EAI Endorsed Transactions
on Smart Cities

Research Article EAI.EU

An Integrated Framework for Virtual Testing of Autonomous Vehicles in Mixed Urban Traffic

B. Caroleo[1,*], J. Sadeghi[1], C. Botta[1], S. Nikneshan[1] and M. Arnone[1]

[1]LINKS Foundation, Via P.C. Boggio 61, 10138, Turin, Italy

Abstract

INTRODUCTION: As cities gradually begin integrating autonomous vehicles into existing transport systems, it becomes essential to assess their potential impacts on traffic dynamics and safety in a comprehensive and systematic manner — particularly through tools that can anticipate impacts before actual on-road deployment.
OBJECTIVES: This paper aims to develop a data-driven and modular framework to evaluate the integration of autonomous mobility solutions in mixed traffic conditions.
METHODS: A data-driven approach combining sensor data collected during autonomous shuttle trials with video-based behavioural analysis of road users and calibrated traffic microsimulation is employed to perform ex-ante assessment of different deployment scenarios.
RESULTS: The framework enables the evaluation of the impacts of autonomous mobility solutions on traffic performance and safety, providing insights across multiple scenarios.
CONCLUSION: The framework supports informed decision-making and enhances the understanding of how autonomous mobility can be effectively integrated into urban environments.

Keywords: Cooperative Connected and Automated Mobility (CCAM), Autonomous Vehicles (AV), Traffic Management, Traffic Simulation, Virtual Testing, Urban Transportation

Received on 29 April 2025, accepted on 23 April 2026, published on 27 April 2026

doi: 10.4108/eetsc.9193

1. Introduction

Urban mobility is experiencing an era of transformation with the rise of Autonomous Vehicles (AVs) and Cooperative, Connected, and Automated Mobility (CCAM). Recent studies have been shifted toward autonomous transportation, particularly public autonomous shuttles using minibus technology which are expected to become an essential element of future urban mobility.

Current autonomous shuttle experiences are often limited to controlled environments, thus the main challenge is to integrate them into mixed traffic, where they interact with pedestrians, cyclists, and human-driven vehicles.

Integrating autonomous vehicles in urban context requires careful planning to ensure safety and efficiency. As public autonomous shuttles operate on predefined authorized routes, traffic simulation frameworks play a crucial role in evaluating their impacts by allowing researchers and planners to model and analyse various traffic scenarios. Microscopic traffic simulation models are useful for studying individual vehicle dynamics, interactions, and localized traffic phenomena [1], providing insights also in emissions, travel times, and safety metrics.

The city of Turin (Italy) serves as a test environment for studies on autonomous mobility within several European and national projects such as the European project SHOW [2], IN2CCAM [3], and the national project ToMove [4].

*Corresponding author. Email: brunella.caroleo@linksfoundation.com

Various types of data were collected from these trials, including sensor, infrastructure, and environmental data, providing a comprehensive dataset that supports simulation, validation, and improvement of AV systems in diverse urban scenarios.

Traffic simulation models often rely on aggregated data, which fails to capture the nuances of vehicle interactions. This is particularly challenging for modelling interactions between autonomous and human-driven vehicles, due to their distinct behavioural patterns. AVs rely on sensor-driven decision-making, whereas human drivers exhibit variable responses influenced by cognitive and environmental factors. Recent tracking technologies make it possible to continuously collect vehicle telemetry data, such as speed and location, enabling more realistic traffic analysis.

This work has two main objectives. First, to provide a comprehensive framework based on data from urban infrastructure and autonomous vehicle experimentations, simulating the behaviours and interactions of these vehicles with other road users. Second, to provide a tool that allows decision-makers to test AV deployment in the simulation environment before the actual deployment, thus assessing the impact of introducing AVs in specific areas and preventing potential issues regarding safety.

2. Related works

The exploration of Autonomous Vehicles (AVs) and their integration into urban mobility is a multifaceted endeavour that has captivated the attention of researchers, engineers, and urban planners worldwide. Considerable research has been dedicated to understanding the implications of AVs in urban environments. Early studies primarily focused on the technological aspects, such as sensor fusion, perception algorithms, and decision-making processes of autonomous systems [5]. As AV technology matured, attention shifted towards assessing the impact of these vehicles on traffic flow, congestion, and overall urban mobility [6][7][8].

The domain of autonomous shuttle services, in particular, has gained significant interest due to its potential to provide efficient last-mile connectivity and enhance public transportation systems. Initiatives such as the European SHOW project [2] have explored the integration of autonomous shuttles into public transport systems [9][10], emphasizing their role in advancing the understanding and deployment of autonomous shuttles in real-world urban settings [11].

To align the proposed research with the most recent scientific trends, it is essential to engage with the emerging digital twin (DT) paradigm and its associated ethical considerations. Recent advancements highlight the virtualization of motorway dynamics through run-time synchronized models that create a continuous synergy between simulation platforms and real-time traffic data streams [12]. Recent studies have further extended DT approaches to urban-scale environments, integrating multi-source data and real-time simulation frameworks to support predictive traffic management and AV deployment strategies [13]. High-fidelity modelling has also progressed through the integration of detailed vehicle dynamics and environmental factors directly into microscopic simulators to identify safety-critical issues [14], with emerging explorations into Generative AI (GenAI) for autonomous scenario augmentation [15]. Moreover, advanced frameworks are beginning to combine infrastructure-based sensing (e.g., camera systems) with reinforcement learning techniques to optimize traffic control policies within digital environments [16]. Parallel to these technical developments, recent literature emphasizes that the ethics of urban data use and democratic governance must be treated as central pillars of smart city DTs [17]. This includes the necessity for ethics-aware risk control frameworks that prioritize the protection of vulnerable road users [18] and the establishment of standards for cross-disciplinary collaboration between planners and ethicists [19]. From a policy perspective, these challenges also involve data governance, privacy, and equitable access to mobility services, particularly when dealing with large-scale high-resolution mobility datasets. The proposed framework supports these requirements by ensuring ethical data management through the use of anonymized trajectory data.

Traffic simulation has been used in numerous simulation research scenarios of AVs, highlighting its versatility and wide-ranging applications: it provides a helpful tool to answer complex research questions, to evaluate or test traffic management strategies and their impacts [20]. Simulation offers a safe and cost-effective method of testing and validating autonomous algorithms and control systems before real-world deployment. These simulations provide a crucial bridge between theory and practice, allowing researchers to refine AVs technologies in controlled virtual environments. Traffic simulators replicate situations and evaluate the consequences of different strategies for using AVs. In the majority of microscopic simulation research, AVs have been distinguished from human-driven vehicles based on their driving patterns, highlighting the pivotal role of driving behaviour in modelling AVs [21]. Understanding the AV driving behaviour is essential for creating realistic and effective simulations that mirror real-world conditions. Using microsimulation, some researchers studied the impact of different AVs penetration on traffic parameters [22] and on urban traffic flow and road capacity [7], highlighting potential risk that if AVs increase overall car use, they could strain traffic management. In the context of communication, technologies like Cooperative Awareness Messages (CAM) - defined by the European Telecommunications Standards Institute (ETSI) [23]- facilitate real-time data exchange between vehicles and infrastructure, further enriching the available datasets. Additionally, microsimulation has emerged as a pivotal tool for understanding the complex interactions between AVs and other road users in realistic traffic scenarios. Existing works delve into the development of microsimulation environments for AVs, often utilizing platforms such as SUMO (Simulation of Urban MObility) [24]. The assessment of impacts related to autonomous services, safety, and overall traffic dynamics becomes feasible through this simulation, guiding the identification of optimal strategies in various contexts. Recent digital twin-based simulation frameworks further enhance these capabilities by integrating real-time traffic data streams into microscopic environments, improving both prediction accuracy and responsiveness, although they often rely on aggregated rather than high-resolution trajectory data [13].

The safety of AVs and their interactions with pedestrians and conventional vehicles have been extensively investigated. Studies have utilized surrogate safety measures like Time to Collision (TTC) and Post Encroachment Time (PET) to evaluate and compare the safety of AV-pedestrian interactions with conventional vehicle-pedestrian interactions [25][26]. These approaches are increasingly complemented by video-based interaction analysis, enabling a more detailed characterization of micro-level conflicts and behavioural patterns in real-world urban environments.

In parallel, recent interdisciplinary research has focused on developing "human-like" automated vehicles whose behavioural decision-making aligns with human driving habits and cognitive expectations [27]. Empirical studies reveal that human drivers exhibit significantly higher stress levels during hard-braking scenarios with conventional vehicles compared to AVs, with behaviour and headway choices varying by driver age [28]. Trust remains a central factor, as shown by trials evaluating how pedestrians and cyclists perceive AV manoeuvres in shared spaces [25]. Innovative infrastructure-mediated communication has also emerged to foster safer navigation for pedestrians [29], especially at unsignalized crosswalks where pedestrians often exhibit discomfort when interacting with automated shuttles [8]. More recent studies further highlight the complexity of behavioural adaptation in mixed traffic, showing that interactions with AVs may generate uncertainty, reduced predictability, or lower trust in specific scenarios, particularly in multi-agent environments [29]. These findings stress the importance of incorporating perception, risk awareness, and decision-making processes into behavioural models.

From an operational and policy perspective, the impact of autonomous shuttles can be assessed across different infrastructure strategies, such as dynamic or separated lanes [30], while vehicle connectivity is being investigated to manage dynamic responses to unplanned urban events (e.g., road closures or accidents) via TraCI-mediated rerouting algorithms [31]. This approach is further refined by considering heterogeneous driving styles to ensure that virtual models accurately reflect real-world diversity [32] and by calibrating virtual environments based on specific automated shuttle service influence on urban networks [33].

In summary, the landscape of AV research is vast and continually evolving, but research contributions in the field of virtual testing for autonomous vehicles have been sectoral, and an overall view is lacking. There are some scientific contributions that try to give an overview by cross-referencing the contributions that may come from different domains [34], but there is a lack of real data from connected / autonomous vehicles, and thus no concrete results from such integrated frameworks. In particular, many existing approaches rely on synthetic or aggregated data and lack validation based on high-frequency real-world trajectory data in mixed urban traffic conditions.

Specifically, the proposed approach distinguishes its substantive novelty through three key dimensions. First, it leverages high-frequency real-world AV trajectory data to ensure behavioural realism often absent in synthetic studies. Second, it integrates video-based analysis with sensor data for vehicle tracking [35], enabling interaction-level analysis and precise calibration in unsignalized contexts. Third, it provides a modular testing environment that bridges individual trajectory detail with system-level policy evaluation. Additionally, the framework enables dynamic co-simulation, supporting ex-ante testing of traffic management and policy scenarios.

3. The case study

The City of Turin (Italy) with a robust CCAM ecosystem encompassing local institutions, research centres, the Traffic Control Centre, and the public transport operator served as a compelling case study regarding the integration of autonomous shuttle in urban transportation. During August–October 2022, as part of the H2020 SHOW project [2], two SAE level 4 autonomous shuttles (provided by Navya [36]) were deployed in mixed traffic.

The pilot was designed to offer a flexible, on-demand transport service for up to 15 passengers, circulating on a 5 km authorized route near the hospital area. Navya's shuttle was equipped with LIDAR and onboard sensors to navigate autonomously. Data collected from these sensors during the pre-demonstration phase, before passenger service began, was used in this study. The dataset included kinematic data sampled at 1 Hz. This level of granularity enabled the accurate replication of real-world behaviours within the simulation. In addition to trajectory data, traffic signal states transmitted via Signal Phase and Timing Extended Message (SPATEM), and road network features such as crosswalks and intersections were used. The use of these data into SUMO simulations ensured validity of the resulting models.

In addition, some videos were recorded along the section of the shuttle route at an unsignalized crosswalk with moderate traffic (Figure 1). This allowed the study of natural interactions between road users, including pedestrians, vehicles, and the AV. Two Garmin VIRB™ action cameras (1080p HD, 30 fps) mounted on a pole at a height of 10.80 m, were used for this purpose.

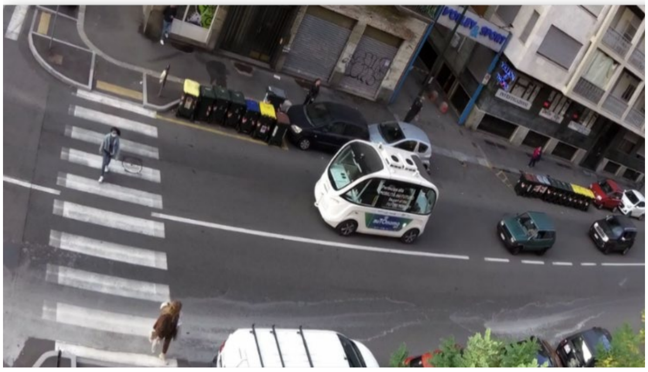

Figure 1. Snapshot of the monitoring area

4. Methodology

The proposed framework consists of three main phases: data gathering, data analysis and modelling, and traffic microsimulation (Figure 2). This modular structure allows flexibility across diverse urban contexts. The Data Gathering phase collects inputs from multiple range of data sources, including sensor, infrastructure, and camera-based data. The Data Analysis and Modelling phase preprocesses, validates, and enriches this data through statistical methods and machine learning techniques. The Traffic Microsimulation phase integrates the processed data into SUMO to simulate mixed-traffic scenarios and evaluate the impacts of AV deployment. The following subsections describe each phase of the Turin case study.

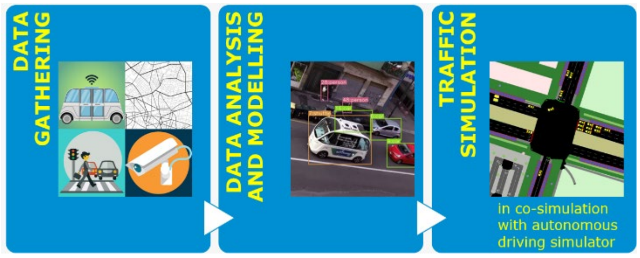

Figure 2. Proposed framework

4.1. Data gathering

The foundation of the proposed framework is based on an extensive and detailed data collection process. The types of data gathered in this case study are as follows:

(i) **Sensor data** from the autonomous vehicle constituted the primary source. In accordance with the Italian Smart Road Decree, which defines the minimum requirements for autonomous vehicles operations in Italy, the shuttles logged time-series data at regular intervals. The collected data included operational and kinematic variables such as WGS84 position, operation mode (autonomous or manual), instantaneous speed and acceleration, vehicle dynamics, steering and braking states, lightening and signalling status and other system-level indicators, as well as V2X messages received and transmitted. High-frequency operational logging of this kind is an intrinsic byproduct of AV operation itself and is therefore expected to be available wherever autonomous vehicles are deployed.

(ii) **Infrastructure data** from signal groups along the experimentation route were made available through the City's Traffic Control Centre. These signal groups transmitted phase and timing information via SPATEM (Signal Phase and Timing Extended Message), providing the time-to-green and time-to-red countdowns for each intersection.

(iii) **Road network data** were supplied by the Municipality of Turin in collaboration with the Traffic Operation Centre. This included the road geometry for the authorised zone, cross-referenced with OpenStreetMap, along with traffic demand information, and relevant data for the authorized zone. Public transport routes and stops were obtained from the local operator in GTFS format, allowing their incorporation into the simulation as an additional traffic layer.

(iv) **Camera recordings** were collected at a single unsignalized crosswalk along the shuttle route during different days. Interactions among different road users in the scene were captured. This location was deliberately chosen because the absence of signal control allowed the observation of unregulated crossing behaviour — the type of interaction that poses the greatest uncertainty for AV deployment in urban areas.

4.2. Data analysis and modelling

Once data is collected in the previous phase, two parallel analysis streams are carried out: a sensor-based analysis of AV telemetry data and a camera-based analysis of video recordings, as shown in Figures 3 and 5.

Sensor-based analysis

In the case study presented in this paper, high-frequency data from AV sensors were pre-processed and validated using statistical techniques in Python, ensuring the reliability of the dataset. Compliance with legislative requirements was also systematically checked to ensure adherence to safety and operational standards. Speed profiles were then derived from AV telemetry data to inform microsimulation settings and to support model validation.

Anomaly detection was then performed using the Median Absolute Deviation (MAD) [37], a robust statistical method for identifying outliers. The analysis focused on abrupt changes in vehicle speed, which were indicative of irregular behaviour during the operation.

MAD was selected due to its robustness to outliers and its flexibility, allowing the detection threshold to be adjusted according to the analysis requirements. Feature relationships were also examined through correlation analysis, providing insights into dependencies among variables and supporting a better understanding of AV operational dynamics.

Finally, a PowerBI dashboard has been created to visualize sensor-based data, offering a user-friendly, interactive interface. This tool visualizes key metrics and trends, enabling stakeholders and end users to explore the data. Sensor data during the shuttle experimentation were periodically uploaded to PowerBI as structured datasets following a batch processing rather than semi-real-time approach.

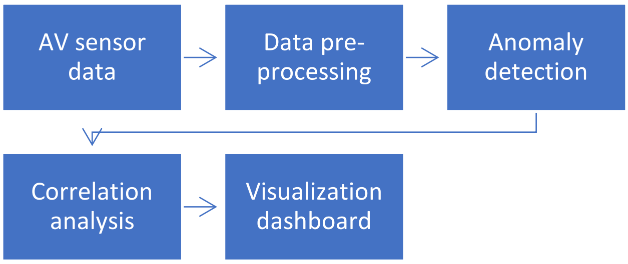

Figure 3. Sensor-based analysis workflow

Camera-based analysis

Given the absence of dedicated lanes for AVs in the case study area, understanding their behaviour in interactions with other road users is imperative. Camera recordings were used for data enrichment as they covered a limited section of the route for a limited period of time. Due to the wide-angle lenses, videos were affected by radial distortion, which causes straight lines to appear slightly curved. To address this, camera calibration was performed to estimate the intrinsic camera matrix and lens distortion coefficients, which were then used to correct radial distortion in the recorded frames. This allowed geometric correction of the frames and improved the accuracy of road user positioning in the scene.

After distortion correction, object detection and tracking was applied to extract and follow road users within the video frames, enabling subsequent interaction analysis (Figure 4). For object detection, YOLOv7 [38] was utilized. Since autonomous shuttle was not included in the default detection configuration, a customized YOLOv7 model was trained using a dataset collected for this study, containing labelled instances of autonomous shuttle which enabled their reliable detection alongside conventional vehicles. Object tracking was then performed using SORT [39], a tracking-by-detection algorithm based on Kalman filtering and the Hungarian assignment method. Each detected object was assigned a unique identifier that remained consistent across frames, ensuring consistency in tracking and trajectory construction. This step was crucial in detailed examination of movement patterns and interactions between objects within the dynamic traffic environment.

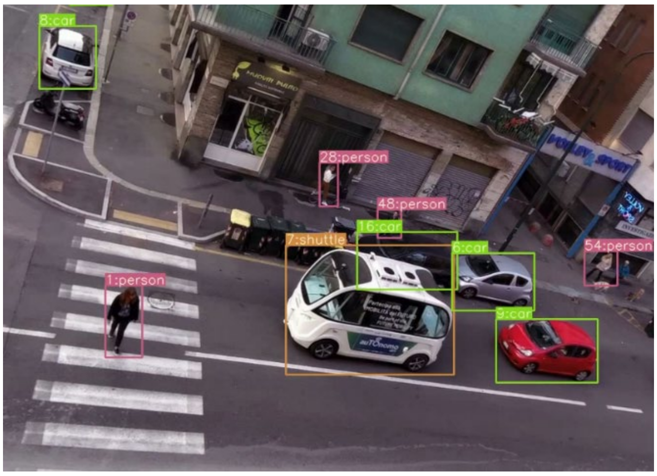

Figure 4. Detection and tracking results on a video frame

In the following, conflict analysis was carried out using the spatial-temporal trajectories extracted from the detection and tracking step. For each pairwise encounter between a pedestrian and a vehicle — whether conventional or autonomous — two surrogate safety measures were computed: Time-to-Collision (TTC), which measures the time remaining before a potential collision, and Post-Encroachment Time (PET), which captures the time gap between a vehicle and a pedestrian passing through the same spatial point. The resulting distributions were then compared across vehicle types using the Kolmogorov–Smirnov (K-S) test to determine whether statistically significant differences exist in how each vehicle type interacts with pedestrians at the unsignalized crosswalk. Beyond conflict analysis, the trajectory data also served a second purpose: speed profiles were extracted by vehicle type and used to inform the parameterisation of car-following behaviour in the microsimulation phase, ensuring that the distinct driving patterns of the shuttle and conventional vehicles were reflected in the SUMO environment.

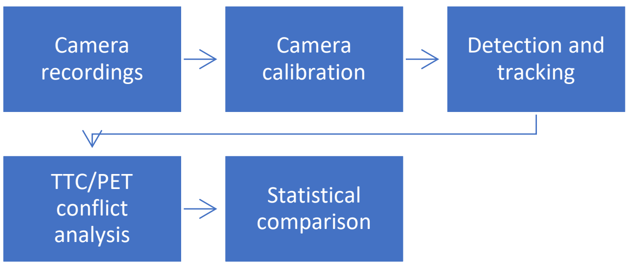

Figure 5. Camera-based analysis workflow

4.3. Traffic microsimulation

After data collection and analysis, the traffic microsimulation phase was developed to replicate traffic conditions in a defined urban zone and to support future evaluation of the impact of AV introduction. Traffic microsimulation, rather than macro or meso-simulation, was used to simulate individual vehicle movements and interactions.

In this study, the open-source SUMO (Simulation of Urban Mobility) simulator [24] was selected, due to its active user and developer community, and compatibility with Python through the TraCI (Traffic Control Interface) [40] API. TraCI allows real-time interaction with the running simulation and customization of various traffic scenarios, providing more flexibility. This approach enables a wide range of applications, from testing traffic management strategies to developing autonomous vehicle algorithms, making SUMO particularly effective compared to other simulators. By incorporating AV telemetry data into SUMO, realistic traffic conditions can be reproduced and evaluated in mixed urban traffic where both conventional and autonomous vehicles exist. To represent the behavioural differences between autonomous and conventional vehicles, derived from the data analysis and modelling phase, distinct car-following and lane-changing models were defined. Human drivers were modelled with standard parameters reflecting visual perception and variable reaction times, while AV parameters were adjusted to reflect sensor-based decision-making.

High-Performance Computing was used to handle the computational demands of running large-network simulations with real-time trajectory injection, accelerating execution time compared with standard hardware. While this significantly reduced processing time, the same simulations could still be performed without HPC, with longer runtimes depending on the scenario size and duration.

The following steps were carried out in SUMO to simulate the selected study area in Turin (Figure 6).

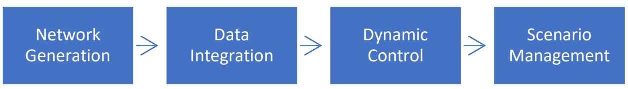

Figure 6. Overview of SUMO simulation workflow

(i) Network generation and traffic demand: e simulation network was generated using OpenStreetMap (OSM) data and refined to accurately reflect the real road infrastructure, including lane configurations, traffic signals, and speed limits for each road segment. An origin–destination (OD) matrix representing passenger demand within the study area was used to generate baseline traffic demand. The OD matrix was converted into individual vehicle routes using SUMO built-in routing tools, which were then used to initialise traffic flows in the simulation environment [24].

(ii) Data integration: Recorded AV telemetry data and speed profiles for both autonomous and conventional vehicles were integrated into the simulation environment using TraCI. Vehicle trajectories were assigned unique identifiers and mapped onto the simulation network. Coordinate transformations between geographic (latitude/longitude) and SUMO's Cartesian coordinate system were performed using built-in functions to ensure spatial consistency. Vehicle classes and behavioural parameters (e.g., maximum speed, acceleration, and physical characteristics) were defined in XML configuration files and incorporated into the simulation setup. In addition, pedestrian and public transport flows were defined to represent mixed traffic conditions.

(iii) Dynamic control: Vehicle trajectories were replayed in SUMO using TraCI. The $MoveToXY$ function was used to update vehicle positions according to recorded trajectory points, ensuring alignment with the simulation network. Between successive update steps, vehicle movement was governed by SUMO's internal mobility models, while external trajectory inputs were applied at each simulation timestep. This procedure enabled the temporal reconstruction of observed vehicle movements within the simulated environment.

(iv) Scenario management and performance monitoring: Simulation scenarios and control logic were implemented using TraCI during runtime. This included the activation of predefined interventions, specification of their spatial and temporal conditions, and management of lane permissions, intersection behaviour, and routing adjustments where applicable. Throughout the simulation, key performance indicators (KPIs) were continuously recorded to evaluate system behaviour under different traffic configurations and operational scenarios.

5. Results

This section presents the main results from each phase of the methodology, based on the collected and processed data.

The first data gathering phase has proven to be a valuable asset for the proposed study. The dataset, encompassing diverse data sources offers a comprehensive snapshot of real-world scenarios which serves as a foundation for subsequent steps of the framework. On-board sensors on the two autonomous shuttles recorded operational data continuously throughout the three-month period (August-October 2022), from which a dataset of approximately 1,157,000 records was derived at a sampling rate of 1 Hz. Regarding the infrastructure data, signal phase and timing data from around 14 signal groups along the route—transmitted to the shuttle via SPATEM during the trials—were recorded onboard and used to calibrate traffic signal timings in the SUMO model, aligning the simulation with real-world conditions experienced by the shuttle. Camera recordings totalled roughly 10 hours of footage across multiple days at the unsignalized crosswalk, capturing naturalistic interactions between pedestrians, conventional vehicles, and the shuttle.

As regards the data analysis and modelling phase, the anomaly detection analysis successfully identified deviations in AV speed that were flagged as potential anomalies. These events may correspond to unexpected situations, such as sudden interactions with obstacles or irregular vehicle responses. The use of MAD enabled the identification of both pronounced and subtle deviations by adjusting detection thresholds, providing flexibility in highlighting events of interest.

The dashboard is deployed via the Power BI Service and is accessible to authorised stakeholders through a standard web URL, with no local installation or data-analysis expertise required (Figure 7). The interface is specifically designed for non-technical end users: interactive filters — by date, operational zone, and key variables — allow to explore key performance metrics through self-explanatory charts, maps, and summary cards. This makes the tool directly usable by researchers, policymakers, and transport operators, a feature that is often missing in technically oriented AV assessment frameworks.

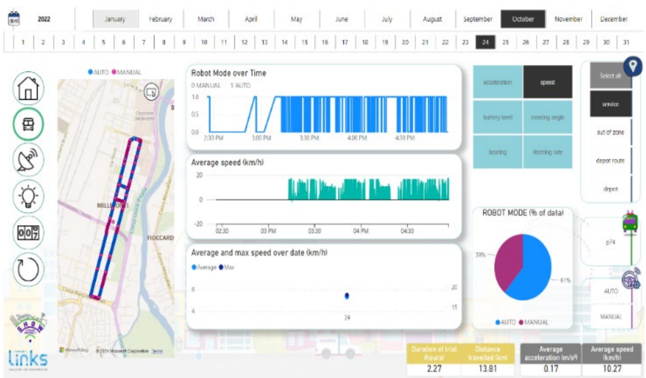

Figure 7. Snapshot of Power BI dashboard

The dashboard consists of four main pages, the content of which is described below.

- Vehicle Dynamics: This page presents crucial metrics such as speed, steering rate and angle, bearing, and battery level. These features are analysed over time to understand shuttle dynamics. It also reports the share of time spent in different operational modes, including autonomous operation.
- Operational Status: This page visualises various sensor and system status such as GNSS correction, door state, and battery status. Temperature data (engine, indoor, outdoor) is also displayed. This page helps in monitoring and assessing the shuttle's operational conditions.
- Lighting and Signalling Systems: This page focuses on the signalling components, including the state of blinkers, brake lights, and reverse lights. The analysis here aims to ensure that these safety components are functioning correctly, providing a statistical representation of their operation.
- Fleet Overview: This page aggregates data across the two shuttles, showing key metrics like distance travelled and trip duration in both manual and autonomous modes. It allows stakeholders to evaluate performance and compliance with legislative requirements.

Regarding the camera-based analysis, object detection and tracking methods, using customized YOLOv7 and SORT, have been employed to identify and track vehicles, pedestrians, and shuttles within the environment.

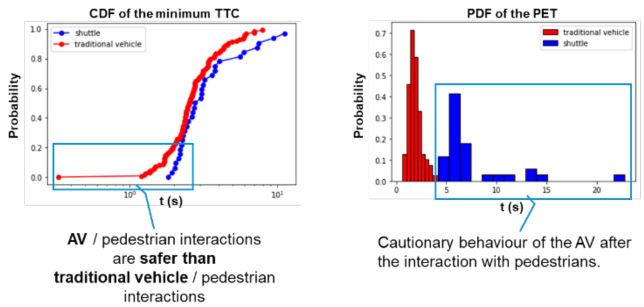

Figure 8. Cumulative Distribution Function of minimum TTC (left) and Probability Density Function of PET (right): autonomous shuttle (blue) vs. traditional vehicles (red)

To study the interactions in real traffic, camera-based data from an unsignalized crosswalk zone were used, where specifically pairwise interactions of pedestrians with both conventional and autonomous vehicles were observed. The analysis focused on two surrogate safety measures (SSMs) to quantify conflict severity: Time-to-Collision (TTC) and Post-Encroachment Time (PET). The analysis comprised a total of 33 pedestrian-AV interactions and 135 pedestrian- conventional vehicle interactions, and the resulting distributions were compared using the Kolmogorov–Smirnov test (Figure 8).

Regarding TTC, based on the cumulative distributions for the two vehicle types, no statistically significant difference was found. However, the tail of the distribution showed a notable divergence. In particular, the distribution for conventional vehicles shows a faster increase at very low TTC values, meaning that human drivers more frequently found themselves in near-miss situations with pedestrians. The shuttle, by contrast, almost never reached those critical thresholds. Regarding PET, the histogram indicates that most interactions with conventional vehicles occur at low PET values, with the majority below 10 seconds. The shuttle's distribution, on the other hand, is visibly shifted to the right and spread across a wider range, extending well beyond 20 seconds in several cases. This difference was statistically significant and reflects the shuttle's tendency to stop earlier and resume movement more gradually, leaving a substantially larger temporal gap between its passage and the pedestrian's.

What emerges from these two measures points to a clear trade-off. While TTC did not reveal major differences, PET demonstrated that AVs adopted more conservative driving behaviour such as braking earlier and waiting longer, which improved pedestrian safety. However, this same caution can increase the time vehicles occupy road space, potentially affecting traffic flow and network capacity in mixed-traffic conditions. The results underline the importance of considering vehicle type to balance safety benefits with operational impacts when evaluating AV deployment in urban environments.

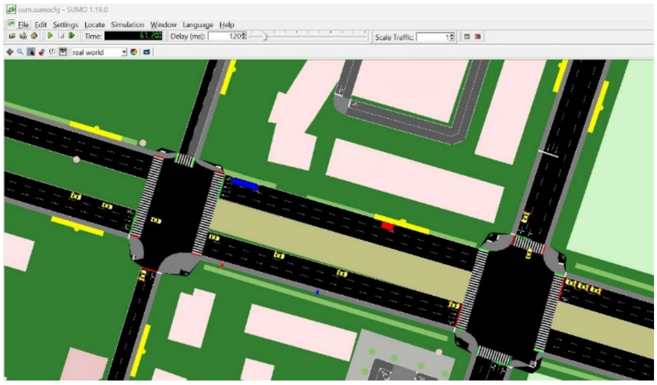

Figure 9. Snapshot of SUMO simulation in the study area: AV (red), conventional vehicles (yellow), public transport (blue)

The key outcome of the traffic simulation phase is the development of a 'baseline environment' to test further traffic management strategies in the dedicated area. The shuttle's real-world trajectories, derived from the 1 Hz sensor data, were dynamically injected into the simulation via TraCI and synchronised with OD-based conventional traffic flows and public transport routes operating on the same network. Figure 9 shows the resulting environment with all three traffic layers coexisting. Traffic signal timing in the simulation was configured using the infrastructure data collected during the trials, ensuring consistency between the simulated conditions and those actually encountered by the shuttle. The incorporation of real-world mobility data into the simulation ensures high degree of accuracy, enabling it to serve as a virtual testing platform for evaluating AV deployment scenarios under realistic mixed-traffic conditions without additional on-road experimentation. The baseline scenario provides a reference for systematically comparing future interventions, such as alternative routing strategies, signal priority schemes, dedicated lane configurations, and varying AV penetration levels. Initial tests confirmed that the simulation runs efficiently on HPC infrastructure, with execution times reduced to minutes for scenarios that would otherwise require hours on standard hardware.

6. Conclusions

This paper presented a data-driven framework for assessing the integration of autonomous vehicles into mixed urban traffic, developed and tested using three months of real-world data from autonomous shuttle trials conducted in Turin as part of the European SHOW project.

Five contributions distinguish this work. First, the framework is grounded in actual operational data rather than synthetic or estimated parameters. The AV telemetry dataset, infrastructure signal timing, road network information, and external video recordings collectively provide an empirical basis that most existing simulation studies lack. Second, the video analysis pipeline — combining YOLOv7-based detection, SORT tracking, and extraction of surrogate safety measures — demonstrated that the shuttle produces significantly higher PET values and fewer near-miss TTC episodes than conventional vehicles when interacting with pedestrians at an unsignalized crosswalk. These findings offer concrete, field-validated evidence that AVs adopt more conservative behaviour in conflict situations, although the same caution raises questions about potential impacts on network capacity that will need to be examined as deployments scale up. Third, the anomaly detection component, based on the Median Absolute Deviation method with adjustable thresholds, proved effective at flagging unexpected operational events in the sensor data. Fourth, the Power BI dashboard provides a non-technical interface through which city planners, transport operators, and regulators can explore AV performance data without requiring programming expertise — an aspect that is often overlooked in technically oriented frameworks. Finally, the microsimulation phase produced a functioning baseline scenario in SUMO, where the shuttle's recorded trajectories were dynamically replayed via TraCI alongside OD-based conventional traffic and public transport flows, creating a representative mixed-traffic environment against which future interventions can be compared.

Several limitations should be acknowledged. High-resolution microsimulations with real-time trajectory injection are computationally demanding; longer processing times can be expected under less computationally intensive setups. The dataset, while rich, is specific to Turin's urban layout and CCAM ecosystem. Cities with less mature traffic infrastructure or limited access to AV trial data would require adaptations to the data gathering process; in such contexts, kinematic and safety parameters could be sourced from published research datasets, traffic simulators, or pilot studies in similar traffic contexts. Signal timing data are typically retrievable from municipal Traffic Control Centres through standard protocols (SPATEM, UTMC, or direct XML exports), and OD matrices can be reconstructed from census or mobile-phone-based origin-destination studies. The transferability of the quantitative findings — particularly the surrogate safety measures — would need to be re-validated in the new environment. The framework itself is agnostic to the source of these inputs, which is what makes it transferable. What would require re-establishment in a new city is not the framework but the empirical calibration of its inputs against local conditions. The safety analysis, though statistically significant, was based on a relatively small sample at a single crosswalk — and broader validation across multiple locations and traffic contexts would strengthen the generalisability of the findings.

The current microsimulation relies on standard car-following and lane-changing models for conventional vehicles, which do not fully capture the variability and unpredictability of human driving behaviour. While AVs are typically distinguished from human-driven vehicles through operational parameters such as reaction time, minimum gaps, and acceleration profiles, these simplified representations do not account for the broader cognitive and perceptual factors influencing human decision-making. A promising direction for future research is the integration of more advanced behavioural models that better represent driver heterogeneity, situational awareness, and adaptive responses to mixed-traffic conditions. In addition, insights from the video-based analysis of pedestrian–vehicle interactions presented in Section 4.2 can support the development of more realistic pedestrian models, particularly in terms of crossing behaviour and gap acceptance in the presence of autonomous vehicles. Specifically, the extracted TTC and PET distributions can be used as calibration targets within SUMO to better represent pedestrian decision-making processes.

Beyond model-level improvements, future work will focus on data expansion and transferability. Ongoing projects, including ToMove [4], are expected to provide additional CCAM datasets from new trial campaigns, enabling cross-city comparisons and enhancing the generalisability of the proposed framework.

Acknowledgements

The authors would like to thank the collaborators of the Politecnico di Torino for their precious support: Prof. Silvia Chiusano, Prof. Marco Bassani, Prof. Francesco Deflorio, and Andrea Avignone. This paper is partially supported and funded by European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 875530, project SHOW (SHared automation Operating models for Worldwide adoption), and by Horizon Europe Innovation Actions Programme under Grant Agreement No 101076791, project IN2CCAM (Enhancing Integration and Interoperability of CCAM eco-system).

References

[1] European Telecommunications Standards Institute (ETSI). Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions. Sophia Antipolis: ETSI; 2009 Jun. Report No.: ETSI TR 102 638 V1.1.1.

[2] SHOW Project [Internet]. [cited 2026 Apr 19]. Available from: https://show-project.eu/

[3] IN2CCAM Project [Internet]. [cited 2026 Apr 19]. Available from: https://in2ccam.eu/

[4] To Move - Torino City Lab [Internet]. [cited 2026 Apr 19]. Available from: https://torinocitylab.it/to-move/

[5] Anderson JM, Kalra N, Stanley KD, Sorensen P, Samaras C, Oluwatola OA. Autonomous vehicle technology: a guide for policymakers. Santa Monica (CA): RAND Corporation; 2016 Mar.

[6] Guériau M, Dusparic I. Quantifying the impact of connected and autonomous vehicles on traffic efficiency and safety in mixed traffic. In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC); 2020 Sep 20-23; Rhodes, Greece. IEEE; 2020. p. 1-8. doi: 10.1109/ITSC45102.2020.9294174.

[7] Yu H, Tak S, Park M, Yeo H. Impact of autonomous-vehicle-only lanes in mixed traffic conditions. Transp Res Rec. 2019 Sep;2673(9):430-439. doi: 10.1177/0361198119847475.

[8] Avignone A, Bassani M, Borgogno B, Caroleo B, Chiusano S, Princiotto F. Evaluating unsignalized crosswalk safety in the age of autonomous vehicles. Comput Ind. 2025 May;167:104259. doi: 10.1016/j.compind.2025.104259.

[9] Nemoto EH, Issaoui R, Korbee D, Jaroudi I, Fournier G. How to measure the impacts of shared automated electric vehicles on urban mobility. Transp Res Part D Transp Environ. 2021 Apr;93:102766. doi: 10.1016/j.trd.2021.102766.

[10] Anund A, Ludovic R, Caroleo B, Hardestam H, Dahlman A, Skogsmo I, et al. Lessons learned from setting up a demonstration site with autonomous shuttle operation – based on experience from three cities in Europe. J Urban Mobil. 2022;2:100021. doi: 10.1016/j.urbmob.2022.100021.

[11] Thorhauge M, Fjendbo Jensen A, Rich J. Effects of autonomous first- and last mile transport in the transport chain. Transp Res Interdiscip Perspect. 2022 Sep;15:100623. doi: 10.1016/j.trip.2022.100623.

[12] Kušić K, Schumann R, Ivanjko E. A digital twin in transportation: Real-time synergy of traffic data streams and simulation for virtualizing motorway dynamics. Adv Eng Inform. 2023;55:101858. doi: 10.1016/j.aei.2022.101858.

[13] Bagabaldo AR, Hackl J. Digital Twins for Intelligent Intersections: A Literature Review. arXiv. 2025 Oct 14. arXiv:2510.05374. Available from: https://arxiv.org/abs/2510.05374.

[14] Xu G, Chen J, Wang Z, et al. Enhancing traffic safety analysis with digital twin technology: integrating vehicle dynamics and environmental factors into microscopic traffic simulation. Sci Rep. 2025;15:44404. doi: 10.1038/s41598-025-31930-1.

[15] Xu H, Omitaomu F, Sabri S, Zlatanova S, Li X, Song Y. Leveraging generative AI for urban digital twins: a scoping review on the autonomous generation of urban data, scenarios, designs, and 3D city models for smart city advancement. Urban Informatics. 2024;3:29. doi: 10.1007/s44212-024-00060-w.

[16] Azfar T, Li J, Yu H, Cheu RL, Lv Y, Ke R. Deep learning-based computer vision methods for complex traffic environments perception: a review. Data Sci Transp. 2024 Jan 8;6(1):1. doi: 10.1007/s42421-023-00086-7.

[17] Helbing D, Sánchez-Vaquerizo JA. Digital twins: potentials, ethical issues and limitations. In: Handbook on

the politics and governance of big data and artificial intelligence. Edward Elgar Publishing; 2023. p. 64-104. doi: 10.4337/9781800887374.00013.
[18] Li D, Okhrin O. Ethics-Aware Safe Reinforcement Learning for Rare-Event Risk Control in Interactive Urban Driving. arXiv. 2025 Aug 22. arXiv:2508.14926.
[19] Suresh Babu CV, Akkash Anniyappa CS, Raut A. Toward seamless mobility: Integrating connected and autonomous vehicles in smart cities through digital twins. In: Iyer S, Nayyar A, Paul A, Naved M, editors. Digital Twins for Smart Cities and Villages. Elsevier; 2025. p. 169-187. doi: 10.1016/B978-0-443-28884-5.00008-7.
[20] Lopez PA, Wiessner E, Behrisch M, Bieker-Walz L, Erdmann J, Flötteröd YP, et al. Microscopic traffic simulation using SUMO. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC); 2018 Nov 4-7; Maui, HI. IEEE; 2018. p. 2575-2582. doi: 10.1109/ITSC.2018.8569938.
[21] Raju N, Farah H. Evolution of traffic microsimulation and its use for modeling connected and automated vehicles. J Adv Transp. 2021;2021:2444363. doi: 10.1155/2021/2444363.
[22] Lu Q, Tettamanti T, Hörcher D, Varga I. The impact of autonomous vehicles on urban traffic network capacity: an experimental analysis by microscopic traffic simulation. Transp Lett. 2020;12(8):540-549. doi: 10.1080/19427867.2019.1662561.
[23] European Telecommunications Standards Institute (ETSI). Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions. Sophia Antipolis: ETSI; 2009 Jun. Report No.: ETSI TR 102 638 V1.1.1.
[24] SUMO Documentation [Internet]. [cited 2026 Apr 19]. Available from: https://sumo.dlr.de/docs/index.html
[25] Parkin J, Crawford F, Flower J, Alford C, Morgan P, Parkhurst G. Cyclist and pedestrian trust in automated vehicles: an on-road and simulator trial. Int J Sustain Transp. 2023;17(7):762-774. doi: 10.1080/15568318.2022.2093147.
[26] Papadoulis A, Quddus M, Imprialou M. Evaluating the safety impact of connected and autonomous vehicles on motorways. Accid Anal Prev. 2019 Mar;124:12-22. doi: 10.1016/j.aap.2018.12.019.
[27] Xie J, Qin Y, Zhang Y, Chen T, Wang B, Zhang Q, et al. Towards human-like automated vehicles: review and perspectives on behavioural decision making and intelligent motion planning. Transp Saf Environ. 2025;7(1):tdae005. doi: 10.1093/tse/tdae005.
[28] Chand C, Jashami H, Wang H, Hurwitz D. Evaluation of the human interaction with automated vehicles on highways. Traffic Saf Res. 2025;9:e000078. doi: 10.55329/xwwy8052.
[29] Chauhan V, Anubhav, Chang CM, Su X, Nakazato J, Javanmardi E, et al. Towards the future of pedestrian–AV interaction: Human perception vs. LLM insights on Smart Pole Interaction Unit in shared spaces. Int J Hum Comput Stud. 2025;205:103628. doi: 10.1016/j.ijhcs.2025.103628.
[30] Sadeghi J, Caroleo B, Botta C, Arnone M. Evaluating Traffic Control Strategies for Autonomous Shuttle in Different AV Penetration, Using SUMO Traffic Simulation. In: Kocian A, Milazzo P, Henriques Martins AL, Nanni M, Pappalardo L, editors. Intelligent Transport Systems. Cham: Springer Nature Switzerland; 2025. p. 332-343.
[31] Sadeghi J, Botta C, Caroleo B, Arnone M. Vehicle Connectivity and Dynamic Traffic Response to Unplanned Urban Events. Urban Sci. 2025;9:409. doi: 10.3390/urbansci9100409.
[32] Shi Y, Wu T, Guo T, Huo J, Gu Z, Dai Y, et al. Data-driven traffic simulation optimization considering heterogeneous driving styles. Commun Transp Res. 2025;5:100181. doi: 10.1016/j.commtr.2025.100181.
[33] Roy S, Hasan S, Nahmias-Biran B. Autonomous Shuttle Service in Sub-Urban Mixed Traffic Conditions: Microscopic Simulation-Based Impact Assessment. Smart Cities. 2026;9(2):29. doi: 10.3390/smartcities9020029.
[34] Shi Y, Liu Z, Wang Z, Ye J, Tong W, Liu Z. An integrated traffic and vehicle co-simulation testing framework for connected and autonomous vehicles. IEEE Intell Transp Syst Mag. 2022 Nov;14(6):26-40. doi: 10.1109/MITS.2022.3188566.
[35] Ventura R, Roussou S, Ziakopoulos A, Barabino B, Yannis G. Using computer vision and street-level videos for pedestrian-vehicle tracking and behaviour analysis. Transp Res Interdiscip Perspect. 2025;30:101366. doi: 10.1016/j.trip.2025.101366.
[36] NAVYA. Self-Driving Shuttle for Passenger Transportation [Internet]. [cited 2026 Apr 19]. Available from: https://www.navya.tech/en/solutions/moving-people/self-driving-shuttle-for-passenger-transportation/
[37] Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology. 2013 Jul;49(4):764–766.
[38] Wang CY, Bochkovskiy A, Liao HYM. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv. 2022 Jul 6. arXiv:2207.02696.
[39] Bewley A, Ge Z, Ott L, Ramos F, Upcroft B. Simple Online and Realtime Tracking. In: 2016 IEEE International Conference on Image Processing (ICIP). 2016 Sep 25–28; Phoenix, AZ, USA. p. 3464–3468.
[40] Eclipse Foundation. TraCI: Traffic Control Interface for SUMO [Internet]. [cited 2026 Apr 21]. Available from: https://sumo.dlr.de/docs/TraCI.html