SUMO User Conference 2021
https://doi.org/10.52825/scp.v2i.120

Published: 29 Jun. 2022

Automated Calibration of Traffic Demand and Traffic Lights in SUMO Using Real-World Observations

Michael Harth[1,2], Marcel Langer[1,2] and Klaus Bogenberger[2]
[1]AUDI AG
[2]Technical University of Munich
michael.harth@audi.de, marcel.langer@audi.de,
klaus.bogenberger@tum.de

Abstract

Virtual traffic environments allow for evaluations of automated driving functions as well as future mobility services. As a key component of this virtual proving ground, a traffic flow simulation is necessary to represent real-world traffic conditions. Real-world observations, such as historical traffic counts and traffic light state information, provide a basis for the representation of these conditions in the simulation. In this work, we therefore propose a scalable approach to transfer real-world data, exemplarily taken from the German city Ingolstadt, to a virtual environment for a calibration of a traffic flow simulation in SUMO. To recreate measured traffic properties such as traffic counts or traffic light programs into the simulation, the measurement sites must first be allocated in the virtual environment. For the allocation of historical real-world data, a matching procedure is applied, in order to associate real-world measurements with their corresponding locations in the virtual environment. The calibration incorporates the replication of realistic traffic light programs as well as the adjustment of simulated traffic flows. The proposed calibration procedure allows for an automated creation of a calibrated traffic flow simulation of an arbitrary road network given historical real-world observations.

1 Introduction

The release of a fully functional automated driving system implies a high effort in testing and development. As real-world tests are time-consuming, expensive, not replicable and potentially dangerous, testing is primarily shifted towards a virtual environment. In order to replace real-world tests by virtual experiments, there are high demands towards a realistic representation of the real world in simulation. Therefore, the approach has been pursued to couple a submicroscopic vehicle simulation and a microscopic traffic simulation, e.g. [1]–[4]. In the approach, traffic simulation provides a temporally and locally realistic traffic system for an autonomous vehicle under test. Submicroscopic simulation controls the specific behavior and interactions of the involved traffic participants only in the direct environment of the tested vehicle. For achieving valid, realistic traffic surroundings, the traffic system must be accurately calibrated with real-world data. Therefore, traffic properties like traffic counts as well as traffic light programs must be obtained and assigned to the simulation, in order to recreate realistic traffic conditions. In order to prove simulation validity and especially enable the comparison of simulated and real-world results, an accurate representation of the real world conditions is inevitable.

The main contribution of this work is a scalable calibration process of a microscopic traffic simulation using the microscopic traffic simulator SUMO [5]. We present a tool chain, which enables real-world data acquisition for a requested time interval, the allocation of these measurements in the virtual world as well as the calibration of traffic flows and traffic lights. The validation of this comprehensive approach takes into account the comparison of simulated measures to the corresponding real world observations as well as the ability of the calibrated traffic lights to cope with the provided traffic demand. The developed tool chain allows for an application of the calibration method to any arbitrary traffic simulation in SUMO given necessary historical data. As our focus is mainly on the replication of a historical traffic scenario, we will not take into account calibration of driver models.

Section 2 provides an overview on related work regarding map association, traffic light emulation and traffic assignment. In Section 3, the allocation of real-world data and the calibration procedure are described in detail. The evaluation of the procedure is presented in Section 4. Section 5 summarizes the main contributions and provides an outlook regarding future work.

2 Related Work

The proposed approach to calibrate the traffic flow simulation in SUMO using real-world observations can mainly be separated into three predominant tasks. The first task consists of the representation of the virtual road network and allocation of real-world data to the virtual world. In a next step, traffic light programs must be created and implemented to the simulation. In a last step, traffic must be assigned based on observed traffic measurements. The following literature review addresses these relevant steps.

2.1 Map association

The allocation of real-world data in the virtual environment can be broken down to a map-matching problem. According to Quddus et al. [6], map-matching procedures can be classified to four different categories. Geometric map-matching algorithms take into account geometric information of a road network by neglecting logical connections between roads. Bernstein and Kornhauser [7] distinguish point-to-point matching, point-curve-matching and curve-to-curve matching as representatives of geometric map-matching algorithms. In contrast to geometric algorithms, topological map-matching techniques consider the geometry as well as the linkage of network elements. The third category of map-matching algorithms, probabilistic algorithms, create an elliptical or rectangular confidence region around a coordinate to be matched. For the determination of the corresponding network element, the confidence region is superimposed on the network. Advanced map-matching algorithms apply several methods like Kalman Filter or an Extended Kalman Filter, a flexible state-space model in combination with a particle filter, a fuzzy logic model or an implementation of Bayesian inference.

2.2 Traffic light emulation

In the literature, there are two approaches for the creation of simulated traffic lights. The first approach treats the emulation of existing traffic light programs for a realistic representation in simulation [8]–[14]. Besides, the second concept is the optimization of existing traffic light programs in order to improve certain measures, e.g. traffic flow, using reinforcement learning or evolutionary algorithms [15]–[18]. As our purpose is on building a detailed representation of the real world, we only provide an overview on literature on the first approach to estimate and emulate traffic light programs.

Schönfelder et al. [8] reconstruct traffic light programs of fixed-time traffic light cycles using camera data from probe vehicles approaching an intersection with a known cycle time. The procedure requires only few and less quality input data for high accuracy results. Axer and Friedrich [9], [10] provide an approach for the estimation of traffic light programs and the cycle durations of fixed-time as well as traffic actuated traffic lights based on sparse simulated floating car data (FCD). Besides, Wang and Jiang [11] use recorded FCD trajectories of vehicles passing a signalized intersection to derive traffic light programs and cycle times. Rostami-Shahrbakaki et al. [12] apply a machine learning approach for determining signal timings of traffic lights based on FCD of connected vehicles. They achieve high accuracy in estimating cycle time at low percentages of connected vehicles (~ 5 - 10 %). For detecting green time durations, higher penetration rates of connected vehicles (~ ≥ 15 %) are necessary to obtain higher precision estimations. The patent of Wolf et al. [13] introduces a procedure using a genetic algorithm to emulate future traffic light signals from historical signal timing information accessed by an interface to the traffic light backend. Weisheit [14] applies a Support Vector Machine on historical signal timing data for predicting future traffic light states.

2.3 Traffic assignment

For the assignment of traffic to a simulation network, trip-based as well as activity-based methods are applied. The four-step model [19] describes a trip-based approach, which divides a simulated region into several traffic assignment zones. These zones serve as origins and destinations of trips in the network. The first step of the model represents the trip generation, in which the frequency of originating and arriving trips is determined for each zone by different trip purposes based on socio-economic statistical measures. In the second step, called trip distribution, trip interchanges between zones are defined in order to meet the frequencies of the first step. The result of the second step is an origin-destination (OD) matrix incorporating the number of trips from each origin to every destination in the network. There are various methods to improve an OD matrix based on observed traffic counts. Van Zuylen and Willumsen [20] introduce information minimization and entropy maximization for estimation of an OD matrix using traffic counts. Further procedures apply generalized least squares [21], [22], a Bayesian estimator [23] and maximum likelihood estimation [24]. The third step of the four-step model implies the transportation mode choice for each trip. Finally, the last step incorporates the route assignment for each trip according to the mode-specific network. For route assignment, Wardrop introduces two concepts regarding the optimization objective. The user equilibrium defines the state, in which all vehicles are assigned to their individual fastest route. Besides, the system optimum describes an equilibrium, in which the total travel times are minimal.

Activity-based models consider traffic assignment on a person- and household-related level rather than on a zone-based level such as the four-step model. The route for each person is affiliated with certain activities and comprises all trips over an entire day. Apart from general socio-economic statistical data, household surveys are an essential basis in order to realistically represent individual travel behavior. Lobo et al. [25] introduce an activity based approach for a simulation of Ingolstadt called InTAS. The city is divided into several districts. According to regional statistics regarding the number of inhabitants, commuters, workplaces, education possibilities and demography, traffic is assigned to the network. However, the traffic assignment does not use traffic counts to generate the demand. Additionally, the traffic lights in the network are created manually instead of using real-world data.

3 Methodology

As introduced in the previous section, calibration of a traffic flow simulation is split up to map association tasks, traffic light program emulation and traffic flow calibration. For the calibration of the traffic flow simulation in SUMO, we use a network comprising the entire city of Ingolstadt, Germany, with an extent of about 10 km by 10 km. It covers more than 12000 edges, including urban, rural and highway roads, as well as 5600 junctions with 120 traffic light actuated intersections. The network is based on an import from OpenStreetMap converted to a SUMO network. Figure 1 depicts a section of the road network.

Figure 1: Section of the road network of Ingolstadt

The following Figure 2 provides an overview on the necessary steps for the calibration purpose, which will be explained in detail in the following sections.

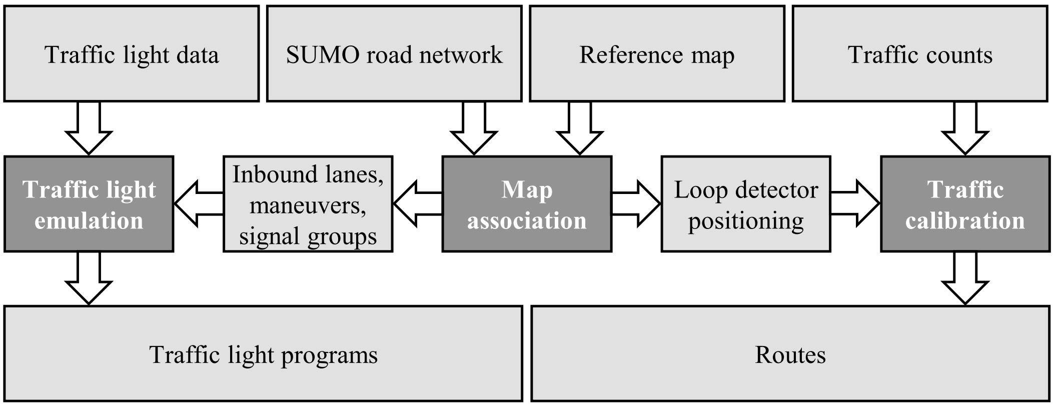

Figure 2: Overview on the calibration procedure

3.1 Map association

A key part of the calibration procedure of the traffic simulation is the linkage between real observations and their representation in the virtual world. For the map association tasks, we apply a framework based on a parsed representation of the SUMO network towards a GeoDataFrame, introduced by the Python library geopandas [26]. The framework enables map-matching tasks, which exceed SUMO basic functionalities, by not only applying geometrical matching, see section 2.1, but also considering the logical connections and properties of the map, e.g. lane successors, lane types etc. The application range of the framework is not only limited to matching tasks related to the calibration of a traffic simulation. Furthermore, the framework can also be used for applications such as vehicle trajectory matching or public transport routing. We plan to provide the developed framework as an open-source Python package.

For the calibration of the simulation, real signal groups of traffic lights must be mapped to the inbound lanes and in a second step to the succeeding connecting lanes of traffic light actuated junctions in SUMO. Signal groups provide signal state information for multiple associated lanes. Furthermore, loop detectors are allocated to the corresponding lanes given their geographical positions. For both applications, allocation is based on a reference topology representing all inbound lanes for each traffic light actuated intersection in Ingolstadt including associated stopping lines, signal groups as well as possible maneuvers. Besides, the geographical positions of all loop detectors can be extracted from the reference map.

In order to assign signal groups to the SUMO network, in a first step, each traffic light actuated junction from SUMO must be identified in the reference topology by simply comparing the geometrical centers of the respective junctions considering small deviations (~10 m) of both maps. As signal group association in the reference topology is stored as an attribute of inbound lanes, the attribute in SUMO is linked to connecting lanes within a junction. For the association of signal groups to SUMO, the inbound lanes of the respective junction in the SUMO network must be derived and matched to the inbound lanes from the reference. To prevent errors in the allocation process, the number of inbound roads and lanes of each traffic light actuated intersection must be the same for both maps. Therefore, these numbers are extracted and compared for the inbound roads and lanes. As there are deviations between the reference and the SUMO network, corresponding lanes from each map do not generally match geometrically. Therefore, the matching of the junction geometries is performed in two steps. In the first step, the inbound roads of the topology are matched with the inbound roads of the SUMO map. Once the pairing is done on the road level, the lanes within each individual pair of roads are matched between both maps. In order to complete this two-step matching, detailed information regarding the association between individual lanes belonging to the same roads is required.

For the SUMO network, the number of inbound roads and lanes can be determined directly by the unique road identifier. The intersection reference map also provides information on which lanes form the associated roads. In order to group inbound lanes from an older version of the reference map without any knowledge on inbound roads given the number of inbound roads from SUMO, additionally hierarchical agglomerative clustering is applied. A comparison of both techniques shows that the clustering produces the same groups of inbound lanes given in the reference map with knowledge about inbound roads. In case of different numbers of inbound roads or lanes due to deviations of both maps, these differences are logged for a manual inspection of both maps and for corrections, which are essential to any of the following steps to allocate signal group information in SUMO. For an association of the inbound roads of both maps, the aggregated geometric centers of the stopping lines of all lanes of each inbound road are compared. The pairing between the two maps is carried out by minimizing the Minkowski distance (p-norm) between each pair of inbound roads from both maps. The Minkowski distance enables an emphasized weighting of shorter distances by applying an exponent, e.g. squaring, in order to weight larger distances with an exponentially higher value. Finally, for the allocation of inbound lanes, all lanes within each pair of inbound roads are considered and again the overall Minkowski distance between the inbound lanes of both the junction in the reference topology and in the SUMO network is minimized. The overall distance is selected as the quality parameter of the matching process as large deviations occur between lanes of the SUMO and the reference map. Furthermore, the application of the overall distance enables an optimal matching for the combination of all corresponding lanes rather than only for several specific lanes. Figure 3 shows the results of the clustering and pairing procedure for an exemplary intersection in the SUMO network and the reference map.

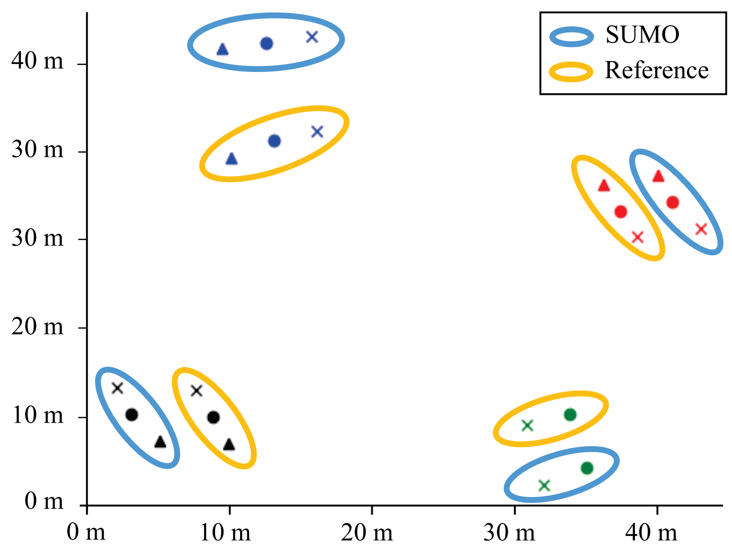

Figure 3: Clustering of inbound roads (blue and yellow ellipses) and pairing of inbound lanes (colored symbolic markers)

In Figure 3, the symbolic markers represent the geographical positions of the centers of the stopping lines of the inbound lanes. The inner markers are derived from the SUMO network and the outer markers result from the reference map. The color of the markers indicates the association to inbound roads, i.e. markers with equal color indicate the ends of all lanes from a single road. Furthermore, equal symbols having equal colors provide information on related pairs of inbound lanes comprehending both maps.

For the placement of loop detectors, a similar procedure of grouping and pairing is applied. Loop detectors are grouped by their inbound road. For the aforementioned older version of the reference map without information on inbound roads, detectors are grouped by using the agglomerative hierarchical clustering on the geographical positions. After grouping, the detectors are matched to their nearest SUMO inbound lanes by minimizing the Minkowski distance of the whole group. In contrast to the intersection matching, detectors are also matched and paired if the number of detectors and the number of inbound lanes differ. In case of more detectors than lanes, multiple detectors are placed on the same lane according to their geometrical position.

3.2 Traffic light emulation

In order to cover the strong influence of traffic control on traffic behavior, a realistic representation of traffic light programs is inevitable. Real traffic light programs in Ingolstadt are controlled traffic adaptively and incorporate features like green phase prolonging as well as public transport prioritization. For the proposed approach, average green times observed in reality are represented as fixed-time traffic light programs in the simulation. The individual effects on traffic of green phase prolonging and public transport prioritization are ignored, but their impact is still considered as average green time adaptions. Furthermore, different traffic light programs over a day are taken into account by hourly generating specific traffic light programs for each traffic light. Therefore, we provide an approach recreating these programs based on the same historical signal timing data used in [13]. Signal timing data can be requested for any traffic light actuated intersection in Ingolstadt for a defined time interval via an interface to the traffic light backend. The data are available for 75 traffic lights and contain the temporal intervals of green and red phases for each signal group of every specific traffic light with a resolution of one second. All remaining traffic lights in the network cannot be calibrated and are therefore adapted manually with static traffic light programs in order to cope with the traffic demand.

A realistic estimation of the traffic light programs requires determining the switching order and mean green time durations of all signal groups. For each signal program, one signal group is declared as the reference. For the requested time interval, this reference signal group is the first in the data to switch to a green state on the condition that all other signal groups have switched to red at least in the time step before. Next, the most common switching order, in which all signal groups have switched to green at least once and which starts with the reference signal group, is determined. Furthermore, the mean green and red time as well as the overall green time during the one-hour interval are calculated for each signal group. Incomplete signal timing phases in the recordings are ignored in the calculation. Based on the number of occurrences n_i in the most common order as well as the mean green time $\bar{t}_{g,i}$ and the proportion of the green time in the whole one-hour interval $p_{total,i}$, the cycle time $t_{cycle,i}$ for the emulated traffic light program is estimated for each signal group i individually according to equation 1.

$$t_{cycle,i} = \frac{n_i \cdot \bar{t}_{g,i}}{p_{total,i}} \tag{1}$$

Based on the individual cycle time estimations, the overall cycle time t_{cycle} for the emulated traffic light program is calculated as the mean of all $t_{cycle,i}$ weighted by $p_{total,i}$.

$$t_{cycle} = \frac{\sum(t_{cycle,i} \cdot p_{total,i})}{\sum p_{total,i}} \tag{2}$$

The weights are applied to estimate a general cycle time, which takes into account the relevance of certain signal groups for the whole cycle with respect to the green time in the entire one-hour interval. In order to fit into the determined cycle time, the emulated green times $\tilde{t}_{g,i}$ of each signal group are adopted based on their average green times $\bar{t}_{g,i}$ and red times $\bar{t}_{r,i}$.

$$\tilde{t}_{g,i} = \frac{\bar{t}_{g,i}}{(\bar{t}_{g,i} + \bar{t}_{r,i})} \cdot t_{cycle} \tag{3}$$

The signal switching order is created by applying the emulated green times under consideration of average overlaps of green times in the whole interval. If there is no overlap between two consecutive signal groups, a clearing time of 6 s is included in between.

For the implementation of the estimated traffic light programs in SUMO, all inbound lanes, which are linked with the relevant signal groups, must be associated to the SUMO map as explained in the map association section. As a result, the related inbound lanes of the reference topology and the SUMO net are available. As traffic light state information in SUMO is stored as an attribute of connecting lanes within each intersection, the succeeding connections of each inbound lane are determined. Furthermore, the feasible maneuvers for each connection are compared to the maneuvers linked to the signal groups in order to achieve a unique allocation of connections to a related signal group from reality. The generation of the necessary SUMO additional file, which incorporates the traffic light program, requires an iteration through one cycle of each traffic light. Yellow phases are accounted as a part of the red signal in the historical traffic light data. The yellow light duration is implemented after switches from green to red light according to German legislation [27]. For every change in state of any signal group, a new line in the SUMO additional file is created covering its duration as well as the current states of the corresponding connecting lanes.

3.3 Traffic assignment

For the calibration of simulated traffic flows, a two-step procedure is considered, which covers the determination of an initial pool of routes from OD relations and further calibration based on observed link counts. In order to obtain regional traffic flows, according to the first two steps of the four-step model [19], an initial OD matrix for Ingolstadt is determined based on statistical information. The initial OD matrix includes regional statistics from Ingolstadt regarding the distribution of inhabitants and workplaces in different municipalities as well as information on the attractiveness of municipalities with respect to free-time opportunities, shopping areas and educational establishments. Furthermore, we access aggregated data from Audi regarding residences and working time models of all 44000 Audi employees in Ingolstadt [28] considering data protection requirements that ensure no transparency of individual residences. The data from Audi are highly accurate and cover a significant number of trips in Ingolstadt having 136000 inhabitants [29]. As there are no local distributions regarding other trip purposes than work, their temporal distributions and trip lengths, as well as modal split, also Germany-wide statistics [30] are included. Finally, statistics of commuters traveling from and to Ingolstadt every working day are considered. Aggregating all this information, a procedure is established to create OD matrices taking into account the Ingolstadt municipalities, different trip purposes that are work, business, free-time, shopping and education, the temporal distribution of these trip purposes as well as the length of each trip with respect to the modal split. Modal split is separated to vehicles, public transport, bicycles and pedestrians. As a result, the procedure generates one OD matrix for every hour of an average working day and each mode of transport. The elements of each OD matrix can be converted to SUMO trips and routes using built-in executables *od2trips* and *duarouter* according to the route assignment of the four-step model. As the pool of resulting routes strongly affects the quality of the calibration outcome, the selection of roads in traffic assignment is weighted based on the road priority. Nevertheless, the pool of resulting routes is inaccurate due to statistical deviations from real local conditions and daily traffic characteristics making further traffic calibration inevitable. The calibration is conducted only for vehicle routes, as no further observation data regarding the other modes are available.

The pool of routes generated from statistical data is used as a basis for an hourly traffic assignment based on observed traffic counts in the next step of the calibration. The real-world counts are aggregated for each road with induction loop detectors. Additionally, the number of vehicles that would pass the same roads in the routes generated from statistical data is calculated. For the calibration, routes are copied two times and then successively removed until no further improvement can be achieved. The difference between the evaluated counts from routes and the observed counts for each road serves as a quality measure for the calibration. High values of the quality measure indicate a large error. This removal of routes is carried out in two steps. In a first step, only routes, which exclusively pass roads with an overestimation with respect to the observed counts, i.e. the worst quality measure, are removed at random. This procedure is repeated until no more routes meeting the criterion are found. In this step of the route selection, all routes are removed that pass only locations with an overestimation of traffic. In the second step, routes are also excluded if they have a positive impact on the overall error. This leads to a balancing of over- and underestimation of the error across all roads. The second step is repeated iteratively until any removal would lead to an increasing overall error value. The final set of routes contains anything between zero to three copies of each route taken from the initial pool of routes generated from statistical data. Since calibration is conducted on an hourly basis, routes for each hour are temporally shuffled inside the interval to prevent copied routes to start at the same point in time.

4 Results

For a validation of the calibration process, simulated observations are compared to their corresponding real-world counterparts. The results of the simulation are therefore analyzed in comparison to the input measurements. The evaluation has the objective to verify the functionality of the proposed calibration process. In the following, the quality of map association, traffic light emulation as well as traffic assignment is assessed.

4.1 Map association

For the map association of the signal groups to the virtual environment, all of the 75 considered traffic light actuated intersections can be fully matched showing the same number of inbound roads and lanes as well as the corresponding maneuvers for traffic light emulation. 18 intersections require slight manual changes in the SUMO network beforehand regarding inbound lanes or maneuvers in order to match the logic of the reference topology.

The association of the detectors between the two maps ends up with more errors than the intersection association due to geometric divergences between the two maps. Furthermore, the application of the agglomerative hierarchical clustering of detectors is error prone due to the missing number of clusters and strongly varying distances between detectors. Without knowledge on inbound roads, detectors are misplaced to outbound roads on occasion. Therefore, only the results of the method incorporating knowledge about inbound roads are presented. From the 529 detectors available in the road network, all are placed in the network. 210 of the 529 detectors must be replaced manually due to misalignment, while 319 are assigned correctly. In case of an equal number of detectors in a group of inbound lanes, the association of the detectors to the corresponding lanes works without errors as every lane gets occupied by one detector in the pairing process. Misalignments result from a higher number of lanes on a road than to be matched detectors in combination with map deviations. Shifted groups of detectors are paired with the closest set of inbound lanes, which may result in a shifted placement on the lanes. All in all, the 529 detectors are placed on 254 roads. For future simulation setups, the use of high definition map data is suggested as this makes an accurate geometric matching of detector locations to the simulation map possible.

4.2 Traffic light emulation

In the described approach, traffic light programs are reconstructed as fixed-time programs based on average green and red time durations of the different signal groups. In reality, traffic lights in Ingolstadt apply green phase prolonging for traffic adaptive control and public transport prioritization resulting in a discrepancy between real and simulated traffic lights. However, simulated traffic lights implicitly take into account the impact of these effects on the average green times. For the validation, emulated and real traffic lights are compared regarding the percentage of green time of each signal group with respect to the cycle time of the corresponding traffic light. In order to cover varying switching behavior of the real traffic lights along with cycle time deviations, the green time ratio of each real signal group is determined over the whole one-hour interval in which the calibration takes place. It is assumed that similar green time ratios lead to a comparable impact on the traffic.

Validation of emulated traffic lights is conducted for one defined week from Monday to Friday. Figure 4 indicates the distributions of absolute green time ratio errors in percent between emulated and real signal groups. All weekdays are aggregated and evaluated hourly for time intervals from 7 a.m. to 8 a.m., 9 a.m. to 10 a.m., 4 p.m. to 5 p.m. and 6 p.m. to 7 p.m.

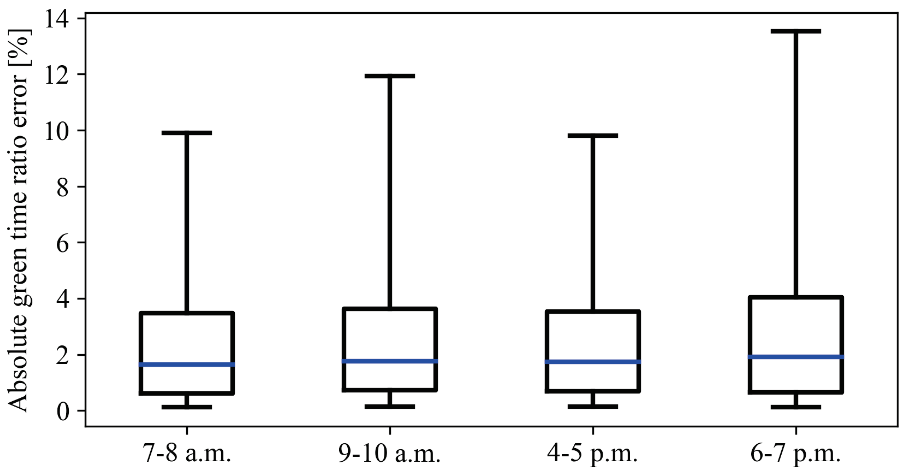

Figure 4: Distributions of absolute errors of the ratio of green time duration with respect to the cycle time

The determined median values of the deviation between simulated and observed green time ratios lie between 1.6 % and 1.9 %. The values indicate that the emulated traffic light programs are capable to adapt a realistic traffic light behavior for different time intervals. The distributions of errors in peak traffic hours from 7 a.m. to 8 a.m. and 4 p.m. to 5 p.m. range from 0 % to 10 %. In time periods with a lower traffic demand between 9 a.m. and 10 a.m. as well as 6 p.m. and 7 p.m., the proportion of public transport, which strongly influences the switching order and green time duration, is larger compared to the other hours. This increased variety of real traffic light programs leads to a wider distribution of the error value.

In the calibration process of traffic volumes in microscopic traffic simulation, traffic lights are often tuned to match the calibrated demand. In this publication, the traffic lights are created from real-world data together with the traffic counts. Inaccuracies in the calibration procedure will therefore quickly show in the simulation, since an overestimation of traffic volumes or an underestimation of the capacities at signalized intersections will lead to unrealistic congestions. However, the calibrated simulation shows that the emulated traffic lights are able to cope with the traffic demand even in peak hours without leading to excessive congestion or deadlock situations in the simulated network. This capability further indicates a realistic representation of traffic lights in Ingolstadt.

4.3 Traffic assignment

Traffic volumes generated by the proposed approach are validated by a comparison of simulated traffic counts with their real counterparts both aggregated on road level. In reality, there exist only measurements at specific loop detector locations so that only the counts of these measurement points from reality and simulation can be compared. Although these data are used for calibration of the simulation, there is no further available data for the validation purpose of traffic flows.

For the validity analysis of the simulated traffic counts, the absolute value of the relative error e_{rel} is calculated according to equation (4):

$$e_{rel} = \frac{|count_{real} - count_{sim}|}{count_{real}}, \tag{4}$$

where $count_{real}$ represents the real-world traffic count and $count_{sim}$ is its simulated counterpart both aggregated road-wise for each of the 254 roads containing detectors.

The proposed calibration approach uses detector counts of one hour to calibrate routes starting within the corresponding hour in order to meet the counts again. Consequently, the procedure leads to an offset of routes, which start in the corresponding hour but cannot finish during the one-hour interval. The issue especially arises in peak traffic, as the higher number of vehicles potentially leads to more congestion, which increases travel time and causes incomplete trips. Therefore, simulated traffic counts are evaluated in a prolonged timespan, which enables all generated vehicles to complete their journey through the network. Figure 5 shows the relative errors for one week from Monday to Friday which are separated hourly in time spans from 7 a.m. to 8 a.m., 9 a.m. to 10 a.m., 4 p.m. to 5 p.m. and 6 p.m. to 7 p.m.

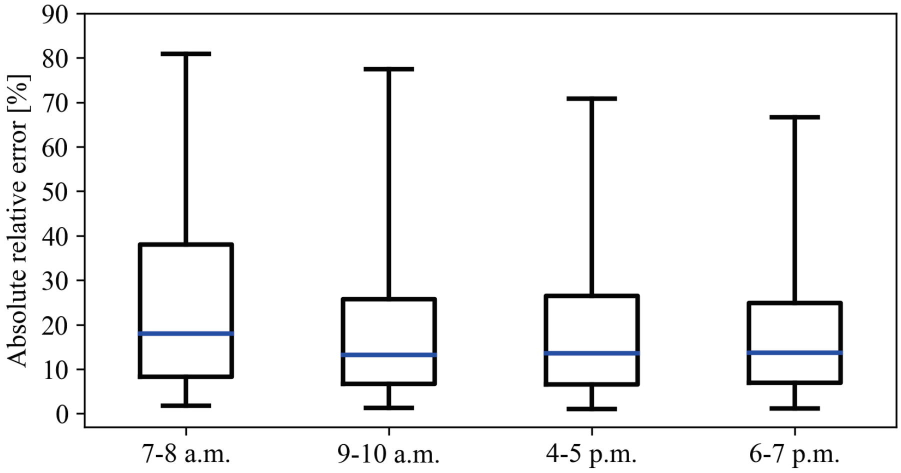

Figure 5: Distributions of absolute relative errors of simulated traffic counts aggregated to road level after a clearing period

All distributions presented in Figure 5 show median values between 13 % and 18 %. The error distribution of values of peak traffic in the afternoon between 4 p.m. to 5 p.m. shows similar results compared to the hours from 9 a.m. to 10 a.m. and 6 p.m. to 7 p.m. with lower demand. Only the distribution for peak traffic in the morning from 7 a.m. to 8 a.m. reveals higher error values. There are mainly two reasons for the remaining errors, which result from the calibration, compared to reality, i.e. corrupt input data and rerouting. The calibration procedure balances the over- and underestimation of routes passing the detector loops in order to achieve an overall error of 0. However, the observed traffic counts show several unrealistically low values, which strongly influence the calibration. Given an exemplary road network with three junctions linking three successive main roads in a line each containing detectors, unrealistically low detector counts of the road in between the two other roads would induce the traffic assignment algorithm to adjust an underestimation of the two roads with correct detector counts and an overestimation of the road with the lower counts. This would lead to a balanced overall error but reveals in an increased absolute error. Especially in peak traffic hours, overestimation of traffic exceeds the capacity of a share of roads resulting in congestion and deadlock situations. In order to cope with the overestimation induced by inaccurate input data, the rerouting probability of SUMO vehicles is adapted to 20 %, which enables 20 % of the vehicles to take different routes than assigned to reach their destinations. In comparison, Lobo et al. [25] apply a more than four times higher rerouting probability to handle congestion in the network.

The absolute values of the observed traffic counts vary over a wide range. However, absolute errors of small traffic count values add to higher percentage deviations than equal absolute errors of high counts. Therefore, GEH analysis is applied additionally in order to assess the calibration performance under consideration of a non-linear weighting of errors. The GEH statistic is given by

$$\mathrm{GEH} = \sqrt{\frac{2 \cdot (\mathrm{count}_{\mathrm{real}} - \mathrm{count}_{\mathrm{sim}})^2}{\mathrm{count}_{\mathrm{real}} + \mathrm{count}_{\mathrm{sim}}}}. \tag{5}$$

Figure 6 depicts the distributions resulting from GEH analysis regarding a clearing period.

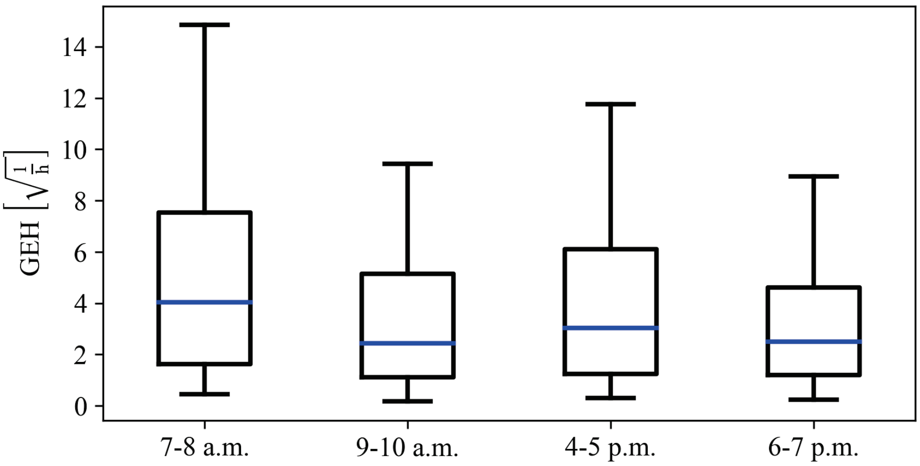

Figure 6: Distributions of GEH analysis of simulated traffic counts aggregated to road level after a clearing period

Figure 6 reveals wider distributions of the GEH statistic in peak traffic hours from 7 a.m. to 8 a.m. and 4 p.m. to 5 p.m. than for the remaining hours with less traffic. The median GEH values between 2.4 and 4 indicate a good performance of the calibration procedure.

In order to tackle the issue regarding the offset of routes when starting the simulation under initial conditions without any traffic, additionally, the impact of a one-hour warm-up phase without a clearing period is investigated. The warm-up on the one hand fills up the network at the beginning of each analyzed interval. On the other hand, the overlapping traffic is originally calibrated for the previous time interval. Figure 7 displays the resulting distributions of the absolute value of relative errors of the traffic counts for the previously specified time intervals regarding a warm-up phase.

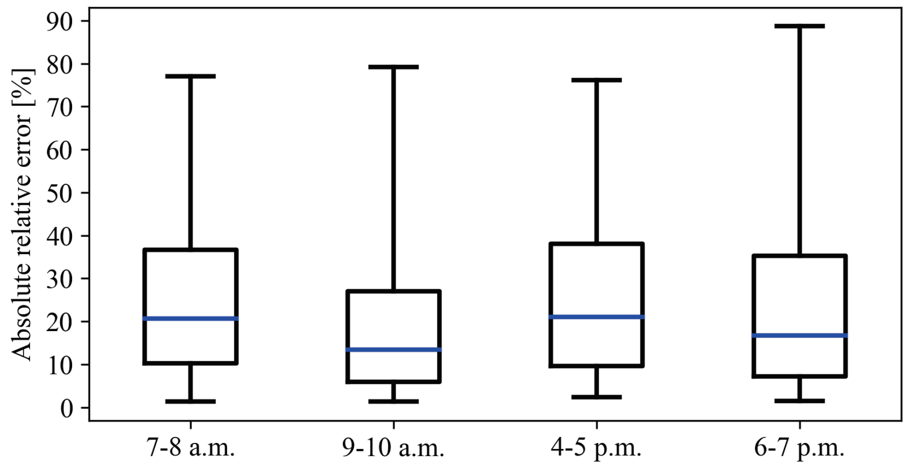

Figure 7: Distributions of absolute relative errors of simulated traffic counts aggregated to road level with a previous warm-up phase

With an initial warm-up phase, small deviations between traffic in peak hours and traffic with lower demand are visible. Compared to Figure 5, the slightly higher error values when considering a warm-up phase might result from different traffic volumes in the corresponding two consecutive time intervals, which increase the observed error. Hence, the traffic from 9 a.m. to 10 a.m. shows similar characteristics compared to the previous hour from 8 a.m. to 9 a.m. and therefore yields the lowest error. In order to improve the presented calibration method, the temporal offset should be considered by a partial overlap of calibrated routes beyond the corresponding hour.

5 Conclusion

In this work, a scalable approach for the calibration of a microscopic traffic simulation based on real-world observation data for a requested time interval was presented. Therefore, information of traffic lights and loop detectors for this time interval was assigned to the virtual world. The inbound roads and lanes to intersections were associated correctly to the map by a robust determination of inbound roads and pairing of lanes. However, the detector placement was error prone affected by map inaccuracies and the number of lanes exceeding the number of detectors. Traffic light programs were emulated as fixed-time traffic lights and implemented to the simulation. Simulated traffic lights were evaluated regarding the ratio of green times with respect to the cycle times showing median errors between 1.6 % and 1.9 %. Furthermore, emulated traffic lights proved the ability to cope with the calibrated traffic demand. For the calibration of traffic flows, a two-step procedure was introduced. In a first step, an initial pool of routes was generated based on OD relations determined from statistics. In a second step, the routes in the pool were copied twice and subsequently removed in order to meet the observed traffic counts in the network. The procedure exhibited an offset of a share of routes calibrated with traffic counts of a one-hour interval, which could not be completed within the interval. Considering a clearing period for vehicles to finish their trips, the evaluation resulted in median absolute error values between 13 % and 18 % with respect to observed traffic counts.

In the future, we will apply a time component to the traffic flow calibration in order to shift calibrated routes to the corresponding previous hour. Further improvement can potentially be achieved by a more comprehensive and diverse initial pool of routes. Additionally, different optimization algorithms like Cadyts [31]–[33] can be applied for a comparison of traffic flow calibration procedures. In order to increase the accuracy of the simulation, fixed-time traffic lights must be replaced by a traffic adaptive traffic light control using rule-based or data-driven techniques. Finally, individual driver behavior has a strong impact on the whole traffic system. Therefore, we will analyze and calibrate driver models with real-world data to achieve a more accurate representation of the real world.

We plan to publish the simulation of Ingolstadt as an open-source project to make it available for further applications especially regarding testing and development of automated driving. Furthermore, we intend to publish the code framework for map-matching to the SUMO network as a Python library based on geopandas [26], which provides a broad range of possible matching applications.

6 Acknowledgements

This work is supported by the German Federal Ministry of Transport and Digital Infrastructure (BMVI) within the Automated and Connected Driving funding program under grant no. 16AVF2145A (SAVe) with a funding amount of 3.885.755,00 €.

The authors thank the colleagues from GEVAS, TTS and AUDI AG for their support and for providing the traffic count and traffic light data, on which the calibration is based on.

References

[1] M. Semrau, "Untersuchung zur Modellierung von chinesischem Fahrverhalten auf Autobahnen für den Test pilotierter Fahrfunktionen," TU Braunschweig, 2018.

[2] M. Semrau, J. Erdmann, B. Friedrich, and R. Waldmann, "Simulation framework for testing ADAS in Chinese traffic situations," *SUMO 2016 Conference Proceedings*, pp. 103–113, 2016.

[3] J. Olstam and R. Elyasi-Pour, "Combining traffic and vehicle simulation for enhanced evaluations of powertrain related ADAS for trucks," *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 851–856, 2013.

[4] D. Krajzewicz, A. Richter, M. Behrisch, and J. Erdmann, "Abbildung des Umgebungsverkehrs in einem Fahrsimulator," pp. 1–10, 2013.

[5] P. A. Lopez *et al.*, "Microscopic Traffic Simulation using SUMO," *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 2575–2582, 2018.

[6] M. A. Quddus, W. Y. Ochieng, and R. B. Noland, "Current map-matching algorithms for transport applications: State-of-the art and future research directions," *Transportation Research Part C: Emerging Technologies*, vol. 15, no. 5, pp. 312–328, 2007.

[7] D. Bernstein and A. Kornhauser, "An introduction to map matching for personal navigation assistants," *Technology*, pp. 587–602, 1996.

[8] M. Schönfelder, V. Protschky, and T. Bäck, "Reconstructing fixed time traffic light cycles by camera data analytics," *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 1–7, 2018.

[9] S. Axer and B. Friedrich, "A Methodology for Signal Timing Estimation Based on Low Frequency Floating Car Data: Analysis of Needed Sample Sizes and Influencing Factors," *Transportation Research Procedia*, vol. 15, pp. 220–232, 2016.

[10] S. Axer and B. Friedrich, "Signal timing estimation based on low frequency floating car data," *Transportation Research Procedia*, vol. 25, pp. 1645–1661, 2017.

[11] C. Wang and S. Jiang, "Traffic signal phases' estimation by floating car data," *2012 12th International Conference on ITS Telecommunications, ITST 2012*, pp. 568–573, 2012.

[12] M. Rostami-Shahrbabaki, K. Bogenberger, A. A. Safavi, and A. Moemeni, "Intersection SPaT estimation by means of single-source connected cehicle data," *TRB 99th Annual Meeting*, 2020.

[13] J. C. Wolf, M. Zweck, and C. Rucker, "Traffic Signal Emulation Using Genetic Algorithm," US Patent 10,255,805, 2019.

[14] T. Weisheit, "Algorithmenentwicklung zur Prognose von Schaltzeitpunkten an verkehrsabhangigen Lichtsignalanlagen," *Heureka*, vol. FGSV 002/1, pp. 320–339, 2014.

[15] Z. Shen, K. Wang, and F. Zhu, "Agent-based traffic simulation and traffic signal timing optimization with GPU," *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 145–150, 2011.

[16] R. Blokpoel and J. Vreeswijk, "Uses of Probe Vehicle Data in Traffic Light Control,"

Transportation Research Procedia, vol. 14, pp. 4572–4581, 2016.
[17] S. S. Mousavi, M. Schukat, and E. Howley, "Traffic light control using deep policy-gradient and value-function-based reinforcement learning," *IET Intelligent Transport Systems*, vol. 11, no. 7, pp. 417–423, 2017.
[18] A. Hussain, T. Wang, and C. Jiahua, "Optimizing Traffic Lights with Multi-agent Deep Reinforcement Learning and V2X communication," no. 1, pp. 1–6, 2020.
[19] M. G. McNally, "The Four Step Model," *Handbook of Transport Modelling*, pp. 35–52, 2000.
[20] H. J. Van Zuylen and L. G. Willumsen, "The most likely trip matrix estimated from traffic counts," *Transportation Research Part B: Methodological*, vol. 14, no. 3, pp. 281–293, 1980.
[21] M. G. H. Bell, "The estimation of origin-destination matrices by constrained generalised least squares," *Transportation Research Part B: Methodological*, vol. 25, no. 1, pp. 13–22, 1991.
[22] E. Cascetta, "Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator," *Transportation Research Part B*, vol. 18, no. 4–5, pp. 289–299, 1984.
[23] M. J. Maher, "Inferences on trip matrices from observations on link volumes: A Bayesian statistical approach," *Transportation Research Part B: Methodological*, vol. 17, no. 6, pp. 435–447, 1983.
[24] H. Spiess, "A maximum likelihood model for estimating origin-destination matrices," *Transportation Research Part B: Methodological*, vol. 21, no. 5, pp. 395–412, 1987.
[25] S. C. Lobo, S. Neumeier, E. M. G. Fernandez, and C. Facchi, "InTAS -- The Ingolstadt Traffic Scenario for SUMO," in *SUMO User Conference 2020*, 2020, pp. 1–20.
[26] "Geopandas," 2021. [Online]. Available: https://geopandas.org/. [Accessed: 17-May-2021].
[27] A. V. Forschungsgesellschaft für Straßen- und Verkehrswesen -FGSV- Köln, Ed., *Richtlinien für Lichtsignalanlagen. RiLSA. Lichtzeichenanlagen für den Straßenverkehr.* Germany, 2010.
[28] AUDI AG, "Audi Werk Ingolstadt." [Online]. Available: https://www.audi-mediacenter.com/de/ingolstadt-195. [Accessed: 31-Jul-2020].
[29] Stadt Ingolstadt, "Ingolstadt in Zahlen 2018/2019," 2019. [Online]. Available: https://www.ingolstadt.de/media/custom/465_1995_1.PDF?1533818349. [Accessed: 31-Jul-2020].
[30] infas Institut für angewandte Sozialwissenschaft GmbH, "Mobilität in Tabellen (MiT 2017)," 2017. [Online]. Available: https://mobilitaet-in-tabellen.dlr.de/mit/login.html?brd. [Accessed: 20-Jul-2020].
[31] G. Flötteröd, M. Bierlaire, and K. Nagel, "Bayesian calibration of dynamic traffic simulations," 2009.
[32] G. Flötteröd, Y. Chen, and K. Nagel, "Behavioral Calibration and Analysis of a Large-Scale Travel Microsimulation," *Networks and Spatial Economics*, vol. 12, no. 4, pp. 481–502, 2012.
[33] G. Flötteröd and R. Liu, "Disaggregate path flow estimation in an iterated dynamic traffic assignment microsimulation," *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, vol. 18, no. 2, pp. 204–214, 2014.