

Online calibration with SUMO for network-wide traffic and emission monitoring – Case study ITS Huainan

Yun-Pang Flötteröd¹ and Michael Behrisch¹

¹Institute of Transportation Systems, German Aerospace Center
yun-pang.floetteroed@dlr.de, michael.behrisch@dlr.de

Abstract

Currently, the city of Huainan, China, is constructing its intelligent transportation system, and traffic and environmental monitoring system for efficiently and effectively monitoring and managing city traffic. DLR's traffic information platform KeepMoving is adopted in the aforementioned system, where the existing online calibration module with SUMO has been extended in order to provide current and predicted traffic states and the resultant emission information in real-time. Accordingly, a comprehensive city-wide traffic situation can be captured and shown at the KeepMoving portal, as a decision support tool for traffic management personnel. The comparison between the real and simulated data shows a promising calibration result.

1 Introduction

Rapid economic growth and technology development facilitate people's mobilities and expand their life and social cycles. Accordingly, the usage of private vehicles has continually increased since years and resulted in traffic congestions especially in dense urban areas. Different traffic management strategies have been applied with use of real-time data collected by different kinds of sensors, such as loop-based, microwave-based or camera-based detectors, drones and floating car data (FCD). With regard to high investment and maintenance costs the aforementioned sensors are often only located at limited places with heavy traffic. Therefore, real-time traffic states can be monitored locally, but not networkwide. To deal with this issue, traffic simulation tools have been applied as a part of traffic monitoring and management system in some cities.

As known, good data, such as traffic demand, road network, signal control and traffic management strategies, is the basic prerequisite for setting up a representative traffic simulation. Normally, it takes lots of time to generate/update general traffic demand data when using household survey methods. Moreover, traffic demand varies from time to time due to travelers' decisions on departure time, transport mode, destination and their trip purposes when observing it at a finer level. Once there are

changes in urban planning, area development and network structure, traffic demand and its distribution will also be significantly affected. Such situations occur quite often in many cities in China currently. In this case, it is difficult to solely use traffic simulation as complement for network-wide traffic monitoring. Online simulation and calibration with real-time data is thus a great extension to the traffic managers toolbox. Following several research works executed within the DLR fundamental projects, such as DELPHI, VABENE and VABENE++ (Behrisch, et al., 2010) (Erdmann, 2012) (Bieker, Behrisch, & Ruppe, 2012)) and the on-going DLR project D.Move, SUMO (Alvarez Lopez, et al., 2018) can be integrated into a real-time traffic information platform and provide both real-time simulated and predicted traffic data. Moreover, a module to online calibrate flows and speeds with use of sensor and FCD data is implemented together with data processing.

This paper describes the simulation part of the traffic and environmental monitoring system, which efficiently and effectively monitors and prospectively manages the city traffic of Huainan, China. Moreover, the data processing part is also briefly explained, since its resultant data is used for calibration and prediction. DLR's traffic information platform KeepMoving (Brockfeld, 2014) is adopted in the aforementioned system, where SUMO should provide calibrated traffic data and the resultant emission information for real-time and predicted traffic state especially at places where no sensors and FCD are available. In order to achieve the goal, the online calibration module is modified for fitting the information system and the used Oracle database is extended for generating and delivering emission information. In the following, the framework and the concept used for data processing and online calibration is firstly introduced. The data preparation for simulating the traffic in Huainan is then explained. After that, the calibration results are presented, and the conclusions and remarks are then made at the end.

2 Framework and methodology

The applied online process consists of two parts, i.e. data processing and simulation with calibration and prediction. Their workflows are illustrated in Figure 1 and Figure 2 respectively. These two components are executed by the same script, but with different option settings. Besides, they can share the same configuration file, in which many settings can be defined, such as time period, data type(s), data receiving interval and aggregation interval, data duplication check, simulation inputs, calibration data types, data quality threshold, interval, simulation type (micro/meso), data cleaning interval, prediction, database connection. A database, e.g. PostgreSQL or Oracle, is required for data processing, exchange and output saving. The whole process chain is implemented in Python, and has been customized and enhanced for the Huainan ITS system.

The data process part include data correction, data aggregation, data fusion and data extrapolation. Currently, edge-based FCD and traffic measurements (speed and flow) either from stationary sensors or other data providers can be considered in the data processing, where flow types include passenger cars and trucks. As mentioned before, data processing and online calibration and prediction needs to be started separately. In the former case, speed average at each detector group is weighted with respect to the number of flows, and the interval-based data quality indicator is calculated according to the number of available data records in a pre-defined interval. If two or more detector groups locate at the same edge, average values will be derived and used. In the latter case, required data and data format for calibration needs to be in the pre-defined tables of the respective database. Moreover, route-based daily traffic demand is required as base. The related edge-based route sets/distributions and traffic demand per pre-defined interval for calibration are then generated accordingly and used together with the given network as inputs in simulation. In addition, detector group locations and the corresponding location-matching in the investigated sumo network are also required so that SUMO knows exactly where speeds and flows should be calibrated.

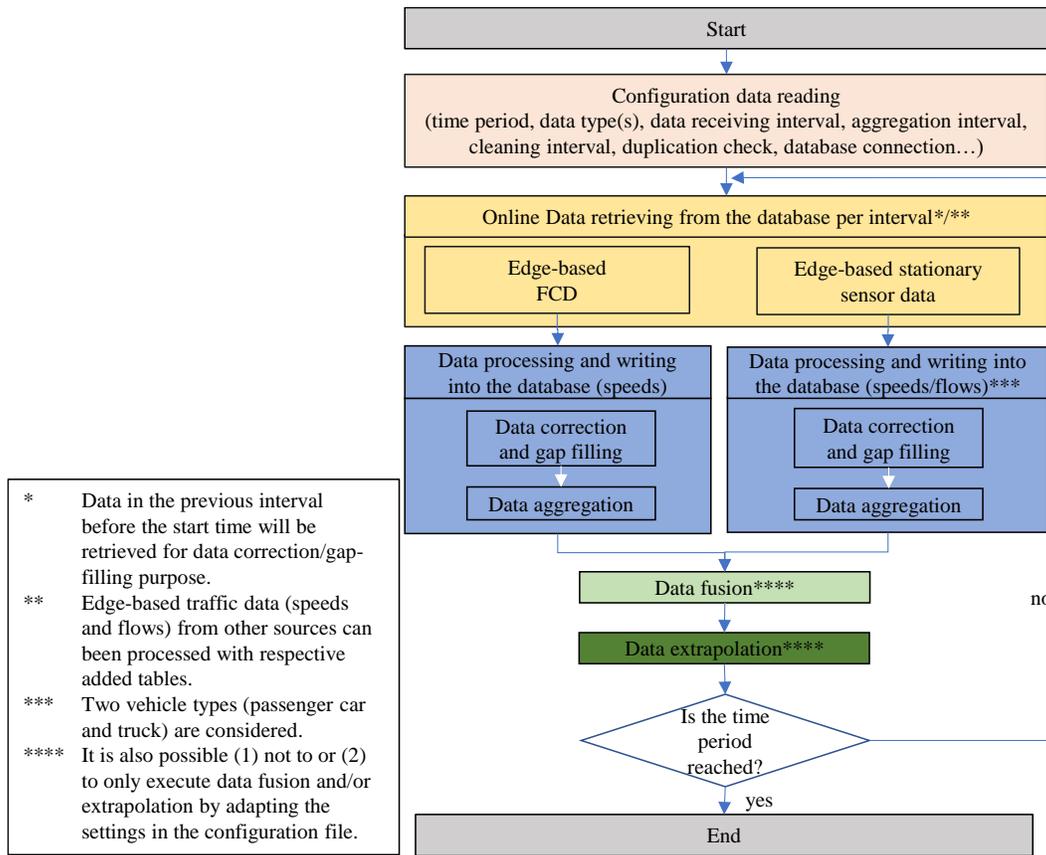


Figure 1: Overview of the online data correction and gap-filling process

When the whole process is started, data availability is checked during data processing. If data does not arrive in the database in time, the data process will be hold until the pre-defined time threshold is reached. Once data is prepared and filtered with the given quality indicator, the simulation will use the simulated traffic state from the previous interval as reference and try to get flow and speed data per interval in real-time for calibration, where it is possible to choose to use aggregated or fused data. For traffic prediction, data extrapolated from the fused data is used. If no data is available, simulation will still continue to run without calibration. With these data the on-line calibration will be executed for each pre-defined time interval, e.g. 10 minutes. Based on that, traffic prediction will be made for the pre-defined duration, e.g. 30 minutes. When travelling speeds in the simulation differ from the respective measured edge speed averages, the related maximal speeds will be adjusted in the corresponding intervals accordingly. When simulated flows are greater than the related measured flows, route sampling actions will be executed in the respective edge-based route sets. Vehicles with the sampled routes and the current edge travelling speeds will then be inserted to the corresponding edges in the simulation for eliminating the differences between the simulated and measured flows. If the flow situation is the other way around, vehicles will be sampled and then directly removed from the network. SUMO tries to delay the insertion process of new vehicles as long as possible in order not to force the

later removal of vehicles coming from the “real” demand (are not a calibration result). In addition, failure on vehicle insertion could happen when no space on the respective edges is available.

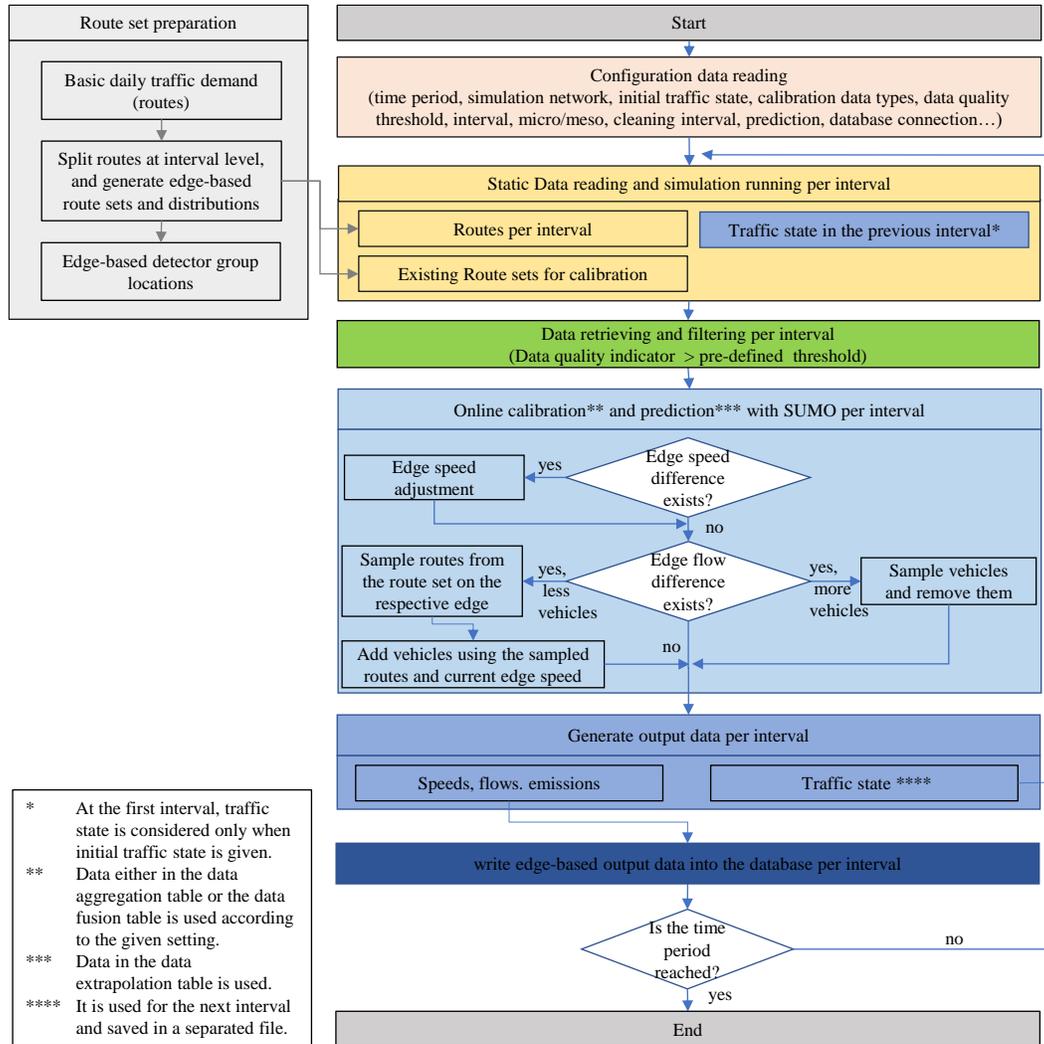


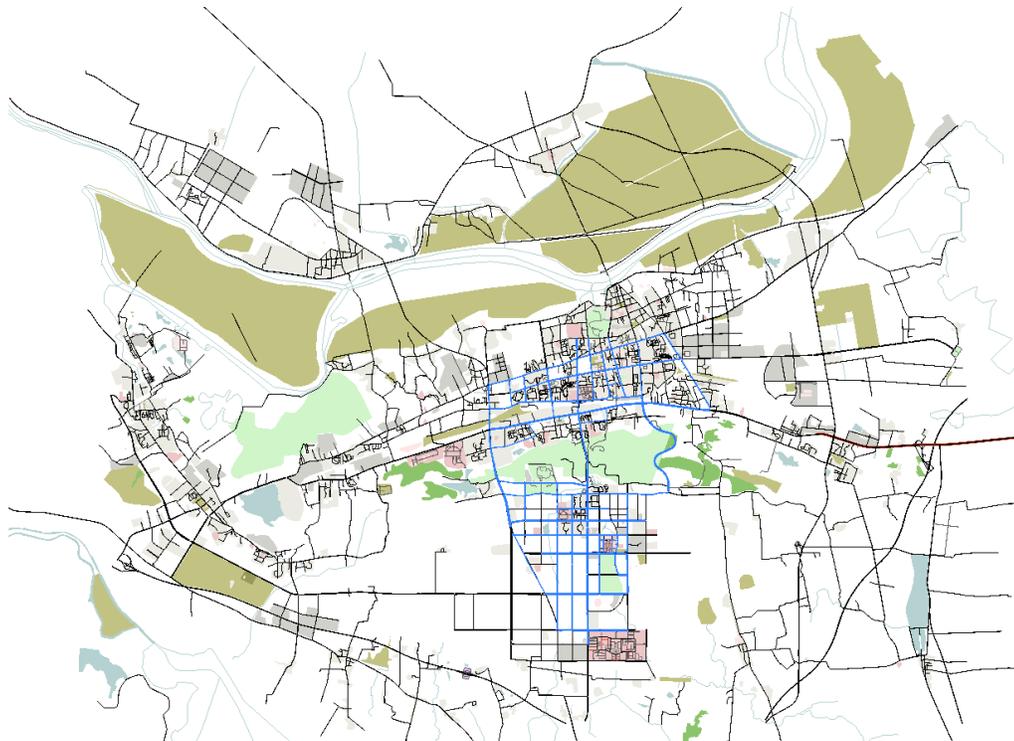
Figure 2: Overview of the online simulation and calibration process

As output, the simulated and calibrated data about traffic efficiency (flows and speeds) and emissions (CO₂, CO, NO_x, PM_x, HC) will be directly written into the database every interval according to the customized setting. Traffic state of each interval will be saved on the used server as well and used as reference for the simulation in the coming interval. The whole process will be iteratively executed until the pre-defined time period is reached. The simulated and predicted data can be used as basis for complementing real-time traffic state in the city and visualized in a selected platform, such as the KeepMoving Portal. In order to keep the whole processing chain continually running, a mechanism is built so that the processing chain will be automatically started if the server is restarted for any reason.

3 Data preparation

3.1 Simulation network

The OpenStreetMap (OSM) database, released in the 3rd quarter, 2019, was used as base for setting up the simulation network with SUMO's netconvert. Due to the limited road geometry and signal information in OSM the network has been further manually adjusted according to satellite pictures from Google Maps and images from Baidu Maps. Moreover, detailed geometry information for the roads in the old town and new central areas, obtained from the project partner, has been used for refining the network. The overview of the simulation network and the detailed road geometry example are shown in Figure 3 and Figure 4 **Fehler! Verweisquelle konnte nicht gefunden werden.** respectively. The total network contains about 10000 edges and 5000 nodes with a total length of about 2600 km.



* the available detailed road geometry information is marked in blue

Figure 3: Simulation network layout

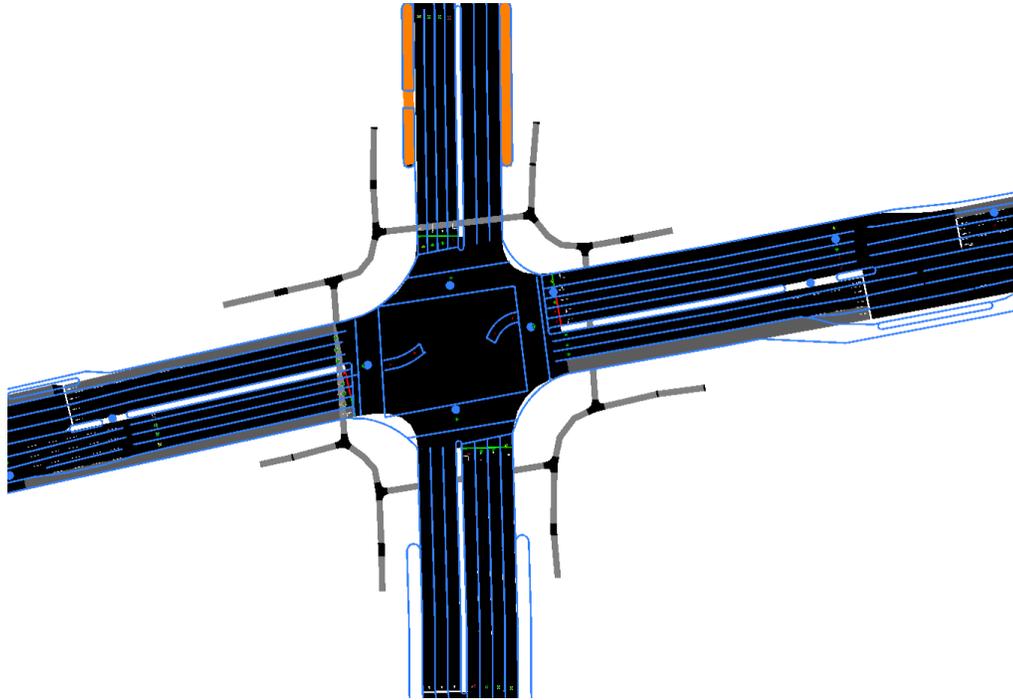


Figure 4: Illustration example of an intersection with detailed geometry information

3.2 Traffic demand and zone connections

The OD-Matrix data for the Huainan City is available and is generated by the transportation planning work, conducted in 2009 (Huainan City Planning Bureau; School of Transportation of the Southeast University, 2010). 81 traffic analysis zones (TAZ) were defined, as shown in Figure 5. An OD matrix prediction for 2020 is also available. However, this OD matrix is based on 26 TAZ, aggregated from the aforementioned 81 TAZ and for one peak hour period. Based on the estimated peak-hour traffic demand for 2020 and the populations of the small traffic analysis zones, used in the transportation planning from 2009, the peak-hour traffic demand was disaggregated into 81 traffic analysis zones. With the time-series data from the detectors 5-min flow distributions both weekdays and weekends were generated. Together with these flow distributions along a whole day and the relevant survey result, made in 2009, the daily traffic demand was derived and the departure times of the trips in each zone were then refined accordingly. Totally, the number of estimated daily trips is around 1.34 million.

Moreover, traffic analysis zone data is also available and converted from the transportation planning software TransCAD (Caliper Mapping & Transportation Software Solutions, 2021) format to the shapefile format. When further converting the shape file into SUMO format, the map is distorted to a certain degree and the respective correction work had been carried out manually. Together with the prepared network and the TAZ polygons the roads with a maximum travelling speed 50 km/h in each zone are selected with SUMO's tool `edgesInDistrict.py` and used as zone connections for traffic assignment. Some connections in TAZ 63, 64, 65 and 66 were manually adjusted due to the restricted road accesses.

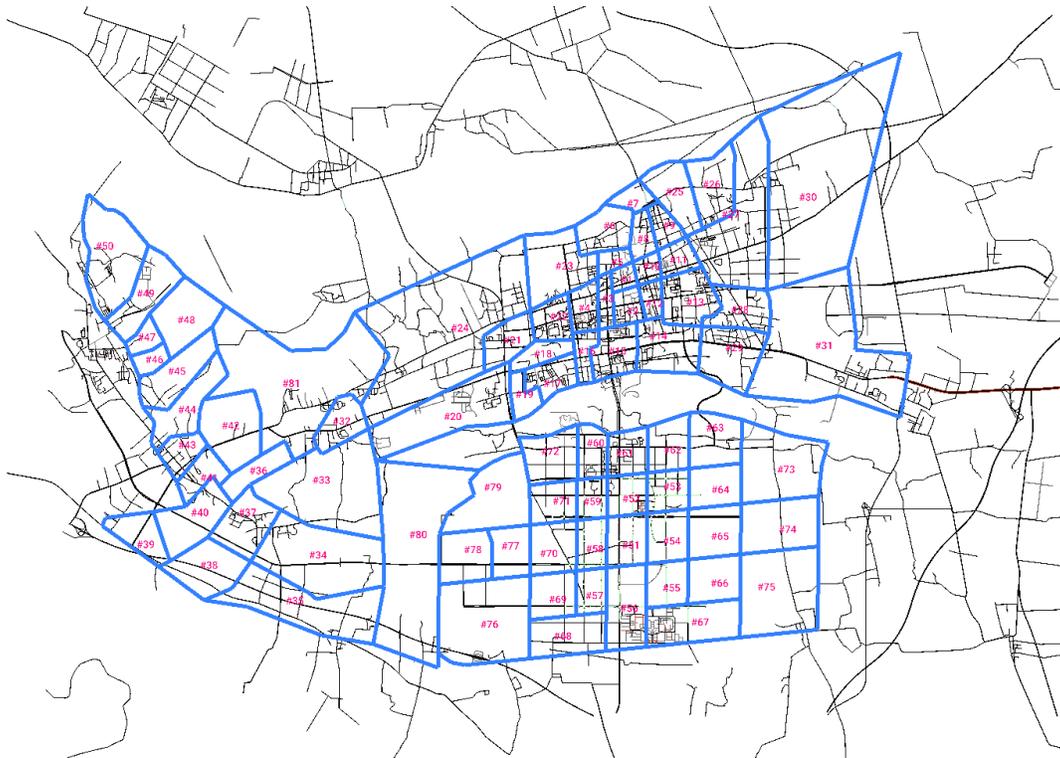
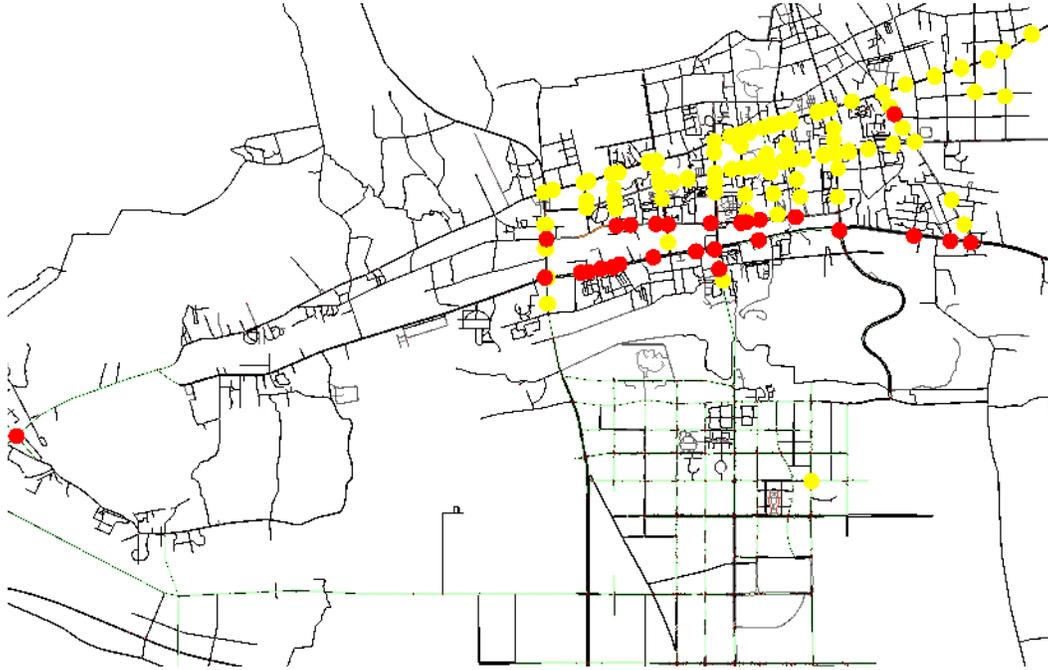


Figure 5: Distribution of the 81 traffic analysis zones in the City Huainan

3.3 Stationary sensor data

In order to enable the simulation to consider sensor data for calibration it is necessary to match the sensors to the respective lanes in the simulation network. Currently, 709 sensors are connected to the database. 509 of them are microwave-based sensors and the rest of the sensors are video-camera based sensors. Each detector is either deployed on or connected to a lane and detected data is transmitted back to the database every minute. Detectors on the same edge and closed to each other are grouped as a detector group. Figure 6 gives an overview of the detector groups' locations and indicates that most of the detector groups are in the old down town area. Only one detector group locates in the southern and western area and the new central area respectively. It is expected to have some more sensors installed in the new central area when the respective development and construction works are finished. It implies that the simulation and calibration performance with respect to the reality will be severely limited outside the old town area.



*: red ones are video-camera based sensors, while yellow ones are microwave-based sensors.

Figure 6: Locations of the stationary sensor groups

4 Simulation setting and execution

With the concern of the real-time operational readiness mesoscopic simulation is applied (Eissfeldt, 2004). In order to get route-based daily traffic demand a traffic assignment with the trip-based traffic demand, TAZ information, zone connections, described in Section 3.2, and DUAROUTER was executed. The resultant route file was then further used as described in Section 2. Based on the pre-defined interval, the given daily route file will be split into several files, i.e. 900 sec for the case study ITS Huainan. Only the temporally corresponding route files will be used in each simulation interval. Moreover, the routes and the route distribution on each calibrated edge will be generated as well.

When executing the simulation, a python script, as main script, is called to start the whole calibration process, where the definition of database tables and the general functions to get schemas, insert data into the database, update data and compare data can be found there. A configure file will then be read for mainly getting the following information:

1. start and end simulation/calibration times,
2. input files and their locations,
3. the intervals for data processing, calibration and prediction,
4. maximum delay for receiving sensor data,
5. data source for calibration, i.e. data either from FCD, sensors or data fusion result,

6. calibration parameter, i.e. only flow or both flow and speed,
7. action, i.e. either data processing or simulation with calibration and prediction,
8. data quality threshold,
9. look-back time for prediction,
10. prediction action and period
11. database connection information, as well as
12. durations to keep the data in the database and the output files on the used server.

For the data processing and simulation in the Huainan project, the intervals in Point 3 are set to 5 minutes, which corresponds to the interval used at the KeepMoving platform. 2 minutes is used as data transmission delay, so that the data processing will be hold at most for 2 minutes if corresponding data does not arrive in the database. Moreover, both fused speed and flow data is selected as data source for calibration. When running the simulation, the traffic state for each interval is saved and used as input for the next simulation interval in order to properly reflect the respective traffic situation in the network. Regarding traffic prediction, the look-back time for prediction and the prediction period are 10 and 30 minutes respectively. It means that traffic prediction at 5-min interval for 30 minutes will be made according to every past 10-min simulation result. With the concern of the limited reliability of basic daily traffic demand, due to on-going road constructions and urban development in Huainan, only data with a quality indicator larger than 0 is considered in the online calibration which will exclude all completely disabled detectors.

Currently, the maximum experienced simulation duration for each 5- min interval is around 130 s, whilst the common simulation duration is between 70 and 90 sec. The total required time from data processing to the output, writing to the database, is less than 5 minutes, corresponding to the expected performance.

5 Preliminary results

The establishment of the whole ITS system and environmental monitoring system is newly finished. Currently, it is under the testing and maintenance phase. 2 weeks data from 2021.01.14 to 2021.01.31 has been used to examine the calibration performance. It is noticed that online 5-min time-series data is not yet entirely available for most of detectors currently. According to the used data set, it is quite common that data gaps exist for some short periods. Sometimes, data gaps also happen for longer periods. When looking at the calibration result at 5-min interval, around 80% of the calibrated flows and more than 95% of the calibrated speeds have an absolute relative error (ARE) less than or equal to 15%. At the hourly flow level, the GEH statistic is used to evaluate the hourly flows, aggregated from the 5-min flows. The result shows that around 92% of the calibrated flows have a GEH-value less than or equal to 5, which indicates a good fit.

Moreover, four edges are selected as examples for further investigating the calibration performance along a whole day. In Figure 7 (a) and (b), it shows that, as expected, the simulation and calibration result is quite promising for both flow and speed, even when speed fluctuates largely. When traffic congestion raises in the simulation, the calibrator cannot insert more vehicles into the network for reflecting the flows in the reality. The simulated flows are then lower than the measured ones, such as the situation indicated in Figure 7 (c). Such case may sometimes occur due to other reasons, such as that travel speeds are set too low due to the low accuracy of detected speeds and lower quality threshold. More investigation on it is then needed. Figure 7 (d) shows that the sensor data is only partly available

(until the early afternoon). With the FCD supplement, the related edge speeds in the late afternoon can still be properly calibrated.

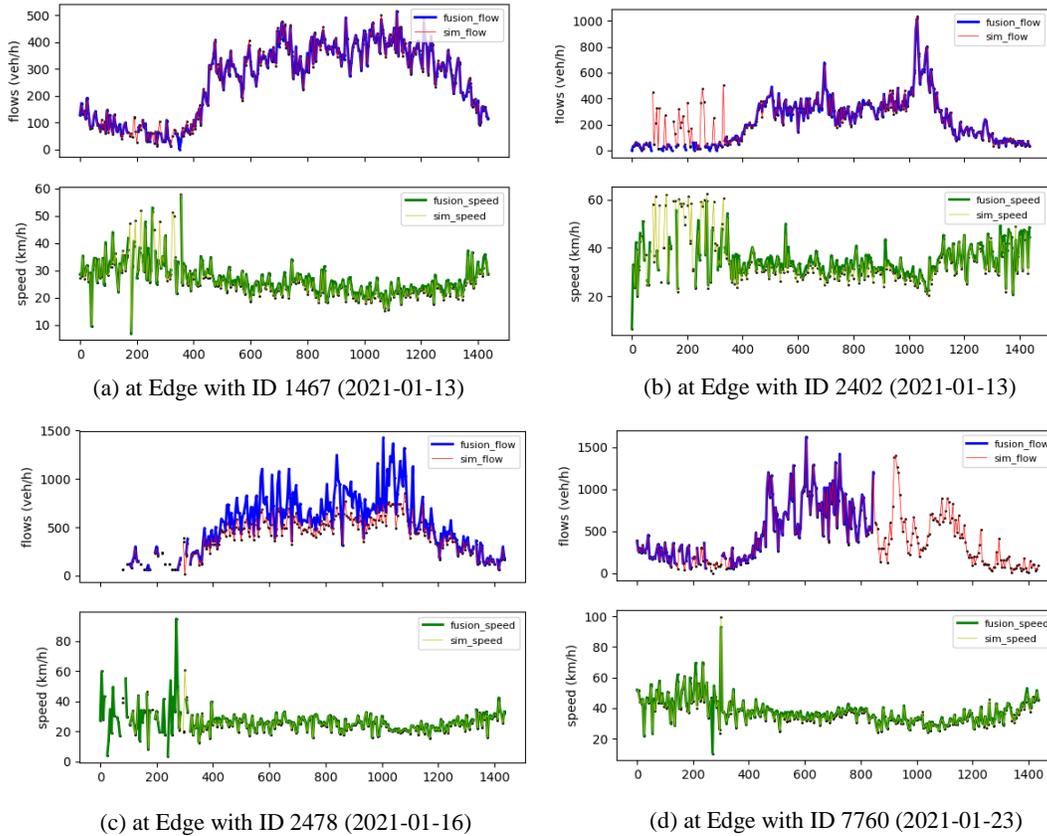


Figure 7: Comparison of real and simulated data for the selected edges

6 Conclusion and remarks

An online continuing calibration and prediction process with SUMO and its enhancement has been described in this paper. Generally speaking, the simulation keeps a defensive manner when conducting online calibration, i.e. the simulated state (flow and speed) is preferred to be untouched until the end of each interval is approached. This mechanism may result in failed vehicle insertion and affect the calculation of the number of vehicles in the respective intervals sometimes. For example, two vehicles need to be inserted at Interval x . To be sure that no other vehicles will be going to enter the observed edge in Interval x , a vehicle insertion action will not be done until the end of Interval x . Sometimes, such action cannot be carried out due to lack of space on the related edge, and affects calibration performance. Under this circumstance, the current analysis result indicates that the calibration and the overall computation performances are principally still quite well. Some larger fluctuation in speed occurs, especially at late night or very early morning. More related investigation is needed.

City Huainan is continuously realizing its urban planning and development as well as completing its road network system. Several road construction projects with ITS based infrastructure are on-going. It implies that the related parts in the simulation network need to be continuously updated accordingly for reflecting actual traffic state. Moreover, more real-time traffic information should be then available. It is expected that the simulation result can be further improved later on. Due to the existing and coming changes in land use and road network it is indeed necessary to update the daily traffic demand data and, if necessary, the respective TAZ definition either with the conventional demand modelling method or/and with other data sources, such as data from OSM, Wikidata, mobile phones, social media and navigation systems. The overall online calibration and prediction performance will also be benefited by the updated daily traffic demand.

References

- Alvarez Lopez, P., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., . . . Wießner, E. (2018). Microscopic Traffic Simulation using SUMO. *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, (S. 2575-2582). Maui, USA.
- Behrisch, M., Hohloch, M., Junghans, M., Kuhns, G., Krajzewicz, D., & Wang, Y.-P. (2010). Traffic management decision support based on on-line data. *POLIS CONFERENCE - Innovation in transport for sustainable cities and regions*.
- Bieker, L., Behrisch, M., & Ruppe, S. (2012). EmerT – a web based decision support tool for Traffic Management. *19th ITS World Congress 2012, 22-26 October 2012*. Vienna, Austria.
- Brockfeld, E. (2014). Der Reiseassistent KeepMoving. *5. Tagung Mobilitätsmanagement von Morgen "Auf dem Weg zur emissionsarmen Mobilität"*. Berlin, Germany.
- Caliper Mapping & Transportation Software Solutions. (2021). *TransCAD Transportation Planning Software*. Von <https://www.caliper.com/tcovu.htm> abgerufen
- Eissfeldt, N. (2004). *Vehicle-based modelling of traffic . Theory and application to environmental impact modelling*. Dissertation, Universität zu Köln.
- Erdmann, J. (2012). Online-Kalibrierung einer Mikroskopischen Verkehrssimulation. *VIMOS-Kolloquium 2012*. Dresden, Germany.
- Huainan City Planning Bureau; School of Transportation of the Southeast University. (2010). *Urban Comprehensive Transportation Planning of Huainan City (2009 ~2020)*. Huainan, China.