

**ABSTRACT**

Integrating autonomous vehicles (AVs) as sensors presents a novel approach to traffic state estimation, leveraging both moving and parked cars to enhance data collection and analysis. This paper explores the effectiveness of using AVs as traffic sensors, comparing the performance of a Weighted Spatio-Temporal Estimation (WSTE) method against two baseline methods with and without complete historical data. We employ a camera-based detection system to identify participants in a microscopic traffic simulation, explicitly model the parked vehicles, and gather data from moving and parked observers at varying penetration rates. By comparing the estimation accuracy of link-level densities from observers and lanes, we demonstrate that moving and parked vehicles can complement each other using the new camera-based detection method, especially with different penetration rates. Our Weighted Spatio-Temporal Estimation method has also been proven to reach high accuracy, with even one observation from the historical data.

*Keywords*: Autonomous Vehicles as Sensors (AVaS), Autonomous Vehicles, Traffic State Estimation, Detection, Extended Floating Car Data (xFCD)

# INTRODUCTION

## Background and Motivation

Traffic state estimation plays a critical role in traffic engineering. Precise estimation and prediction can significantly benefit traffic control measures and long-term planning. Accurate traffic state estimation is fundamental for improving road safety, reducing congestion, and optimizing transportation infrastructure. In recent years, research methodologies have expanded to include both model-based and data-driven approaches, mainly due to the huge number of available data sets.

Collecting traffic data forms the foundation for practical traffic state estimation. Traditionally, traffic data collection methods, such as manual counting and loop detectors, have been labor-intensive and cost-prohibitive. However, the rise of autonomous vehicles (AVs) and intelligent transportation systems (ITS) has introduced novel methodologies leveraging sensor data from diverse sources. AVs, equipped with advanced sensor technologies such as LiDARs, radars, and cameras, can collect variant traffic data. These vehicles typically serve as moving observers, persistently acquiring real-time traffic data while circulating in complex urban road networks.

Motivated by the fact that the data from AVs are already there, we propose an innovative approach known as Autonomous Vehicles as Sensors (AVaS) (*1*), given AVs' inherent data collection capabilities. This approach utilizes AVs in both motion and parked states to compile valuable data for traffic state estimation, unlike traditional Floating Car Data (FCD) or Extended Floating Car Data (xFCD), which relies solely on vehicles in motion and provides information only about the ego-vehicle, AVaS benefits from both moving and parked observers within the network. Additionally, on-board vehicle sensors offer supplementary information about other road users. Since AVs can be centrally regulated, they can be strategically assigned to specific locations to gather data, mainly when not privately owned, such as Autonomous Mobility on Demand (AMoD) fleets.

By applying the AVaS idea, traffic state estimation can be expanded to include microscopic and macroscopic scales. On the microscopic scale, AVs can estimate traffic states such as densities and link speeds for individual road segments. Concurrently, these data can be aggregated for macroscopic traffic parameters like average speed, volume, and density at the cluster or network level.

This study aims to explore the potential of leveraging AVs as sensors for link-level traffic state estimation, specifically for density estimation. By simulating the data collection process using real sensor attributes (*2*) and applying the AVaS concept to a real-world traffic network in Ingolstadt, Germany, we seek to estimate and predict traffic states using both moving and parked vehicles.

## Literature Review

Traffic state estimation (TSE) methods highly depend on the available data types. While traditional TSE methods rely on stationary detector data, more modern approaches use probe vehicle (PV) data of different types. Table 1 gives an overview of the use of different sensor types for TSE.

Using traffic data from moving vehicles for traffic engineering and control originated in 1954 when Wardrop and Charlesworth (13) introduced the moving observer (MO) method. This technique involved estimating speed and traffic flow through manual observations of surrounding traffic in both directions, including counting the number of vehicles overtaking or being overtaken and oncoming vehicles. With advancements in vehicle sensor technology over the past decades, onboard sensors in Connected and Autonomous vehicles (CAVs) can now acquire data on surrounding traffic conditions, providing extended Floating Car Data (xFCD) for use in ITS applica-

| | Stationary Detectors | Probe Vehicles (PVs) | | |
|---|---|---|---|---|
| | | **Conventional PVs** | **PVs with spacing measurement** | **Automated Vehicles** |
| **Installed Sensors** | Loop detectors, cameras | Global Positioning System (GPS) | GPS & Long Range Radar (LRR) | GPS, LRR, LiDAR, camera |
| **Collected Raw Data** | Vehicle counts over time | PV's trajectory | PV's & proceeding vehicle's trajectory | GPS, LRR, LiDAR, camera |
| **TSE possible** | yes | no | yes | yes |
| **Literature** | (*3, 4*) | (*5–7*) | (*8–11*) | (*1, 12*) |

**TABLE 1**: Overview of the usage of different sensor types for traffic state estimation (TSE)

tions. Table 2 summarizes and classifies simulation studies conducted in the last years on the MO method.

Previous work has expanded the MO method into various aspects. From the perspective of transportation modes, Czogalla and Naumann (15) and Kühnel et al. (17) examined the impact of different types of moving observers on traffic flow estimation, considering both passenger cars and public transit. When considering the calculated traffic parameters, most reviewed literature primarily focuses on determining instantaneous link-based fundamental traffic flow variables, such as traffic volume, speed, and density. Czogalla and Naumann (15) and Schäfer and Hoyer (19), Schäfer (28) also explored the calculation of link-based travel times from recorded traffic parameters. Additionally, Langer et al. (24) derived queue lengths at intersections from relative flow data collected from oncoming traffic flows.

Regarding estimation methods, Florin and Olariu (18) introduced a data-driven approach using machine learning to predict traffic flow characteristics based on continuous data streams from mobile sensors. Guerrieri et al. (22) explored data fusion techniques to enhance the accuracy and completeness of traffic data through multiple moving observers. Furthermore, Ma and Qian (26) integrated machine learning techniques with moving observers to improve prediction accuracy.

Overall, past studies focus on using moving observers for traffic data collection. These studies have offered valuable insights into the potential applications and limitations of MO data but have also revealed specific gaps: 1) Current literature emphasizes data collection during motion, with limited exploration of stationary scenarios; 2) The estimation is based on the processed data, i.e., the data have been collected and are ready for estimation, and 3) Estimation for MFD is generally biased, especially for moving observers. To address these research gaps, we introduced our approach, Autonomous Vehicles as Sensors (AVaS) (*1, 2*). In (*1*), we first introduced this concept and verified the idea in a grid network using the Macroscopic Fundamental Diagram (MFD).

| Source | Year | Observer Vehicle: Passenger Car | Observer Vehicle: Public Transportation | Direction: Same-Directional | Direction: Oncoming Traffic | Recorded Traffic Parameters: Relative Flow (same-directional) | Recorded Traffic Parameters: Relative Flow (oncoming traffic) | Recorded Traffic Parameters: Link Travel Time | Recorded Traffic Parameters: Avg. Speed (of detected vehicles) | Calculated Traffic Parameters: Traffic Volume | Calculated Traffic Parameters: Speed | Calculated Traffic Parameters: Traffic Density | Calculated Traffic Parameters: Travel Time | Calculated Traffic Parameters: Queue Lengths at Intersections |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mulligan and Nicholson (14) | 2002 | x | | | x | | x | x | x | x | | | | |
| Czogalla and Naumann (15) | 2007 | | x | x | x | x | x | | x | | | | x | |
| Wolf et al. (16) | 2008 | | x | | x | | x | | | x | | | | |
| Kühnel et al. (17) | 2009 | x | x | | x | | x | x | x | x | | | | |
| Florin and Olariu (18) | 2016 | x | | x | | x | | x | | x | x | x | | |
| Schäfer and Hoyer (19) | 2017 | x | | x | x | x | x | | | x | x | | x | |
| Schäfer and Hoyer (*20*) | 2017 | x | | x | x | x | x | | | x | x | | x | |
| van Erp et al. (21) | 2018 | x | | x | | x | | | | x | | x | | |
| Guerrieri et al. (22) | 2019 | x | | x | x | x | x | x | | x | x | x | | |
| van Erp et al. (23) | 2019 | x | | x | | x | | | | x | | x | | |
| Langer et al. (24) | 2020 | x | | | x | | x | | | | | | | x |
| van Erp et al. (25) | 2020 | x | | x | x | x | | | | x | | x | | |
| Ma and Qian (26) | 2021 | x | | x | x | x | x | x | | x | x | x | | |
| Florin and Olariu (27) | 2023 | x | | x | x | x | x | | | | | x | | |
| Zhang et al. (1) | 2023 | x | | x | x | x | x | | x | x | x | x | | |

**TABLE 2**: Literature overview of simulation studies on the moving observer method

Based on this result, we extended the concept in (*2*) by simulating the data collection process with sensor attributes and multi-level estimation in a calibrated city-level simulation. However, the limitations include: 1) occlusion is not included, and 2) the point estimation is not accurate enough. In this paper, we try to work on these two limitations further.

### Contributions and Research Questions

To tackle the current research gaps, we utilize an emulation of a camera-based detection within the microscopic simulation, model the parked vehicles as observers, and improve our local traffic density estimation methods based on spatio-temporal correlations. Therefore, our contributions can be summarized as the following three:

1. Implementation of camera-based detection: we apply an emulation of a camera-based detection, a neural network architecture trained on a computer vision-based evaluation if vehicles are detectable by the observer based on (*12*). Compared to distance-based detection methods we used in the previous paper (*1, 2*), it is more realistic as it includes occlusion.
2. Modeling parked vehicles as sensors: We model the parked observers by adding parking lots based on real-world demand and use the parked vehicles to detect the surrounding traffic.
3. Estimation density based on spatio-temporal correlations: instead of focusing on the road section where our observers are located, we utilize all vehicles within a defined area and estimate the link-level density based on the weights from both spatial and temporal perspectives. The results are also compared among different penetration rates.

To achieve these three contributions, we develop our research questions accordingly:

1. How accurately can the camera-based detection method detect surrounding traffic compared to distance-based methods?
2. How much information can parked observers provide to estimate the traffic?
3. How does our Weighted Spatio-Temporal Estimation (WSTE) method compare to baseline methods?

## METHODOLOGY

The methodology can be summarized in Figure 1. With a well-calibrated SUMO simulation, we first simulate both moving and parked observers with different penetration rates. Second, we use two data collection methods, distance- and Neural Network (NN)-based detection methods. With the outputs from moving and parked observers, we applied different estimation methods for point calculation and area estimation. Point estimation targets for the raw data collected from the observers. A ‘detected’ dictionary will be returned at each timestamp for an individual observer. Point estimation refers to estimating the local density based on these data. Area estimation is based on the edges/lanes. Because an edge/lane does not always have an observer, we use an area surrounding the edge/lane to estimate its density. All the observers inside this area will be considered to estimate the densities. Note that the highlighted modules and parameters in the flowchart represent a sensitivity analysis. Therefore, we will compare between:

1. Moving observers vs. parked observers
2. Different penetration rates
3. Distance- vs. NN-based detection
4. Different area estimation methods

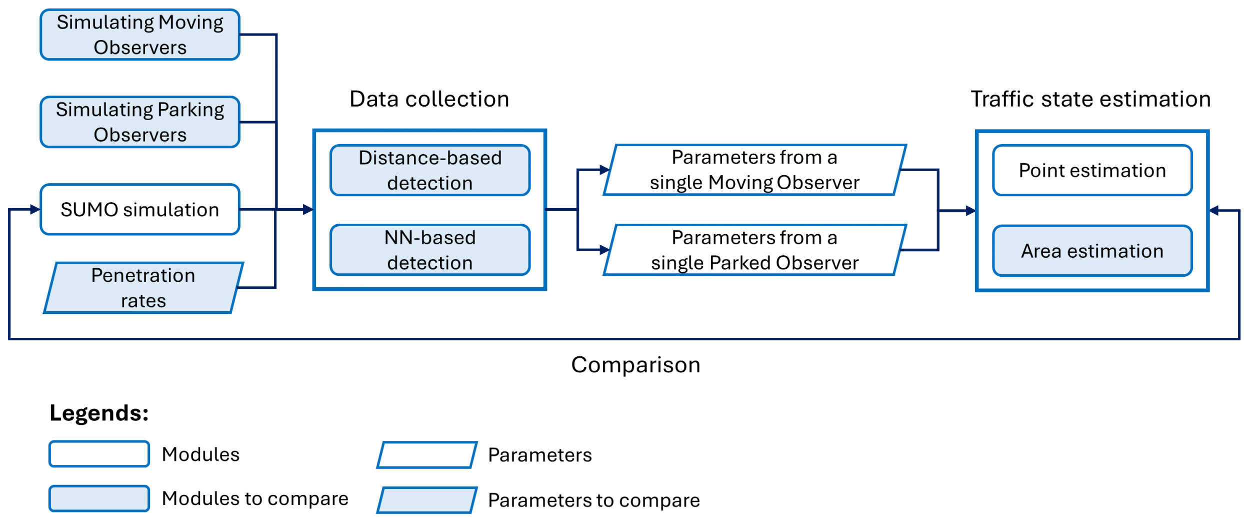



**FIGURE 1**: Flowchart of the methodology

### Simulating Moving and Parked Observers

In our research, we detect other traffic participants from both Moving Observers (MO) and Parked Observers (PO). For the MO, we classify each spawned vehicle as an observer or a non-observer based on the investigated penetration rate. The vehicle routes remain unchanged and adhere to the traffic demand specifications defined in (*29*). Since the traffic network of Ingolstadt, created in (*29*), does not include parked vehicles, we adapt the network to enable roadside parking on all edges, as illustrated in Figure 3. We define a parking density, representing the proportion of parking spots randomly occupied by vehicles. During the simulation, parked vehicles do not affect the traffic demand of moving vehicles.

### Detection and Data Collection

To accurately estimate the capabilities of AVaS for traffic state estimation in a microscopic simulation, it is essential to identify which vehicles can be recognized by the corresponding moving and parked observers, as would be the case in a real-world scenario with corresponding sensors and 3D object detection algorithms. As the first detection method, we use a distance-based method as in the previous work (*2*). The Distance-based detection computes the Euclidean distance between the ego and nearby vehicles. A vehicle is added to the detected list if the distance lies within specific thresholds (detection ranges) $d_l^{f/b}$ for lane $l$ in both forward $f$ and backward $b$ directions. In this paper, we select three different distances $[25, 50, 100]$ meters for $d_l^{f/b}$. This time, we do not differentiate vehicles from different lanes or even the opposite direction instead of detecting vehicles in the same lane. This simplification significantly accelerates the simulation and calculation.
However, This distance-based detection overestimates real-world sensors' detection capabilities, as analyzed in (*30*). We apply an alternative approach introduced in (*12*) to address this. This methodology approximates camera-based detection of other traffic participants, using the KITTI (*31*) thresholds as a reference. We at this moment utilize the NN approach presented in the paper for fast inference. To implement this approach, we first generate a dataset, which is subsequently used to train the neural networks. Following the methodology outlined in (*12*), the dataset is created by converting the two-dimensional SUMO simulation into a three-dimensional point-cloud

representation. This involves utilizing the polynomial representations of buildings from SUMO and employing realistic three-dimensional meshes for the vehicles, as shown in Figure 2. We attach emulated camera sensors to both parked and moving vehicles within this representation. In this study, we have mounted four camera sensors on the observers, specifically at the front, rear, left, and right sides. The point cloud data of the vehicles and buildings are subsequently projected onto the camera plane of the emulated sensors. This projection produces a depth-image-like representation on the camera plane, including object assignments for the projected points as shown in Figure 2. Using the KITTI thresholds, we can approximate which vehicles would be detectable in a comparable real-world scenario, utilizing real-world RGB-camera sensors and state-of-the-art 3D object detection algorithms. With this so-called computer-vision approach, it is, therefore, possible to determine which other traffic participants are detectable and available for the traffic state estimation by the moving and parking observers. The projection of the point clouds onto the camera plane is computationally expensive, which results in a reduction of the simulation speed of up to 10 seconds for each observer. Consequently, we use the computer vision approach solely to generate a dataset for training neural networks, which then emulate the computer vision process directly from a Bird's Eye View (BEV) representation of the simulation. Each data point consists of the BEV representation, where an example is shown in Figure 2, of the current traffic situation centered around the observer, the two-dimensional vector from the observer to the vehicle for which detectability is to be determined, and the corresponding binary detectable label derived from the computer vision approach. In this fashion, we generate 100,000 data points. With these data points, we then train the NNs where we specifically use an adapted version of the Vision Transformer presented in (*32*). After training, the model achieves an accuracy of over 94% on the test split, i.e., in 94% of the cases, the NNs correctly determine if a vehicle is detectable according to the computer vision approach. The significant benefit of the NN approach is the inference speed, i.e., the duration it takes to determine if a vehicle is detectable by the observer. On average, the NN approach requires 0.05 seconds to determine the detectability of all surrounding vehicles per observer.

**FIGURE 2**: Visualization of the three-dimensional point cloud surrounding the observer under investigation (left), the corresponding depth image captured by the front camera (center), and the BEV representation of the same scene utilized for the NN approach (right).

**Traffic State Estimation**

Point Calculation Based on previous studies, we have demonstrated that estimating and predicting speeds $v$ is relatively more straightforward, whereas estimating densities $k$ is generally more challenging. Consequently, this affects the accuracy of traffic flow calculations, $q = ks$. Therefore, this study specifically focuses on estimating and predicting traffic densities at the link level. For MO, we estimate the real-time lane density $\hat{k}$ (vehicles/km) of an individual moving observer as Equation 1.

$$\hat{k} = \frac{\hat{n}+1}{2 \times (d + \frac{L}{2}) \times \overline{n}_l} \times 1000 \tag{1}$$

where,

- $\hat{n}$: Number of detected vehicles
- $d$: Detection distance (m)
- $L$: Average car length (m)
- $\overline{n}_l$: Number of lanes of the edge

However, since a point estimation can be unstable, we cap the estimated density $\hat{k}$ at the maximum value $k_{max}$ to avoid unrealistic high densities. $k_{max}$ is calculated by the 95% maximum density from the data. For PO, the density estimation is nearly identical, except the PO itself is not included in the count.

$$\hat{k} = \frac{\hat{n}}{2 \times (d + \frac{L}{2}) \times \overline{n}_l} \times 1000 \tag{2}$$

Area estimation Area estimation refers to predicting the density of a specific lane by analyzing the observations from both moving and parked observers within a designated area around that lane. Unlike point estimation, which focuses on a single observation from a moving or parked observer, area estimation aggregates data from all relevant observations within the specified area. To illustrate our Weighted Spatio-Temporal Estimation (WSTE) method, we also develop some baseline methods to compare.

**Baseline, All historical data**

For each lane ID $i$ and time interval $t$, the predicted density $\widehat{k_{i,t}}$ is calculated using exponential decay weights:

$$\widehat{k_{i,t}} = \frac{\sum_{j=1}^{N} \widehat{k_{i,t-j\cdot t_0}} \cdot w_j}{\sum_{j=1}^{N} w_j} \tag{3}$$

where $\widehat{k_{i,t-j\cdot t_0}}$ is the lane density at the past interval $t - j \cdot t_0$ seconds, $w_j$ is the exponential decay weight for the $j$-th past interval, and $N$ is the number of past intervals with available data. $t_0$ represents the aggregated estimation interval. The exponential decay weight $w_j$ is defined as:

$$w_j = \exp(-\lambda \cdot j) \tag{4}$$

Where $\lambda$ is the decay constant determining the decay rate.

If no past densities are available for the lane ID $i$ at the specified past intervals, $\widehat{k_{i,t}}$ remains at the same density in the last interval. This approach ensures the predicted density is weighted more heavily towards recent intervals. This baseline method serves as the high bar for traffic state estimation because we have all the historical data to predict the next interval.

**Baseline, No historical data + MO + PO**

For this baseline method, we calculate the predicted density $\widehat{k_{i,t}}$ as a weighted mean of the densities estimated by MO and PO. The weights are denoted by $w_{\text{mo}}$ and $w_{\text{po}}$, respectively. The equation for $\widehat{k_{i,t}}$ is given by:

$$\hat{k_{i,t}} = \frac{\hat{k_{i,t,mo}} \cdot w_{\text{mo}} + \hat{k_{i,t,po}} \cdot w_{\text{po}}}{w_{\text{mo}} + w_{\text{po}}} \tag{5}$$

Where:

- $\hat{k_{i,t,mo}}$ is the weighted density estimated by MO.
- $\hat{k_{i,t,po}}$ is the weighted density estimated by PO.

The sum of weights $w_{\text{mo}}$ and $w_{\text{po}}$ are generated using a Gaussian distribution within a radius $r$. All MO and PO within the radius $r$ are considered. The distance $d$ is calculated between the MO/PO and the centroid point of the predicted lane. The weights are then calculated based on the distance $d$ using the Gaussian function:

$$w(d) = \exp\left(-\frac{d^2}{2\sigma^2}\right) \tag{6}$$

Where:

- $\sigma$ is the standard deviation of the Gaussian distribution, often set to a value related to the radius $r$.

Since this estimation uses the estimation from MO and PO and no historical data have been considered, we use this method as the baseline for the low bar for traffic state estimation.

**Weighted Spatio-Temporal Estimation (WSTE)**

This method combines the power from both spatial and temporal perspectives. We start with an initial value $k_{t_0}$, which is the true density at the initial timestamp $t_0$. This initial density is from the simulation simulation and is the only information we use from the historical data. We also retrieve the initial observations from MO $\hat{k_{i,t_0,mo}}$ and/or PO and $\hat{k_{i,t_0,po}}$.

At each subsequent timestamp, we calculate the difference ($\Delta_{i,t}$) between the observed density at the current timestamp $\hat{k_{i,t,mo}}$ and $\hat{k_{i,t,po}}$ and the initial density $\hat{k_{i,t_0,mo}}$ and PO and $\hat{k_{i,t_0,po}}$, then use this difference to adjust the initial density proportionally.

The formula for the predicted density $\hat{k_{i,t}}$ at timestamp $t$ is given by:

$$\hat{k_{i,t}} = k_{t_0} \times (1 + \Delta_{i,t,\text{mo}}) \tag{7}$$

or

$$\hat{k_{i,t}} = k_{t_0} \times (1 + \Delta_{i,t,\text{po}}) \tag{8}$$

Where:

- $\hat{k_{i,t}}$ is the predicted density at timestamp $t$.
- $\Delta_{i,t,\text{mo}}$ is the percentage change in density observed by MO between $t_0$ and $t$.
- $\Delta_{i,t,\text{po}}$ is the percentage change in density observed by PO between $t_0$ and $t$.

The percentage delta $\Delta_{i,t}$ is calculated as:

$$\Delta_{i,t,\text{mo}} = \frac{\hat{k_{i,t,\text{mo}}} - \hat{k_{i,t_0,\text{mo}}}}{\hat{k_{i,t_0,\text{mo}}}} \tag{9}$$

or

$$\Delta_{i,t,\text{po}} = \frac{\hat{k_{i,t,\text{po}}} - \hat{k_{i,t_0,\text{po}}}}{\hat{k_{i,t_0,\text{po}}}} \tag{10}$$

Also, we combine these two differences using the weighted method we mentioned for the baseline. The combined predicted density $\hat{k_{i,t}}$ is given by:

$$\hat{k_{i,t}} = \frac{(k_{t_0} \times (1+\Delta_{i,t,\text{mo}})) \cdot w_{\text{mo}} + \left(k_{t_0} \times (1+\Delta_{i,t,\text{po}})\right) \cdot w_{\text{po}}}{w_{\text{mo}} + w_{\text{po}}} \tag{11}$$

By applying these calculations, we can compare the influence from MO-only, PO-only, and the combined influence. Since the point estimation is biased, the idea behind the WSTE is to use the change of observations instead of the observations to estimate the traffic densities.

Metrics In this part, we present the equations and explanations for three commonly used error metrics in traffic state estimation: Root Mean Squared Error (RMSE), Coefficient of Determination ($R^2$), and Mean Absolute Percentage Error (MAPE).

**Root Mean Squared Error (RMSE)**

The RMSE measures the average magnitude of the errors between the predicted values and the actual values. It is defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{12}$$

Where:

- $n$ is the number of observations,
- $y_i$ is the actual value for the $i$-th observation,
- $\hat{y}_i$ is the predicted value for the $i$-th observation.

**Coefficient of Determination ($R^2$)**

The $R^2$ metric, also known as the Coefficient of Determination, measures the proportion of the variance in the dependent variable that is predictable from the independent variable(s), as shown here:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{13}$$

Where:

- $y_i$ is the actual value for the $i$-th observation,
- $\hat{y}_i$ is the predicted value for the $i$-th observation,
- $\bar{y}$ is the mean of the actual values.

**Mean Absolute Percentage Error (MAPE)**

The Mean Absolute Percentage Error (MAPE) evaluates the accuracy of a predictive model by calculating the percentage difference between actual and predicted values. We include MAPE alongside Root Mean Square Error (RMSE) to assess errors in terms of absolute volumes and relative percentages.

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{14}$$

Where:

- $n$ is the number of observations,
- $y_i$ is the actual value for the $i$-th observation,
- $\hat{y}_i$ is the predicted value for the $i$-th observation.

## EXPERIMENTS AND RESULTS

In this section, we first describe the simulation environment. Then, we present and discuss the results regarding three research questions.

### Simulation Description

As in (*2*), we utilize the SUMO simulation in Ingolstadt, Germany, as the whole network was modeled and calibrated in detail (*33*) as shown in Figure 3. Prior calibrations, including map association, traffic light program emulation, and traffic flow calibration, have already been performed on this network in (*29*).

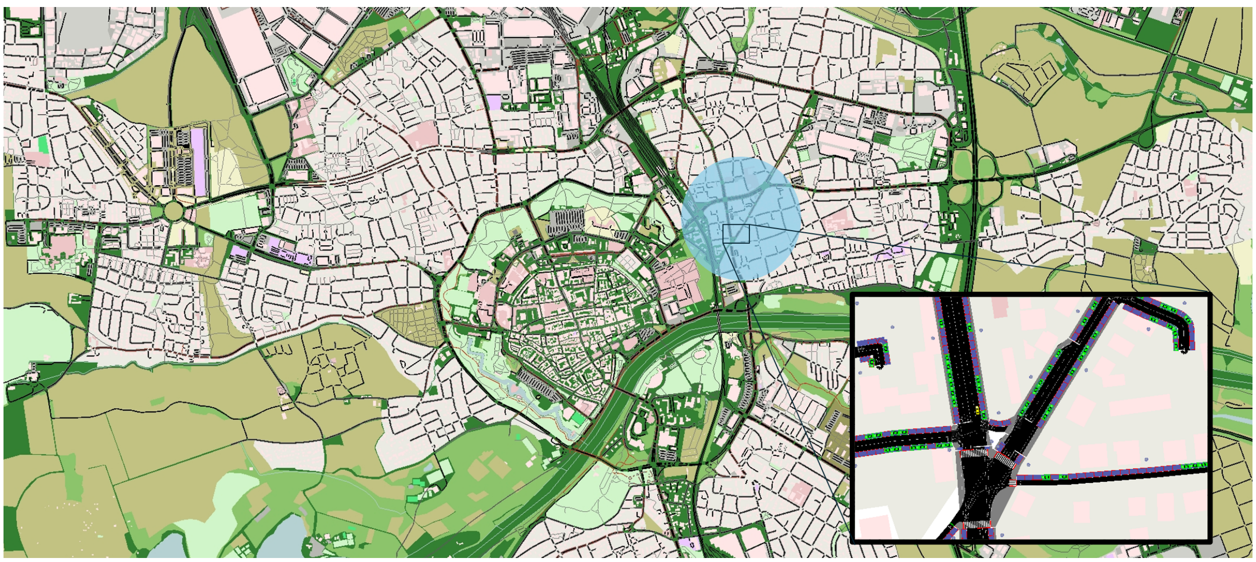


**FIGURE 3**: SUMO simulation of Ingolstadt, as presented in (*29*), with the area investigated in this paper highlighted in blue. Additionally, a close-up view illustrates the parked vehicles within the simulation.

The simulation uses a 5-minute aggregation level to balance efficient results with acceptable temporal resolution. We use the exact interval for all prediction methods, which means $t_0 = 300sec$. Traffic states are evaluated for lanes and edges. We simulate 6 a.m. to 9 a.m. to capture the network accumulation phase while avoiding over-congested networks during the daytime. For both the distance- and NN-based detection, we collect the detected vehicles for the penetration rates [1, 2, 5, 10, 20]% and for a time from 6 a.m. to 9 a.m. with the traffic demand defined in (*29*). The penetration rates apply to both Moving Observers (MO) and Parked Observers (PO); however, for MO, the penetration rate represents the proportion of MO among all vehicles, while for PO, it indicates the percentage of available parking spots occupied. For simplicity, we assume that 50% of all parking spots are available and then multiply the penetration rates to determine the proportion of PO.

For the *Baseline - All historical data method*, we select $N$, which is the number of past intervals with available data, as 5. This means we consider the past 25 minutes to predict the next 5-minute interval. For the *Baseline, No historical data + MO + PO*, we select the radius $r$ as $50m$ to consider all MO and PO within this distance. And the corresponding standard deviation of the

Gaussian distribution $\sigma$ is one-third of the radius $r$, which is $50/3 \approx 13.3m$.

### Results of Detection and Data Collection

First, we compare the point estimation results for different detection methods in Figure 4. Three solid lines stand for the distance-based methods and a dashed line for the camera-based NN method. For moving observers to the left, distance-based detection methods have better performances as they can detect all vehicles without occlusion; however, we notice two phenomena: first, distance-based methods perform the best when considering all the vehicles within 50 meters; second, the error increases after reaching the lowest point as 2% penetration rates. It implies that with high penetration rates, there could be 1) *Repeated Detection*, i.e., one vehicle can be detected by multiple observers, thus causing overestimating, 2) a short detection distance (25-meter) could result in unstable estimation, while a long one (100-meter) could intensify the repeated detection. These can be because we do not collect the information on which lane or direction the detected vehicle is located, as we explained in Subsection 3.2. The emulation of the camera-based detection generally performs worse than the distance-based detection. This is caused by the reduced number of detected vehicles since the method considers occlusion as it would occur by real-world sensors.

However, this trend is different for PO: first, the NN method, which also takes occlusions into account, can outperform the short (25-meter) and long (100-meter) range distance-based methods; second, the errors among different methods drop down the lowest point at roughly 10% penetration rate. The most important reason is the vast difference between MO and PO: whether the observer counts itself into density calculation. From the previous study (*2*), we already proved that MO is biased, especially during an off-peak time, because the section where one MO is located will be at least one vehicle, while most of the section could be empty. This bias can be solved with PO, and *Repeated Detection* could be less common. We also notice that since POs are located at roadside parking spaces, they have fewer possibilities for repeated detection.

Note that, in reality, we can never achieve distance-based detection. AVs still rely on cameras and other sensors to detect the surrounding traffic. Therefore, **we will only apply the NN detection for the remaining analyses**.

### Results of Point Estimation

Figure 5 compares the point estimation for MO and PO with different penetration rates. Due to the essence that the detected vehicles can be spread among different lanes, the point estimation is supposed to be and also proved in the figure to be scattered widely. In general, we can conclude that in Figure 5(a), moving observers tend to overestimate the densities, especially when the demand is relatively low, while parked observers have a better estimation of the traffic densities in Figure 5(b).

Because of the inherent biases in both observers, we do not focus on improving estimating densities for a single observation from an observer. Since the ultimate goal is to estimate or predict the traffic densities at the link level, we try to include these point estimation results in the traffic state estimation with other knowledge.

### Results of Traffic State Estimation

Following the methods of TSE, which we introduce in Section 3.3, we compare two baseline methods and three Weighted Spatio-Temporal Estimation (WSTE) methods in Figure 6. Note that

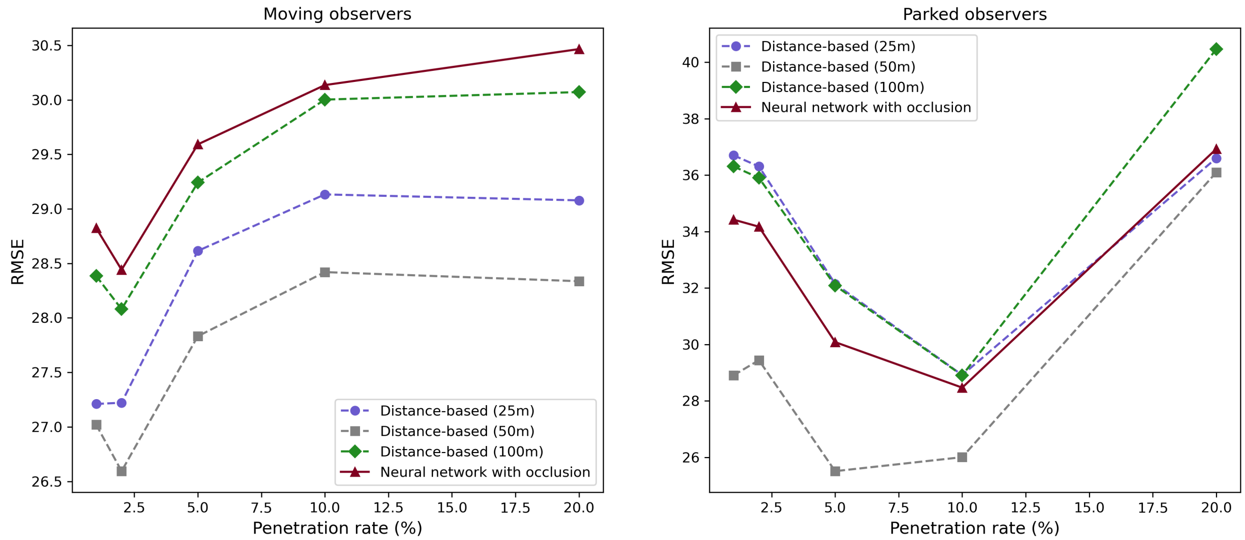


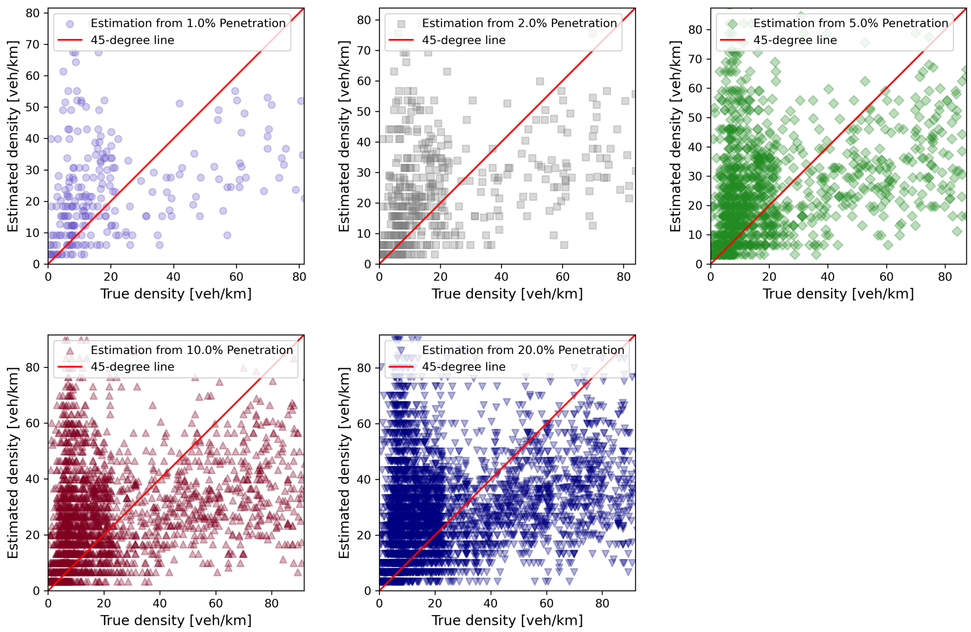

**FIGURE 4**: Comparison among different detection methods: the NN with occlusion method is not accurate enough as distance-based methods in MO, while can achieve a similar level using PO

two baseline methods are visualized in both figures, while in the correct figure for R², the R² value for *Baseline, No historical data + MO + PO* is below 0 and is invisible from the visualization area.

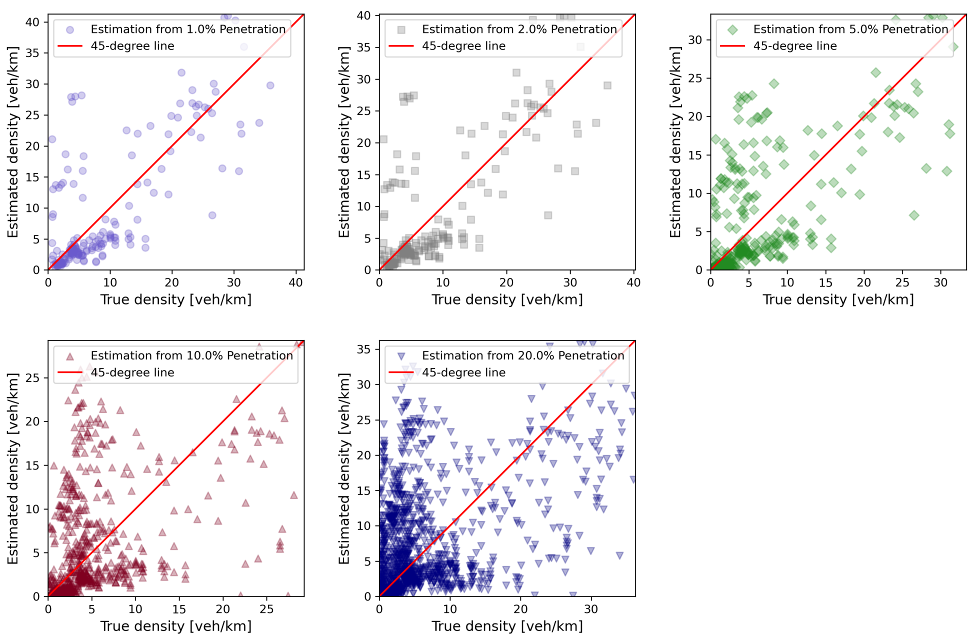
For RMSE, *Baseline, All historical data* leads with the slightest error. In comparison, without historical data, *Baseline, No historical data + MO + PO* still tends to have not too big error thanks to the spatial weights. However, both baselines have no significant improvements in accuracy after the penetration rates reach 5%, implying the limitations of both observers. *Weighted Spatio-Temporal Estimation (WSTE)* have entirely different trends. With only the true density at the first timestamp, *Initial historical data + MO* can reduce the error by increasing penetration rates to 10%. The error increases again at 20%, most probably due to the *Repeated Detection* thus, overestimation. *Initial historical data + PO*, on the other hand, constantly reduces errors with increased penetration rates. The resulting mixed method *Initial historical data + MO + PO* is more towards the *Initial historical data + MO* side, implying the higher weights from the MO side.

Results for R² are different, especially for *Baseline, No historical data + MO + PO*. The negative R² indicates that this baseline method might not compete with an estimation based on the historical average. This might be because without knowing any true densities, the point estimation from AVaS is unstable. For our WSTE method, we can observe that *Initial historical data + PO* benefits from the increased penetration rates, while *Initial historical data + MO* is the opposite. The drop in R² value for MO indicates that the variance due to overestimation (Repeated Detection) becomes bigger than the variation of the traffic state. PO always have available spots even with the increased penetration rates, and this fact makes PO more reliable in avoiding Repeated Detection. That leads to our thinking that we can combine a fleet of MO and PO with different penetration rates in the future TSE study.

An overview of all metrics of all scenarios is listed in Table 3. For *Initial historical data + PO* and *Initial historical data + MO*, the overall best penetration rates are bold, again highlighting




((a)) Results of point estimation for moving observers




((b)) Results of point estimation for parked observers

**FIGURE 5**: Comparison of point estimation for different penetration rates using NN detection methods: MO tend to be more scattered than PO

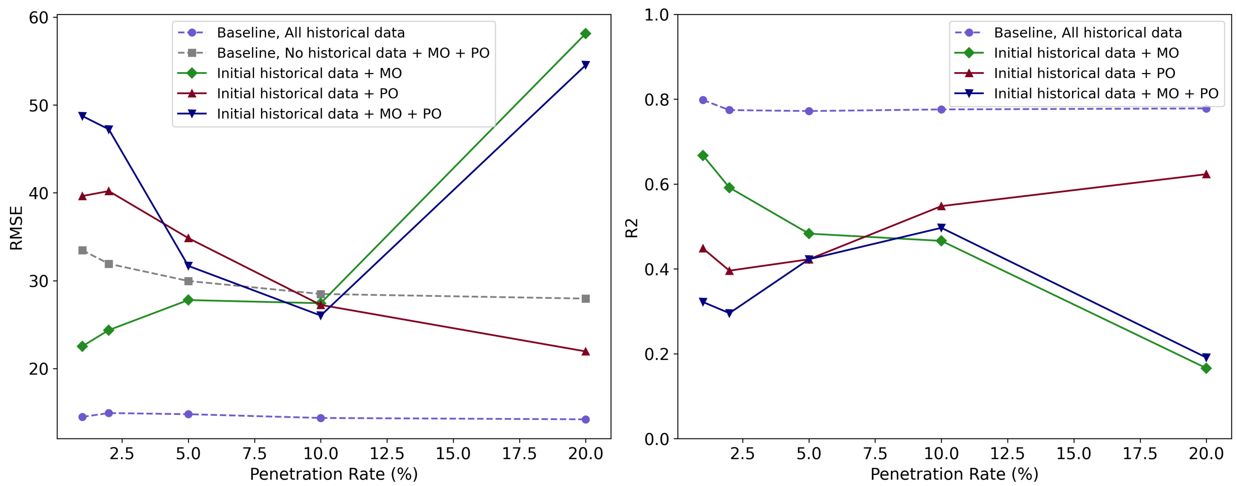



**FIGURE 6**: Comparison among different estimation methods: compared to two baselines with & without all historical data, our Weighted Spatio-Temporal Estimation (WSTE) method can reach high accuracies with RMSE and $R^2$ with just initial historical data.

that TSE can benefit from different penetration rates from MO and PO.

## CONCLUSION

### Summary

Originating from the first concept of AV-as-Sensors (AVaS) (*1*), we simulated the data collection process in (*2*). Using the collected data, this study has explored various approaches for traffic state estimation (TSE). In this study, we 1) implement a new camera-based detection method, 2) model the parked vehicles in SUMO, and 3) estimate the traffic densities using Weighted Spatio-Temporal Estimation for moving and parked observers (MO and PO), compared to two baseline methods. The results show that even with a low penetration rate, MO and/or PO can have a rather good estimation of local densities with just one initial true observation. Thus, we answer our research questions raised in Section 2.3 as follows:

1. *How accurately can the camera-based detection method detect surrounding traffic compared to distance-based methods?* From the results in 4.2, we can conclude that compared to distance-based methods, the camera-based detection method can benefit the traffic state estimation at the same level, especially for PO.
2. *How much information can parked observers provide to estimate the traffic?* Compared to the overestimation from MO, PO can modify the estimation from the results presented in Section 4.3. Also, from the metrics in Section 4.4, we could see that PO overperforms MO, especially with high penetration rates.
3. *How does our Weighted Spatio-Temporal Estimation (WSTE) method compare to baseline methods?* Compared with two baseline methods with all historical data (high bar) and AVaS-only data (low bar), our WSTE method can almost reach the same accuracy with ground truth with only one initial true observation.

| **Condition** | | | | **Penetration Rate (%)** | | | | |
|---|---|---|---|---|---|---|---|---|
| Parked | Moving | Historical Data | Metric | 1 | 2 | 5 | 10 | 20 |
| No | No | All | RMSE | 14.49 | 14.93 | 14.80 | 14.37 | 14.22 |
| | | | $R^2$ | 0.798 | 0.774 | 0.772 | 0.776 | 0.778 |
| | | | MAPE | 0.390 | 0.400 | 0.409 | 0.429 | 0.428 |
| Yes | Yes | No | RMSE | 33.47 | 31.93 | 29.96 | 28.48 | 27.96 |
| | | | $R^2$ | -2.809 | -2.592 | -2.256 | -2.519 | -2.586 |
| | | | MAPE | 0.988 | 0.950 | 0.858 | 0.803 | 0.800 |
| No | Yes | **Initial (WSTE)** | RMSE | **22.53** | 24.37 | 27.79 | 27.43 | 58.13 |
| | | | $R^2$ | **0.667** | 0.592 | 0.483 | 0.466 | 0.166 |
| | | | MAPE | **0.735** | 0.785 | 0.990 | 1.066 | 1.531 |
| Yes | No | **Initial (WSTE)** | RMSE | 39.63 | 40.20 | 34.85 | 27.24 | **21.95** |
| | | | $R^2$ | 0.448 | 0.396 | 0.423 | 0.548 | **0.623** |
| | | | MAPE | 1.024 | 1.053 | 1.101 | 1.096 | **1.294** |
| Yes | Yes | **Initial (WSTE)** | RMSE | 48.74 | 47.23 | 31.68 | 26.03 | 54.54 |
| | | | $R^2$ | 0.322 | 0.296 | 0.423 | 0.497 | 0.191 |
| | | | MAPE | 0.949 | 0.921 | 1.017 | 0.992 | 1.572 |

**TABLE 3**: Sensitivity Analysis considering impact of parked and moving vehicles, historical data, and penetration rates: our WSTE methods (bold) perform best with high penetration rates for parked observers and low penetration rates for moving observers.

**Future Research**

While this study has addressed all the research questions, we admit that this study focuses on traffic densities at the link level, with certain simplifications in the simulation and detection methods. Besides, traffic state estimation is limited to randomly selecting moving and parked vehicles. To address these limitations, we propose future research with different modules in Figure 1.

Module - Simulation Now, the parked vehicles are randomly selected for available parking spots, which are randomly modeled in the simulation as well. In the future, we plan to investigate the amount of contributions to traffic state estimation from different locations. Thus, a future strategic vehicle assignment algorithm can be developed, so that idle vehicles can park where they can provide the most information and data for traffic state estimation.

Module - Data collection The current detection status only uses the information on how many vehicles can be "detected" by the ego observer without further information on which lane or direction they are located. Since the emulation is already based on 3D object detection, in the future, we will utilize this information not only for lane-specific but also for the neighboring lane and opposite-direction traffic density estimation.

Module - Traffic state estimation We already notice in Section 4.4 that moving and parked observers have different trends of performances with increased penetration rates. In the future, instead of using the same penetration rate for both observers, we will conduct the sensitivity analysis for different penetration rates between moving and parked observers. Also, since moving observers are outnumbering parked observers now, the weighted estimation always has higher weights and influences from moving observers. We will update the weighting system in future research as well.

By addressing these different modules, our ultimate goal is to use AV-as-Sensors as a supplementary or even replace the current fixed-location detector system for traffic state estimation. We will study the possibilities of using this concept by integrating it into the operation of Autonomous Mobility-on-Demand systems.

## ACKNOWLEDGEMENTS

This work was funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) and by the European Union in the frame of NextGenerationEU within the project STADT:up (FKZ 19A22006T). [The authors would like to thank the consortium for the successful cooperation.]

## AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: Study conception and design: Yunfei Zhang, Jeremias Gerner, Mario Ilic, Stefanie Schmidtner, Klaus Bogenberger; Data collection: Jeremias Gerner, Mario Ilic; Analysis and interpretation of results: Yunfei Zhang; Draft manuscript preparation: Yunfei Zhang, Jeremias Gerner, Mario Ilic. All authors reviewed the results and approved the final version of the manuscript.

## DECLARATION OF CONFLICTING INTEREST

All authors do not have any conflicts of interest to declare.

## REFERENCES

1. Zhang, Y., M. Ilic, and K. Bogenberger, A Novel Concept of Traffic Data Collection and Utilization: Autonomous Vehicles as a Sensor. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023.
2. Zhang, Y., M. Ilic, and K. Bogenberger, Autonomous Vehicles as a Sensor: Simulating Data Collection Process. *arXiv preprint arXiv:2308.11629*, 2023.
3. Wang, Y. and M. Papageorgiou, Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. *Transportation Research Part B: Methodological*, Vol. 39, No. 2, 2005, pp. 141–167.
4. Thai, J. and A. M. Bayen, State Estimation for Polyhedral Hybrid Systems and Applications to the Godunov Scheme for Highway Traffic Estimation. *IEEE Transactions on Automatic Control*, Vol. 60, No. 2, 2015, pp. 311–326.
5. O'Keeffe, K. P., A. Anjomshoaa, S. H. Strogatz, P. Santi, and C. Ratti, Quantifying the sensing power of vehicle fleets. *Proceedings of the National Academy of Sciences*, Vol. 116, No. 26, 2019, pp. 12752–12757.
6. Herring, R., A. Hofleitner, P. Abbeel, and A. Bayen, Estimating arterial traffic conditions using sparse probe data. In *13th International IEEE Conference on Intelligent Transportation Systems*, 2010, pp. 929–936.
7. Yu, J., M. E. Stettler, P. Angeloudis, S. Hu, and X. M. Chen, Urban network-wide traffic speed estimation with massive ride-sourcing GPS traces. *Transportation Research Part C: Emerging Technologies*, Vol. 112, 2020, pp. 136–152.
8. Mark Richard Wilby, A. B. R. G., Juan José Vinagre Díaz and M. Ángel Sotelo, Lightweight Occupancy Estimation on Freeways Using Extended Floating Car Data. *Journal of Intelligent Transportation Systems*, Vol. 18, No. 2, 2014, pp. 149–163.
9. Seo, T., T. Kusakabe, and Y. Asakura, Traffic State Estimation with the Advanced Probe Vehicles Using Data Assimilation. *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 824–830.
10. Seo, T., T. Kusakabe, and Y. Asakura, Estimation of flow and density using probe vehicles with spacing measurement equipment. *Transportation Research Part C: Emerging Technologies*, Vol. 53, 2015, pp. 134–150.
11. Fountoulakis, M., N. Bekiaris-Liberis, C. Roncoli, I. Papamichail, and M. Papageorgiou, Highway traffic state estimation with mixed connected and conventional vehicles: Microscopic simulation-based testing. *Transportation Research Part C: Emerging Technologies*, Vol. 78, 2017, pp. 13–33.
12. Gerner, J., D. Rößle, D. Cremers, K. Bogenberger, T. Schö, and S. Schmidtner, Enhancing Realistic Floating Car Observers in Microscopic Traffic Simulation. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023.
13. Wardrop, J. G. and G. Charlesworth, A Method of Estimating Speed and Flow of Traffic from a Moving Vehicle. *Proceedings of the Institution of Civil Engineers*, Vol. 3, No. 1, 1954, pp. 158–171.
14. Mulligan, A. and A. Nicholson, Uncertainty in Traffic Flow Estimation Using the Moving-Observer Method, 2002.
15. Czogalla, O. and S. Naumann, TRAVEL TIME ESTIMATION USING FLOATING CAR OBSERVERS, 2007.

16. Wolf, F., S. Naumann, C. Engel, and R. Schönrock, Floating Car Observer–Approaches for Traffic Management Strategies by Analysing Oncoming Vehicles. In *The Second International Workshop on Intelligent Vehicle Control Systems*, SCITEPRESS, 2008, Vol. 2, pp. 83–92.
17. Kühnel, C., C. Leitzke, and R. Hoyer, Evaluation of Floating Car Observer Algorithms using microscopic Traffic Flow Simulation, 2009.
18. Florin, R. and S. Olariu, On a variant of the mobile observer method. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, No. 2, 2016, pp. 441–449.
19. Schäfer, M. and R. Hoyer, Bewegte Beobachter zur mobilen Erfassung netzweiter Verkehrszustände. *Innovative Produkte und Dienstleistungen in der Mobilität: Technische und betriebswirtschaftliche Aspekte*, 2017, pp. 267–283.
20. *Verkehrszustandsschätzung mittels vernetzter Floating Car Observer*, 2017.
21. van Erp, P. B., V. L. Knoop, and S. P. Hoogendoorn, Macroscopic traffic state estimation using relative flows from stationary and moving observers. *Transportation Research Part B: Methodological*, Vol. 114, 2018, pp. 281–299.
22. Guerrieri, M., G. Parla, and R. Mauro, Traffic Flow Variables Estimation: An Automated Procedure Based on Moving Observer Method. Potential Application for Autonomous Vehicles. *Transport and Telecommunication*, Vol. 20, No. 3, 2019, pp. 205–214.
23. van Erp, P. B., V. L. Knoop, and S. P. Hoogendoorn, On the value of relative flow data. *Transportation Research Procedia*, Vol. 38, 2019, pp. 180–200, journal of Transportation and Traffic Theory.
24. Langer, M., T. Schien, M. Harth, R. Kates, and K. Bogenberger, An Improved Moving Observer Method for Traffic Flow Estimation at Signalized Intersections. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 1498–1503.
25. van Erp, P., I. Netherlands Research School for Transport, and Logistics, *Relative Flow Data: New Opportunities for Traffic State Estimation*. TRAIL thesis series, Netherlands TRAIL Research School, 2020.
26. Ma, W. and S. Qian, High-Resolution Traffic Sensing with Probe Autonomous Vehicles: A Data-Driven Approach. *Sensors*, Vol. 21, No. 2, 2021.
27. Florin, R. and S. Olariu, Real-Time Traffic Density Estimation: Putting on-Coming Traffic to Work. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 24, No. 1, 2023, pp. 1374–1383.
28. Schäfer, M., Verkehrszustandsschätzung mittels vernetzter Floating Car Observer. *HEUREKA´17 – Optimierung in Transport und Verkehr*, 2017, conference: HEUREKA´17 – Optimierung in Transport und Verkehr, Stuttgart.
29. Harth, M., M. Langer, and K. Bogenberger, Automated calibration of traffic demand and traffic lights in SUMO using real-world observations. In *SUMO Conference Proceedings*, 2021, Vol. 2, pp. 133–148.
30. Gerner, J., K. Bogenberger, and S. Schmidtner, Investigating Floating Car Observers for Traffic State Estimation through Co-Simulation. In *The 5th Symposium on Management of Future Motorway and Urban Traffic Systems (MFTS 2024)*, Heraklion, Crete, Greece, 2024.
31. Geiger, A., P. Lenz, and R. Urtasun, Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

32. Dosovitskiy, A., L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
33. TUM-VT, *SUMO Ingolstadt*. Online, 2023, GitHub repository.