Alicante-Murcia Freeway Scenario: A High-Accuracy and Large-Scale Traffic Simulation Scenario generated using a Novel Traffic Demand Calibration Method in SUMO

Juan Jesús Gonzalez Delicado, Javier Gozalvez, (Senior Member, IEEE), Jesús Mena-Oreja, Miguel Sepulcre, (Senior Member, IEEE) and Baldomero Coll-Perales

Uwicore Lab., Universidad Miguel Hernández de Elche (UMH), Avda. Universidad s/n, 03202, Elche, Alicante, Spain

Corresponding author: Javier Gozalvez (e-mail: j.gozalvez@umh.es).

This work was supported in part by the Spanish Dirección General de Tráfico under Grant SPIP2017-02204, in part by the Spanish Ministry of Science and Innovation (MCI), the State Research Agency (AEI), and the European Regional Development Fund (FEDER), under Projects TEC2017-88612-R and PID2020-115576RB-I00, and in part by Generalitat Valenciana under Project GV/2021/044.

ABSTRACT The design, testing and optimization of Vehicle to Everything (V2X), connected and automated driving and Intelligent Transportation Systems (ITS) and technologies requires mobility traces and traffic simulation scenarios that can faithfully characterize the vehicular mobility at the macroscopic and microscopic levels under large-scale and complex scenarios. The generation of accurate scenarios and synthetic traces requires a precise modelling approach, and the possibility to validate them against real-world measurements that are generally not available for large-scale scenarios. This limits the open availability of realistic and large-scale traffic simulation scenarios. The purpose of this paper is to present a large-scale and high-accuracy traffic simulation scenario. The scenario has been implemented over the open-source SUMO traffic simulator and is openly released to the community. The scenario accurately models the traffic flow, the traffic speed and the road’s occupancy for 9 full days of traffic over a 97 km freeway section. The scenario models mixed traffic with light and heavy vehicles. The simulation scenario has been validated using a unique dataset provided by the Spanish road authority and a novel learning-based and iterative traffic demand calibration technique for SUMO. This technique, referred to as Clone Feedback, is proposed for the first time in this paper and does not require a pre-calibration to generate realistic traffic demand. Clone Feedback can generate calibrated mixed traffic (light and heavy vehicles) using as input only traffic flow measurements. The results obtained show that Clone Feedback outperforms two reference techniques for calibrating the traffic demand in SUMO.

INDEX TERMS Traffic simulation, freeway, large-scale, SUMO, V2X, CAVs, ITS, traffic demand, calibration, microscopic, macroscopic, traffic flow, traffic speed, road occupancy, Cadys.

I. INTRODUCTION Simulations are commonly utilized to design, test and optimize connected and automated driving, Vehicle to Everything (V2X) communications, and Intelligent Transportation Systems (ITS) solutions. Simulations represent a cost-efficient testing solution [1] under large-scale and complex scenarios compared to analytical evaluations or Field Operational Tests (FOTs). Analytical models generally require important assumptions or simplifications for tractability that hinder their capacity to accurately represent complex and large-scale scenarios. FOTs [2] and testing in proving grounds [3] provide critical implementation and operational feedback, but their impact is usually limited in time and space.

The validity of simulations greatly depends on their capacity to accurately represent the reality, and this accuracy has a significant impact on the outcome of the studies. For example, previous studies demonstrated that the mobility of vehicles has a significant impact on the topology of vehicular networks, and hence on the operation and performance of
vehicular networking protocols [4]. An adequate design, testing and optimization of V2X networks requires then mobility traces and traffic simulation scenarios that can faithfully characterize vehicular mobility at the macroscopic and microscopic levels [5]. Generating accurate traces and scenarios requires a precise modelling of the road topology and configuration, and a careful validation against real-world measurements that are hard to retrieve and are generally not available. This limits the open availability of large-scale and accurate mobility traces and traffic simulation scenarios that are critical for designing, testing and optimizing V2X, connected and automated driving and ITS solutions.

Several microscopic-level traffic simulation platforms are available in the community. Commercial platforms such as VISSIM or PARAMICS can provide high accuracy levels at an economic cost and usually long learning processes. Open-source traffic simulation platforms are more extensively used in research as they are usually simpler to utilize and to expand with new functionalities based on research needs and developments. In addition, existing open-source traffic simulation platforms now include microscopic models of the driver’s behavior and can simulate accurately the movement of individual vehicles. This study is hence focused on open-source platforms, and in particular on SUMO (Simulation of Urban Mobility), a state-of-the-art and highly portable open-source microscopic road traffic simulation platform [6]. SUMO is a reference platform for V2X and CAVs research, and can be interconnected to network simulators such as ns3 and OMNet++. The connection and interaction between SUMO and network simulators is, for example, available in the iTETRIS [7] and Veins [8] frameworks that provide advanced tools for designing and evaluating V2X and CAVs protocols.

SUMO has been used to create traffic simulation scenarios that can generate synthetic mobility traces. These scenarios must accurately model the mobility at the microscopic and macroscopic levels since this will impact the distribution of the road traffic in the scenario, and hence affect the topology of vehicular networks and the interactions between CAVs, among others. However, there is only a limited number of mobility scenarios openly available in the community that can accurately represent real-world traffic flows. This includes the Luxemboug SUMO Traffic Scenario (LuST) [9], the Bologna scenario [10], the TAPAS Cologne scenario [5], and the Monaco SUMO Traffic Scenario (MoST) [11]. These scenarios clearly advanced the state-of-the-art and are widely utilized by the community. However, only a few of them have been validated using real traffic measurements, and the validation has focused mainly on the traffic flow. Accurately reproducing real traffic conditions in simulations requires not only matching the traffic flow but also the traffic speed and the road’s occupancy. In addition, most scenarios model the traffic for a day or a few hours of traffic, and this impacts their usage for Artificial Intelligence (AI) related studies [12][13]. AI studies usually require large and varied datasets for the training, validation and testing phases of deep learning algorithms. While the training phase uses the data to fit the model (learn), the validation phase evaluates the model fit and tunes the model’s hyperparameters, and the testing phase evaluates the model once it is completely trained. The amount of data required for AI studies depends on different factors, such as the complexity of the problem and the complexity of the learning algorithm. In this regard, it is important to note that the amount of data generated with our scenario is significantly higher than the amount data generated with most of the existing scenarios that usually limit the traffic to a few hours or a maximum of 24 hours. Our scenario accurately models the traffic for 9 full days.

In this context, this paper advances the state-of-the-art with the creation and release of a high-accuracy and large-scale traffic SUMO simulation scenario. The scenario realistically reproduces the traffic flow, the traffic speed and the road’s occupancy in mixed traffic conditions (light and heavy vehicles) over 97 km of the A-7 freeway between the cities of Alicante and Murcia in Spain for 9 full days. This makes the Alicante-Murcia scenario one of the largest freeway traffic simulation scenarios available in SUMO. The scenario accurately models the traffic for 9 full days, and hence provides an ideal testing environment to develop AI-based solutions, e.g. for road traffic characterization and prediction [14]. The scenario is openly released to the community and can be downloaded from [15]. The scenario has been built from an extensive dataset provided by the Spanish road authority. The dataset included real traffic flow, traffic speed and road’s occupancy measurements collected during 12 years by 99 induction loops deployed along the scenario, including at some of the 36 on-ramps and 34 off-ramps present in the scenario. The traffic flow is defined as the number of vehicles passing a reference point (e.g. the location of an induction loop or any other type of traffic detector) per unit of time, and is typically measured in vehicles per hour. The road occupancy is the percentage of road surface covered by vehicles, and is computed as the number of vehicles per unit length of the road multiplied by their average length. We selected a total of 9 days where most induction loops were active to create a large but also highly accurate traffic simulation scenario that combines accurate macroscopic traffic flows with fine-grained microscopic vehicular mobility modeling using car-following models.

The scenario has been calibrated and extenively validated using the measurements provided by the induction loops and a new traffic demand calibration technique for SUMO. This technique, referred to as Clone Feedback, is proposed in this paper and has been designed to generate realistic traffic demands. The traffic demand is the set of vehicle trips (or routes) from origin to destination that that are produced in a road during a specific period of time. Generating realistic traffic demands means that the routes followed by vehicles accurately mimic the real mobility of vehicles (e.g. as captured by the measurements from induction loops or any other type
of traffic detector). The proposed calibration technique uses the tool Cadyts available in SUMO, but overcomes its major limitation since it does not require access to an initial accurate traffic demand to start the calibration process. In fact, Clone Feedback does not require a pre-calibration to produce realistic traffic demands, and can generate calibrated mixed traffic (light and heavy vehicles) using as input only the traffic flows measured by the road detectors or sensors.

The rest of the paper is organized as follows. Section II reviews the current state of the art related to the generation of realistic SUMO-based traffic simulation scenarios and traffic demand calibration techniques. The Alicante-Murcia freeway scenario is presented in Section III, and Section IV describes the modelling approach utilized in SUMO to create the traffic simulation scenario. Section V introduces the traffic demand calibration technique available in SUMO using Cadyts. Section VI and Section VII present the proposed traffic demand calibration technique for SUMO, Clone Feedback, that is here presented and that has been used to calibrate the Alicante-Murcia freeway traffic simulation scenario. Section VIII validates the scenario and compares the accuracy achieved with the proposed calibration technique compared to the current state of the art traffic demand calibration techniques in SUMO. Finally, Section IX summarizes the main contributions of this study.

II. RELATED WORK

Several large-scale traffic SUMO simulation scenarios are currently available. One of the first was the TAPAS Cologne scenario that models the city of Cologne and surrounding areas (400 Km²), including urban roads, main highways and freeways [5]. The scenario provides two mobility traces. One trace corresponds to 2 hours of traffic in the peak hour and the second one models traffic for a 24h period. The realism of the scenario is only qualitatively analyzed, and to the author’s knowledge, it has not been validated against real traffic data. This is also the case for the Monaco SUMO Traffic Scenario (MoST) [11] that covers an area of 73 km² in the city of Monaco (including urban roads, main highways and freeways). MoST simulates the morning peak traffic hour (using a normal distribution), and models multimodal traffic. The Luxembourg SUMO Traffic Scenario (LuST) was validated with real traffic data [9]. This scenario models the city of Luxembourg and surroundings areas (156 Km²) including urban roads, main highways and freeways. The scenario provides a mobility trace representing 24h of traffic. The authors validated the trace using 6 million Floating Car Data (FCD) samples, and they demonstrate that the scenario is capable of realistically reproducing the speed of the traffic at the macroscopic level. However, the study does not validate the scenario in terms of the traffic flow and road occupancy or at the microscopic level. The urban Bologna scenario was validated against real traffic flow measurements in [10]. This scenario was developed as part of the European iTETRIS project [7] and models an area of 20 Km² within the city of Bologna. The scenario models one hour of morning peak traffic including passenger vehicles and public bus service. The authors use traffic flow measurements from induction loops to validate the model. However, the study validates the traffic flow but not the traffic speed and the road’s occupancy. Existing large-scale traffic SUMO simulation scenarios have been generally validated against a single traffic variable. However, an accurate representation of the traffic requires a thorough validation with the three fundamental traffic variables: traffic flow, traffic speed and road occupancy. In addition, traffic simulation scenarios should realistically represent the microscopic behavior and interaction of vehicles as well as the traffic at the macroscopic level in large-scale (both in space and time) scenarios that accurately account for the road layout and mixed traffic types [5]. Generating such high-accuracy and large-scale traffic simulation scenarios is the objective of this study. The Alicante-Murcia SUMO freeway traffic simulation scenario that is here presented models the traffic over a large freeway section and over several days. The scenario provides high microscopic-level accuracy, and has been created and validated using realistic macroscopic traffic data sources.

Traffic simulation scenarios are generally created using a calibration process. This process generates the traffic demand for the scenario and fine-tunes the parameters of the traffic simulation model so that the simulated traffic resembles the real traffic. For example, [16][17] fine-tune parameters such as the maximum acceleration/deceleration, the reaction time or the probability to follow the “keep right” rule of the SUMO car-following and lane-changing models. To generate the traffic demand, it is necessary to produce a list of vehicle trips from origin to destination that mimic the real mobility of vehicles. Several techniques can be used to generate the traffic demand, and they differentiate on the type of data they use. For example, some studies [18][11][9] use demographic information and daily activity patterns. This approach generally produces realistic traffic demands at the macroscopic level but can lose accuracy at the microscopic level. Authors used this approach to generate the LuST scenario in [9]. They used the SUMO package tool ACTIVITYGEN to create mobility traces based on demographic data. Demographic statistics were also used to generate the traffic demand for the calibration of the MoST scenario in [11]. An alternative to the use of demographic data is the use of traffic measurements such as those provided by induction loops [19][20][21]. Several studies use these measurements to compute origin-destination (O/D) matrices. These matrices indicate the quantity of vehicles and their routes within the road network during a period of time. Different methods can be used to compute the matrices. The methods differ on the assumptions made about the traffic, for example, whether they consider congestion conditions [19][21] or not [20], and whether the traffic flow is considered to be constant [21] or variable [19][20] during the simulation time. An accurate calibration and representation of the traffic
requires considering potential congestion conditions and variable traffic flows. To the author’s knowledge, only the tool Cadyts [19] can perform in SUMO a dynamic calibration in congested road networks. This tool is actually used in [22] for calibrating traffic over the city of Hefei. To do so, authors import traffic demand generated in VISUM and adjust it using Cadyts and traffic flow measurements. The measurements are obtained from 21 detectors close to major intersections and aggregated over 15 min periods. SUMO offers other tools to generate traffic demand using traffic measurements, for example, DFROUTER and FLOWROUTER. However, these tools are designed only for uncongested road networks and the accuracy of the resulting calibration is then not always guaranteed. DFROUTER is used for example in [20] for the calibration of an urban scenario in the city of Valencia. The study shows that the total traffic of the city can be calibrated using DFROUTER. However, no results are provided about the accuracy of the calibration at the microscopic level, and the analysis focuses on the traffic flow and does not consider the speed or the road’s occupancy. It should be noted that the techniques utilized in [19], [20] and [21] adjust the simulated traffic flow to the input measurements. These techniques do not directly calibrate the speed or the road’s occupancy, so the capacity to accurately reproduce them depends on the quality of the calibration process and the realism of the scenario that is modeled.

III. ALICANTE-MURCIA FREEWAY SCENARIO
The selected scenario is the A-7 freeway between the cities of Alicante and Murcia in the southeastern Spanish coast. This freeway is part of the Mediterranean corridor included in the Trans European Transport Network (TEN-T). It is used by local, national and international traffic as well as freight transport. The selected scenario is 97 km long and has sections with two or three lanes. The scenario includes 36 on-ramps and 34 off-ramps. Thirteen of these ramps are connections to and from other freeways and the rest serve the surrounding areas. The selected scenario serves three mid-sized cities (Alicante, Murcia and Elche) and an important industrial and touristic zone with a total population of around 2 million people. It also serves the Alicante-Elche airport with over 14 million passengers per year. The selected scenario is a very busy freeway in particular near the city of Murcia and between the cities of Alicante and Elche. Some road segments near Murcia experience an Average Daily Traffic (ADT) higher than 100000 vehicles per day (88.8% light vehicles, 11.2% heavy vehicles, 272 foreign vehicles per day). The ADT between Alicante and Elche can be as high as 83000 vehicles per day (93.7% light vehicles, 6.3% heavy vehicles, 964 foreign vehicles per day) [23]. The speed limit is typically 120 km/h on the mainline although near the main cities it is 100 or 80 km/h. The speed limit in the ramps and acceleration/deceleration lanes ranges between 30 and 100 km/h. The pavement is in good conditions and vehicles can safely drive at the speed limit.

The scenario was chosen due to the large number of induction loops available and their distribution. These sensors are managed by the Spanish road authority (Dirección General de Tráfico). The scenario has 99 induction loops in total: 49 in the mainline and 50 in ramp lanes. This means that 71.4% of the 70 ramps in the scenario are monitored. Accurately monitoring the traffic on the entrance/exist ramps is very important to be able to generate a simulation scenario capable to accurately reproduce the traffic in the freeway (mainline and ramps). The ramp lanes are equipped with single induction loops that provide information about the traffic flow and the lanes’ occupancy. The mainline is equipped with double induction loops that provide information about the traffic flow, the lanes’ occupancy, the average speed of vehicles, the average distance between vehicles, the average length of vehicles, and the vehicle type (light or heavy). The induction loops provide measurements every minute. Measurements for 12 years (between 2006 and 2017) have been provided by the national road authority. We exhaustively analyzed and processed this data to generate the large-scale and high-accuracy freeway traffic scenario presented in this paper. In particular, we identified the period of time that is more suitable to model the traffic in the scenario. This period corresponds to the time where most of the detectors were active. Nine days have been finally selected at the beginning of 2015 (from January 16 to January 24, 2015). Figure 1 depicts the average (over all detectors in the scenario) traffic flows experienced on January 23, 2015 at the mainline. This day has the most accurate traffic records: 88 detectors were active with an availability of 99.31%1. The figure represents the total traffic and the traffic of Light Vehicles (LV) and Heavy Vehicles (HV). Figure 1 shows typical traffic patterns with morning and evening rush hours as well as lunch time and night periods. The selected nine days correspond to typical traffic weekly patterns with higher traffic flows measured during the week (especially on Friday) than over the weekend. The selected days also experience traffic volumes similar to the most recent traffic records in the scenario [23].

IV. SCENARIO MODELLING IN SUMO
The selected scenario has been implemented in SUMO, a popular open-source microscopic road traffic simulator. SUMO can simulate different types of vehicles. This study uses the SUMO Krauss car-following model [16] and a four-layer lane changing model (referred to as LC2013 or JE2013) [17] to model the behavior of vehicles. Both models are stochastic to account for the human behavior. The parameter sigma (imperfection of the driver) controls the stochastic nature of the models. A vehicle with sigma=0 represents a

---

1 Availability refers to the amount of time with valid traffic data in the records provided by the road authority with respect to the total time considered. For the remaining selected days, 82 detectors produced valid traffic data with an average availability of 99.03%.
deterministic driver. The Krauss car-following model allows vehicles to drive as fast as they can (up to the desired speed) while keeping a safety distance with the preceding vehicle. SUMO models the desired speed as a normal distribution with a mean equal to the road speed limit multiplied by speedFactor and a standard deviation equal to speedDev. These configurable parameters strongly influence the capacity to accurately reproduce realistic traffic conditions (see Section VIII). The lane change model considers four different reasons for a lane change [17].

The road network has been imported into SUMO using the open-source online maps provider OpenStreetMaps (OSM) [24]. OSM is not specifically designed for traffic simulation so a conversion has been necessary to adapt its information to SUMO (using the SUMO tool NETCONVERT). We have also utilized OSMOSIS (a tool in OSM) [25] for automating the process to extract or modify information in OSM files. We used OSMOSIS to remove from the OSM file all the roads except those tagged as motorway or motorway_link. After importing the road network, the following manual work (using the SUMO tool NETEDIT) was necessary to obtain a precise road network that can accurately simulate realistic traffic:

- We used the SUMO tool NETEDIT to manually edit the road network and delete roads that are not part of the scenario. For example, we only maintained the final sections of the ramps as vehicles enter the freeway. We revised the scenario and restored it to its status in January 2015 (using Google Maps) since the road topology is currently slightly different due to recent road works. Similarly, we revised speed limits for all lanes checking road signs (using photos from Google Maps) in January 2015. We established a 90 km/h speed limit in the mainline for heavy vehicles to be compliant with Spanish regulations.
- All the acceleration and deceleration lanes for on-ramps and off-ramps have been created since OSM does not have this information. Manual work has been necessary to adjust the length and position of the lanes (using Google Maps), the speed limits, the definition of zipper nodes at on-ramps, and the design of combined on-off ramps.

Additional information on the manual modifications that have been necessary for accurately modeling the selected road traffic scenario can be found in [15].

V. TRAFFIC DEMAND CALIBRATION IN SUMO USING CADYTS

The traffic demand is the set of vehicle trips (or routes) from origin to destination that are produced in a road during a specific period of time. The traffic demand can be generated in SUMO using Cadys (Calibration of dynamic traffic simulations) [19], a calibration algorithm based on an iterative Bayesian process. Cadys computes the probabilities to choose routes based on the probabilities that these routes were used in the past and the traffic flow measured in the road network. Cadys adjusts the simulated traffic flow to the input measurements. The resulting speed and occupancy in the road network depend on the quality of the calibration process and the realism of the modeled scenario. Figure 2 represents how Cadys operates in SUMO. The figure shows that Cadys requires two inputs: the traffic flow measurements and the initial trip list. In this study, the traffic flow is averaged every 5 minutes using measurements from the induction loops deployed in the scenario. The initial trip list is a predefined list of trips that Cadys uses as a basis to generate traffic flows that match the measurements. A trip is defined by a departure time, vehicle type, origin, destination, and probabilities of the possible routes from origin to destination. Cadys calibrates the traffic with the following four components illustrated in Figure 2:

1) Initialization: the algorithm reads the traffic flow measurements per detector (induction loop in our study).
2) Choice: Cadys selects from the initial trip list a group of trips that it estimates will produce in the simulator the same traffic flow as measured by the road detectors. For each trip, Cadys selects the most probable route from origin to destination. This process is simpler in a freeway scenario since only one route is generally possible between origin and destination. Additionally, Cadys

![FIGURE 1. Average traffic flow on the mainline (vehicles counted in periods of 5 min) on January 23 2015.]

![FIGURE 2. Cadys-SUMO traffic demand calibration.]

\[2\] We used the combined on-off ramps to improve simulation realism. In these ramps, the acceleration lane of an on-ramp becomes the deceleration lane of a nearby off-ramp. To prevent deadlocks (especially when the acceleration/deceleration lane is short), SUMO recommends using a connection pattern where the off-ramp can be taken from the deceleration lane and from the adjacent lane.
decides how many vehicles take this route. This allows generating more realistic traffic demand that better matches the measured traffic flows.

3) Traffic Microsimulation: SUMO simulates the selected trips and measures the simulated traffic flows at the location of the detectors as well as the travel times per road section.

4) Update: Cadys computes for each detector the difference between the simulated and the real measured traffic flows. It then adjusts the probabilities of choosing routes that cross each detector so that the simulated and measured traffic flows better match in the following iteration.

The algorithm is executed iteratively (following the loop illustrated in Figure 2) in order to reduce the deviation between the simulated and measured traffic flows. For each iteration, Cadys adjusts its selection of trips. The iterative process finishes when the deviation stabilizes and additional iterations do not improve the calibration. In particular, the process ends when the moving average of the deviation (using a window size of 200 iterations) decreases by less than 1% over the last 100 iterations. The calibrated traffic demand is then equal to the selected trips (and their configuration) for the iteration that produced the smallest difference between simulated and real measured traffic flows.

A challenge to calibrate the traffic demand with Cadys is that it requires an initial trip list. Some studies propose using activity patterns [18] or the DUAROUTER tool in SUMO [22] to generate synthetic trip lists. The objective is to generate a pre-calibrated initial trip list that is somehow similar to reality and that Cadys can adjust. However, this pre-calibration is not straightforward. It may require data that is not available and the use of additional traffic modelling techniques that increases the complexity. We propose to address these two limitations with a novel learning-based and iterative traffic demand calibration technique in SUMO that is presented in the following section.

VI. PROPOSED TRAFFIC DEMAND CALIBRATION METHOD: CLONE FEEDBACK

This study proposes Clone Feedback, a new method to generate realistic traffic demand in SUMO. The method uses Cadys but does not require a pre-calibration to generate realistic traffic demand. The proposed method is capable to generate calibrated mixed traffic (light and heavy vehicles) using as input only the traffic flows measured by the road detectors (in our study, induction loops). It also ensures that the routes are realistic, starting at on-ramps and ending at off-ramps. This is in contrast to other calibration techniques [20] that produce routes starting and/or ending in the mainline.

The proposed method is shown in Figure 3 and is based on an iterative execution of the Cadys-SUMO traffic demand calibration process shown in Figure 2. The proposal uses as input the traffic flow measurements and a large randomly generated initial trip list. We choose a uniform random initial trip list for various reasons: 1) it makes no assumptions about traffic conditions and hence does not interfere with Cadys’ calibration process and 2) it can be generated easily in SUMO. Selecting a large initial trip list increases the probability to consider the trips producing the measured traffic flows. In addition, it reduces the calibration time and computational cost since Cadys can only increase traffic demand by a maximum factor of two in each calibration. Our proposal introduces an iteration through a new Clone Feedback module (Figure 3) that adds trips cloned by Cadys to the list of trips that Cadys uses for its calibration process. This is done because Cadys clones the trips that adjust better the simulated traffic flows to the measured ones. In each iteration, we add to the list of trips used by Cadys all the trips that were cloned in all previous iterations. By adding these cloned trips to the initial list, we introduce a learning process and ensure the iterations of the Cadys calibration process consider historical knowledge about the trips that were capable to better adjust the simulated traffic flows to the measured ones. This process helps reducing the difference between the measured and simulated traffic flows with the number of iterations. For example, the deviation is reduced by 38% from first to second iteration, and by 23% from second to third iteration considering the traffic on January 23. It is also important to emphasize that the proposed approach eliminates the need to derive a pre-calibrated initial trip list, and hence simplifies the generation

![Figure 3. Proposed Clone Feedback traffic demand calibration.](image3)

![Figure 4. Mixed traffic calibration.](image4)

3 In contrast, Cadys can reduce traffic demand without restriction in each calibration if necessary.

4 Each cloned trip is added twice.
of simulated traffic demand while achieving a good match between simulated and measured traffic (see Section VIII).

The proposed technique can calibrate mixed traffic scenarios\(^5\). The mix of different types of vehicles has a non-negligible impact on traffic dynamics \([26]\). However, most of the existing SUMO scenarios do not include heavy vehicles \([9],[10],[5],[20],[22]\). Figure 4 shows how our proposal calibrates mixed traffic scenarios. First, an independent calibration is conducted for each type of vehicle (light and heavy in our scenario). This calibration uses the proposed Clone Feedback process using flow measurements and initial trip lists for each type of vehicle separately (each module referred to as Clone Feedback in Figure 4 corresponds to an instantiation of the proposed process illustrated in Figure 3). Once calibrated, the selected trips for each type of vehicle are combined and used as input for a second calibration phase (Figure 4). This second phase is executed using the Cadys-SUMO calibration process depicted in Figure 2. It uses the combined selected trips as initial trip list and the mixed traffic flow measurements as input to the calibration process. This second phase is required for two reasons. First, it is necessary to adjust the variations of the traffic flow resulting from the presence of different types of vehicles in the scenario. Second, not all detectors can detect the type of vehicles (e.g. the detectors in ramps in our scenario).

Many ramp detectors (including those in the Alicante-Murcia freeway) cannot detect the type of vehicle. This could be in principle a drawback for the independent calibration phase in Figure 4 since it is not possible to know the percentage of light and heavy vehicles crossing the detectors. We analyzed two alternatives to address this situation. The first one is to use only measurements from mainline detectors for the independent calibration phase. The second one is to use estimated traffic flow data for the ramp detectors in addition to the measurements from the mainline detectors. To this aim, we use the measurements from the mainline detectors to estimate the overall rate of light and heavy vehicles in the scenario. This rate is then applied to the measurements from the ramp detectors to generate independent flow measurements for light and heavy vehicles. Table 1 shows the deviation (MAE, Mean Absolute Error) between the simulated and measured traffic flows considering both alternatives. The table shows the results for the first (independent) and second (combined) calibration phases. The results for the independent calibration phase consider only average values over mainline detectors since there are no independent measurements for ramps. On the other hand, the results for the combined calibration phase are depicted for mainline detectors only, and for mainline and ramp detectors. Table 1 shows that the first alternative achieves a lower deviation than the second one in the mainline for the independent and combined calibration phases. However, the overall deviation (in mainline and ramps) is better with the second alternative. This indicates that the rate of light and heavy vehicles estimated from mainline detectors matches well the rate of vehicles also crossing the ramp detectors. The table also shows the performance obtained when the combined (2nd) calibration phase is not applied\(^6\). The results show that lower deviations are obtained for the total traffic when the combined calibration (2nd phase) is applied at the expense of a slightly worse calibration of each type of vehicle.

### VII. IMPLEMENTATION AND CONFIGURATION

This section explains how we implemented and configured the proposed calibration technique in SUMO and Cadys for optimizing the traffic demand calibration process that we used to create the Alicante-Murcia SUMO scenario.

#### A. INITIAL TRIP LIST

We generate the initial random trip list using the SUMO tool `randomTrips`. The trips generated include the departure time, vehicle type, origin and destination. We then formatted the trips for Cadys which requires the definition of all possible routes between origin and destination (i.e. the complete list of road sections - called edges in SUMO - of the route), the travel time of every road section in the route, and the probability that each route is selected. This probability is equal to one for the freeway scenario since only one route is possible for a given origin-destination pair. We then simulate the generated trips in SUMO setting the probability of each route equal to one and using the following options: `vehroute-output; vehroute-output.exit-times; vehroute-output.sorted; vehroute-output.dua`. In addition, we use the following configuration of the `randomTrips` parameters to generate trips with the format required by Cadys:

- **Fringe Factor**: this parameter controls the probability that vehicles’ routes start/end at road network edges.

---

\(^5\) SUMO can model different types of vehicles. Our dataset provides information about light and heavy vehicles in the mainline detectors.

\(^6\) Instead, we simulate together in a single iteration of SUMO the trips for light and heavy vehicles selected from the first (independent) calibration phase.
This probability must be 1 for freeway scenarios in order to simulate realistic conditions where vehicles cannot start or end a trip in the middle of the freeway since ramps are the only possible accesses. To this aim, this parameter must be set at a value of various orders of magnitude greater than the number of vehicles generated (10^6 in this study).

- **Period**: randomTrips generates a trip every period between the start and end simulation time. This period must be selected as low as possible in order to generate a large number of trips. However, it should not be too low in order to avoid an increase of the computational cost of SUMO simulations as well as deadlocks resulting from an extremely congested scenario. The period has been set in this study equal to 0.3 s for light vehicles and 0.4 s for heavy vehicles.

- **Edges**: randomTrips is used with the option lanes enabled and with speed-exponent=2. When the lanes option is enabled, the probability that an edge is selected as origin or destination of a trip is weighted by the number of lanes of that edge. This results in more realistic traffic patterns as it generates more trips in multi-lane edges compared to single-lane ones. With speed-exponent=2, the probability that an edge is selected as origin or destination of a trip is weighted by the square of the speed limit of the edge. This generates more trips in faster edges.

- **Trip attributes**: the following attributes have been utilized for a seamless insertion of vehicles in the simulation: departLane=best (inserts vehicles in the lane that allows the longest ride possible without the need to change lanes); departSpeed=max (inserts vehicles at the maximum speed of the edge where they are inserted or the maximum speed that ensures a safe distance with the front vehicle); departPos=last (inserts vehicles at the position of the lane that ensures a safe distance with the front vehicle given its departSpeed).

- **Validate**: this option is activated so that only trips between valid origin-destination pairs in the freeway are produced. We qualify a pair as valid if a vehicle can reach the destination from its origin using the available roads in the scenario.

After launching the simulation with the generated trip list, it is necessary to wait until the road network fills and we can produce valid flows for the calibration. Our trials have shown that 10 minutes of simulation time are enough for this.

**B. CONFIGURATION**

Cadyts is available in the SUMO tool cadytsIterate. We modified this tool so that we can modify certain Cadyts’ parameters. This is necessary to implement the proposed calibration technique. In our implementation, we use the default configuration of Cadyts except for the following parameters:

- **Binsize (5 minutes)**: this parameter represents the duration of the time bins in which the calibration internally collects simulation-related data. The parameter is set equal to the period of the traffic measurements, i.e. 5 min.

- **Override Travel Times** (true): we activate this parameter so that we can consider the vehicles’ travel times in the last iteration during Cadyts’ update phase (Combined calibration phase in Figure 2).

- **Count Last Link**: we activate this parameter so that exiting vehicles cross the upstream detector of their exit edge. This is necessary to correctly compute simulated traffic flows on off-ramps.

- **Stddev**: this parameter represents the smallest allowed standard deviation of the traffic simulation measurements. The parameter is set equal to 1 veh/5min following [19] and [27] that claim that the lower the standard deviation the better the fitting of simulated to measured traffic flows.

- **Regression Inertia**: this parameter describes the inertia of the regressions models that track the simulation behavior. A small value is likely to lead to oscillations and unstable results, whereas a large one leads to slow convergence. The default value is 0.95. However, our tests showed that SUMO simulations are better calibrated with an inertia value of 0.99 since the trip disruptions between iterations are smoother. cadytsIterate has been modified to set this parameter to 0.99.

We also use the default configuration of SUMO except for the following parameters:

- **vClass**: each vehicle type is modelled by a specific vClass that defines the characteristics of the vehicles (e.g. dimensions, acceleration or maximum speed). This parameter can be used to enable/disable the usage of lanes for certain types of vehicles. In this study, light vehicles are modeled with the “passenger” vClass and heavy vehicles with the “trailer” vClass.

- **speedFactor** and **speedDev**: these parameters are used to configure the maximum desired speed of each vehicle. Each vehicle is assigned a maximum desired speed from a normal distribution with mean equal to speedFactor multiplied by the road speed limit, and a standard deviation equal to speedDev. We tested three values for speedFactor: 0.95, 1 and 1.05. A value equal to 1.05 means that the maximum desired speed of the vehicle is on average 5% above the road speed limit. We also tested three values for speedDev (0.05, 0.1 and 0.15) and the impact is analyzed in the following section.

- **Sigma**: driver’s imperfection. This parameter introduces randomness in the vehicle’s behavior. A sigma value equal to 0 represents a deterministic driver. We executed the calibration process with three different values to better adjust the sigma parameter:
0.1, 0.3 and 0.5. We discarded the value of 0.5 since it produced highly congested situations leading to deadlocks in SUMO. sigma values equal to 0.1 and 0.3 produce very similar results in terms of traffic flow and speed but generate significant differences in terms of the road’s occupancy. sigma equal to 0.1 achieves a significantly closer average occupancy to the average occupancy provided by the traffic measurements in the Alicante-Murcia freeway scenario. In fact, the occupancy relative error is 150% higher with a sigma value of 0.3 than with a sigma value of 0.1.

VIII. EVALUATION

This section compares the simulation scenario created with the proposed traffic demand calibration technique to the measurements provided by the Spanish road authority for the complete Alicante-Murcia freeway scenario. To this aim, we compare the traffic flow, the speed and the road’s occupancy. Figure 5 compares the simulated and measured traffic flows. Figure 5a to Figure 5c plot the average traffic flow measured over all mainline detectors over periods of 5 minutes. The figure compares the traffic flow for each type of vehicle (Figure 5b and Figure 5c) and when combining both traffic types (Figure 5a). Figure 5a to Figure 5c correspond then only to the traffic flow at the mainline detectors since the ramp detectors cannot differentiate the type of vehicle. The simulated traffic flow is created with our proposed calibration technique, and the measured one corresponds to the data provided by the Spanish road authority. The figure represents the results for January 23; similar trends were observed for the other 8 days selected to generate the traffic simulation scenario. The results show that the simulated traffic flows accurately match the real measured traffic flows, and the proposed simulation scenario and calibration technique can accurately reproduce traffic flow variations along the day. This includes all traffic patterns with morning and evening rush hours as well as lunch time and night periods. The accuracy of the proposed simulation scenario and calibration technique is shown for the mixed traffic and for each type of vehicle separately. The results obtained show that a better match is achieved with mixed traffic since the second phase of the calibration process adjusts the traffic flow to the total number of vehicles (Figure 4). Figure 5d compares the average traffic flow over ramp detectors only and considering mixed traffic with light and heavy vehicles (ramp detectors cannot differentiate the type of vehicle). The results show similar calibration accuracy at the ramps than at the mainline, which demonstrates the validity of the proposed scenario and calibration technique.

The accuracy of the developed scenario and the proposed calibration technique can be better visualized in Figure 5e. This figure directly compares all the values of the simulated and measured traffic flow. Each point in the figure compares the traffic flow measured in an interval of 5 minutes by one detector with the simulated traffic flow simulated in the same time interval and detector. A perfect fitting would result in all points placed within the bisector line. Figure 5e shows that the fitting is very good with Pearson correlation values (r) very

![Comparison of the simulated and measured traffic on January 23](image)

**TABLE 2.** Speed mean absolute error (m/s) over mainline detectors.

| speedFactor | speedDev | January 23 | 0.95 | 1 | 1.05 |
|-------------|----------|------------|------|---|------|
| 0.05        | -        | 2.87       | -    |   |      |
| 0.1         | 3.18     | 2.67       | 3.05 |   |      |
| 0.15        | -        | 2.87       | -    |   |      |
close to 1. We should also note that we achieve a very good fitting for all the range of measured/simulated traffic flows.

The previous results clearly demonstrate that the scenario we created using the proposed traffic demand calibration technique accurately reproduces the traffic flows in the 97 km long Alicante-Murcia freeway scenario. It is though important to remember that the proposed calibration technique adjusts the simulated traffic flow to the input measurements but does not directly calibrate the speed or the road’s occupancy. The capacity of accurately reproducing the speed and road’s occupancy depends on the quality of the calibration process and the realism of the modeled scenario. This includes the modelling of the road network (Section IV) and the vehicles (Section VII.B). The following results analyze how accurately the simulated model is able to reproduce the traffic speed and the road’s occupancy in the developed freeway scenario. Figure 6 compares the cumulative distribution function (cdf) of the measured traffic speed values to the simulated ones. The results are shown for January 23 for mixed traffic over the mainline detectors. We represent the simulated performance for different values of speedFactor and speedDev since they have a direct impact on the simulated vehicles’ speed. Figure 6a shows that the value of speedFactor (SF) that better adjusts the simulated traffic speed to the measured one is 1. Figure 6b shows that a speedDev (SD) value of 0.1 better adjusts the speed. Table 2 reports the average absolute speed deviation between the measured traffic speed and the simulated one for all the configurations tested of speedFactor and speedDev. The results confirm that the configuration achieving the lowest deviations (2.67 m/s) is speedFactor=1 and speedDev =0.1.

Figure 7 compares the measured and simulated road’s occupancy for different configurations of speedFactor and speedDev. Table 3 reports the deviation between the measured and simulated values for all the configurations tested of the speedFactor and speedDev parameters. The results show once more that our simulation scenario and calibration technique achieve good results and can accurately match the simulated road occupancy levels to those measured by the road detectors for the period under evaluation7. The better match between measured and simulated values is again achieved with speedDev=0.1. speedFactor equal to 1 and 1.05 achieve similar results. Figure 7a shows that speedFactor=1.05 achieves a better match for the road’s occupancy in general, although speedFactor=1 outperforms the other configurations in some highly congested situations8. We still consider that speedFactor=1 is the most appropriate value since it provides the best tradeoff in the calibration of the road’s occupancy and the traffic speed.

Table 4 analyzes the accuracy of our simulation scenario and calibration technique for all 9 days. The table shows that (Figure 5). For the remaining time, the simulated road’s occupancy is generally higher than the measured one even when the simulated traffic flow matches well the measured one. The reason is that the road is highly congested in a few detectors and SUMO cannot model that well congested traffic conditions.

7 Similar trends were observed for the other 8 days included in the scenario.
8 Figure 7 shows that the simulated road’s occupancy is lower than the measured one in the time interval 00:00h-06:00h. This happens because the simulated traffic flow is lower than the measured one during this time period.
we achieve similar performance and trends for all days with some minor differences. The deviation observed in the traffic flow and the road’s occupancy is smaller for days with lower traffic loads (Saturday and Sunday). This is the case because there are fewer trips that need to be calibrated, which eases the calibration process. The speed deviation is more stable along the 9 days period since the values of the traffic speed can be very similar for different traffic loads.

Figure 8 compares the simulated and measured traffic flows in all detectors when using our proposed technique (Clone Feedback) for calibrating the traffic demand and two other reference techniques DFROUTER and Cadys. DFROUTER is provided in the SUMO package, and it is used, for example, in [20] for the calibration of a scenario of the city of Valencia. Cadys is also provided in SUMO, and it is used, for example, in [22] for the calibration of a scenario of the city of Hefei. Cadys needs an initial trip list as input for the calibration. We produce this list randomly in this study for a fair comparison with our proposal. Figure 8 clearly shows that our proposal (Clone Feedback) outperforms the two reference techniques. Our proposal achieves values in Figure 8 that are mostly located near the bisector line (perfect fitting). This is not the case with the two reference techniques that experience a larger dispersion. This is particularly the case for the highest traffic flows where the accuracy achieved with our proposal Clone Feedback clearly outperforms the two reference techniques. This is the case because the use of cloned trips in our proposed technique allows generating higher traffic demands. Table 5 compares the traffic flow, occupancy and speed deviations (MAE metric) for the mainline on January 23. The table clearly shows that the proposed technique for calibrating the traffic demand outperforms existing reference schemes when calibrating the three traffic variables. For example, the deviation in the traffic flow obtained with the proposed technique is nearly one third of the one obtained with DFROUTER, and around half of the deviation obtained with Cadys. DFROUTER achieves the lowest performance since it does not select suitable routes and directly inserts 58% of vehicles in the mainline. Cadys achieves occupation deviation values close to those obtained with our Clone Feedback proposal. However, these results are misleading since these values result from the fact that Cadys is not able to reproduce traffic flows above 4000 veh/h (Figure 8). If Cadys could reproduce larger traffic flows, the occupancy deviation achieved with Cadys would increase.

IX. CONCLUSIONS
This paper has presented a large-scale and high-accuracy traffic simulation scenario for SUMO. The scenario accurately models the traffic flow, speed and road’s occupancy for 9 days of traffic over a 97 km freeway section between Alicante and Murcia in Spain. This is one of the largest traffic SUMO simulation scenarios that is openly released to the community (downloadable from [15]) and is a valuable testbed for V2X, CAVs and ITS research and engineering. The simulation scenario models mixed traffic with light and heavy vehicles, and has been calibrated using a unique dataset provided by the Spanish road authority. This dataset provides measurements from 99 induction loops (detectors) deployed in the scenario (both at mainline and on- and off-ramps). The simulation scenario has been created using a novel learning-based and iterative traffic demand calibration technique that is also presented for the first time in this paper. The proposed technique outperforms the two reference techniques for calibrating the traffic demand in SUMO (DFROUTER and Cadys). The proposed technique accurately matches the simulated traffic flow, traffic speed and road’s occupancy to the measured ones in the scenario. Good modeling accuracy levels are obtained for mixed traffic and when considering separately light and heavy vehicles. Future work includes expanding the scenario to model days in different seasons (with different characteristics of the traffic) and urban/suburban environments. The scenario could also be evolved to include 3D information about the road network, which could improve the realism of e.g., simulations of V2X communications simulations or the estimation of CO2 emissions among others. It could also be evolved with the modeling of automated vehicles [29]. Automated vehicles are expected to augment the road capacity [28], but a detailed analysis is needed to properly understand and quantify their impact on the traffic, especially under mixed traffic scenarios where automated and non-automated vehicles coexist [30].

REFERENCES
[1] M. Guériau, et al., “How to assess the benefits of connected vehicles? A simulation framework for the design of cooperative traffic management strategies,” in Transportation Research Part C: Emerging Technologies, vol. 67, pp. 266-279, 2016, DOI 10.1016/j.trc.2016.01.020.
[2] J. Alfonso, et al. “Cooperative ITS communications architecture: the FOTsis project approach and beyond,” in IET Intelligent Transport Systems, vol. 9, issue 6, pp. 591-598, 2015, DOI: 10.1049/iet-its.2014.0205.
[3] M. Sepulcre, et al., “Cooperative vehicle-to-vehicle active safety testing under challenging conditions,” in Transportation Research Part C: Emerging Technologies, vol. 26, pp. 233-255, 2013, DOI: 10.1016/j.trc.2012.10.003.
[4] J. Harri, et al., “Mobility models for vehicular ad hoc networks: a survey and taxonomy,” in IEEE Communications Surveys & Tutorials, vol. 11, no. 4, pp. 19-41, 2009, DOI: 10.1109/SURV.2009.090403.
[5] S. Uppoor, et al., “Generation and Analysis of a Large-Scale Urban Vehicular Mobility Dataset,” in IEEE Transactions on Mobile Computing, vol. 13, no. 5, pp. 1061-1075, May 2014, doi: 10.1109/TMC.2013.27.
[6] D. Krajzewicz, et al., “Recent development and applications of SUMO-Simulation of Urban Mobility,” in International Journal On Advances in Systems and Measurements, vol. 5, no 3&4, pp. 128-138, 2012.
[7] M. Rondinone, et al., “TETRIS: a modular simulation platform for the large scale evaluation of cooperative ITS applications,” Simulation Modelling Practice and Theory, Vol. 34, pp. 99-125, May 2013.
[8] C. Sommer, et al., “Bidirectionally Connected Network and Road Traffic Simulation for Improved IVC Analysis,” in IEEE Transactions on Mobile Computing, vol. 10, no. 1, pp. 3-15, 2011, DOI: 10.1109/TMC.2010.133.

VOLUME XX, 2017

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
JUAN JESUS GONZALEZ-DELCADO received a Telecommunication Engineering Degree and a Mechanical Engineering Degree from the Universidad Miguel Hernández (UMH) de Elche (Spain) in 2009. He was recognized by the professional organization of Mechanical Engineers and by the UMH with the Best Student award. He was a researcher at UMH working on traffic characterization and prediction under the presence of connected and automated vehicles. He is now working for Talgo, a Spanish train manufacturer, as electrical systems engineer.

JAVIER GOZALVEZ received an electronics engineering degree from the Engineering School ENSERBB (Bordeaux, France), and a PhD in mobile communications from the University of Strathclyde, Glasgow, U.K. He is currently Full Professor and Director of the UWICORE laboratory at the Universidad Miguel Hernández de Elche (UMH) in Spain where he leads research activities in the area of connected and autonomous vehicles, beyond 5G networks, and Industry 5.0. He is an elected member to the Board of Governors of the IEEE Vehicular Technology Society (IEEE VTS) and was the 2016 and 2017 President of the IEEE VTS. He has served as IEEE Distinguished Speaker and IEEE Distinguished Lecturer for the IEEE VTS. He is the Editor in Chief of the IEEE Vehicular Technology Magazine, and the Founder and General Co-Chair for the IEEE Connected and Automated Vehicles Symposium.

JESÚS MENA-OREJA received an Industrial Engineering Degree from the Universidad Miguel Hernández (UMH) de Elche (Spain) in 2015. He was recognized by the professional organization of Industrial Engineers with the Best Student award. He is currently a PhD student at UMH working on the use of Machine Learning and Big Data for traffic characterization and prediction under the presence of connected and automated vehicles.

MIGUEL SEＰULCRE received a Telecommunications Engineering degree in 2004 and a Ph.D. in Communications Technologies in 2010, both from Universidad Miguel Hernández de Elche (UMH), Spain. He was awarded by the COIT (Spanish association of Telecommunications Engineers) with the prize to the best Ph.D. thesis. He has been visiting researcher at ESA (The Netherlands), at the Institute of Technology (Germany), and at Toyota InfoTechnology Center (Japan). He is currently an expert of the ETSI STF 585 on Multi-channel Operation for Release 2 V2X services. He serves as Associate Editor for IEEE Vehicular Technology Magazine and IEEE Communications Letters. He was TPC Co-Chair of IEEE VTC2018-Fall, IEEE/IFIP WONS 2018 and IEEE VNC 2016. He is the Associate Professor at UMH, and member of UWICORE research lab working in wireless vehicular networks and industrial wireless networks.

BALDOMERO COLL-PERALES received his M.Sc. and Ph.D. degrees in Telecommunications Engineering from the Universidad Miguel Hernandez (UMH) de Elche, Spain. He is currently a Juan de la Cierva Research Fellow at the UWICORE laboratory (UMH). He has held visiting research positions at Hyundai Motor Europe Technical Center (Germany), Italian National Research Council (Italy), WINLAB-Rutgers University (USA) and Institute of Telecommunications-King’s College London (UK). His research interests lie in the field of advanced
mobile and wireless communication systems, including the design of
device-centric technologies for future wireless Beyond 5G network and
connected and automated vehicles. He serves/has served as Associate Editor
for the International Journal of Sensor Networks and Springer's
Telecommunication Systems. He has served as Track Co-Chair for IEEE
VTC-Fall 2018, and as member of TPC in over 35 international conferences.