Calibrating Traffic Simulation Models with Limited Field Data: A Case Study on New Jersey Turnpike

M. Bagheri[a*], B. Bartin[b], K. Ozbay[c], P. Jamshidi Ghaleh[d]

[a] PhD Student, Department of Civil Engineering, Ozyegin University, Turkey
mohammad.bagheri@ozu.edu.tr
[b] Associate Professor, Department of Civil Engineering, Ozyegin University, Turkey
bekir.bartin@ozyegin.edu.tr
[c] Professor, Department of Civil and Urban Engineering and Director of the C2SMARTER Cen-ter (Tier 1 UTC funded by USDOT), New York University, New York
kaan.ozbay@nyu.edu
[d] MSc Student, Department of Mechanical Engineering, Ozyegin University, Turkey
parivash.jamshidi@ozu.edu.tr
* Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 29, 2024

Keywords: Calibration and Validation, Microscopic Traffic Simulation, SUMO, Origin-Destination Demand Estimation

1 INTRODUCTION

Traffic simulation models in practice mostly utilize historical data for calibration and validation (C&V). Often, the available real data exhibit limitations, notably in terms of the number of days over which they were collected and their spatial coverage within the network. Ideally, ground truth data from different days should be employed for C&V, but collecting all the required data with the desired spatial and temporal accuracy becomes both costly and time-consuming, and in many cases impossible. Consequently, the only viable recourse, which is often encountered in practice, is to gather existing datasets from multiple sources to be parsimonious with the time and cost associated with data collection (Bartin et al., 2018). Moreover, the developed simulation models are generally used to estimate the impact of future, either non-recurrent or planned, events, such as crashes or work zones. When simulations of planned, future systems are created, validation is not possible since field data cannot be collected. It is only after the anticipated event takes place that one can look back at the model's initial assumptions and evaluate its performance retrospectively. Post-evaluation of traffic simulation models is a neglected yet valuable area of research. It offers modelers a chance to validate their predictions of the future system simulations with actual field data. To that end, the objective of this study is to evaluate the performance of a simulation model in handling disruptions and unexpected events in a traffic network with complete and partial data availability. A case study and a detailed real-world dataset pertaining to New Jersey Turnpike (NJTPK) were used to showcase the importance of ground-truth data in traffic simulation models. The results of this study offer valuable insights into the model's robustness, particularly under nonrecurrent traffic disruptions.

2 METHODOLODY

A six-mile section of the NJTPK from Interchange 12 to 13A was selected as the study network. There are five zones in the network: three interchanges, namely interchange 12, 13 and 13A, and two external zones, southbound and northbound, connecting this six-mile section to the rest of the NJTPK network. Only the northbound travel direction was analyzed, with two available routes between zones, cars-only and cars/trucks/bus lanes. SUMO microscopic traffic simulation software was used to model the study area. The simulation model is shown in Figure 1. Utilizing the NJTPK's electronic toll collection (ETC) database, detailed traffic-related information was extracted, including hourly link volumes, vehicles' departure times, travel times between zones and origin-destination (OD) demand matrices during morning peak hours (7 AM - 9 AM). In addition, this study used the NJTPK traffic management center (TMC)'s incident logs and state police's accident database to select an actual accident scenario.

Four simulation models were developed. The overall structure of these models is presented in Figure 2. **Model I,** a proxy model (akin to a digital twin), was calibrated and validated to depict a typical day (February 5, 2015) based on all available data extracted from the ETC database. In this model, vehicles were released from their respective zones at the exact times as shown in the ETC database by using the API functionalities of SUMO TraCI. **Model II** represents the case where most simulation models are developed in practice, i.e. only limited traffic data are available. Therefore, traffic data from Model I, e.g. hourly traffic volumes and travel times, collected using mainline loop detectors and limited number of probe vehicles, were used as input to Model II for the C&V process. To estimate the OD matrix for Model II in the presence of partially available data, the study adopted a generalized least square formulation (Cascetta and Nguyen, 1988) using the data from detector counts, constrained by link capacities. It should be noted that the C&V of Model I and II, the speed limit of links within the network and the percentage of vehicles on cars-only lanes were used as the two calibration variables.

Model III was devised as a proxy model for a nonrecurrent event, which simulated an accident scenario that took place on February 27, 2013, at 7:10 AM between Interchanges 13 and 13A (Figure 1) for a duration of 30 minutes, based on the TMC incident logs. Model I calibrated network was used but this time vehicles were released from their respective zones at the exact times shown in the ETC database of February 27, 2013. Model III was used to observe if the proxy model could be used to predict the impact of an accident on travel times by comparing the simulated versus the observed travel times from the ETC database for that specific day. Lastly, **Model IV** was devised to assess the efficacy of Model II in capturing the network dynamics during an accident. The same accident scenario was incorporated in Model II, which was calibrated with limited ground-truth data collected on a regular day i.e. February 5, 2015.

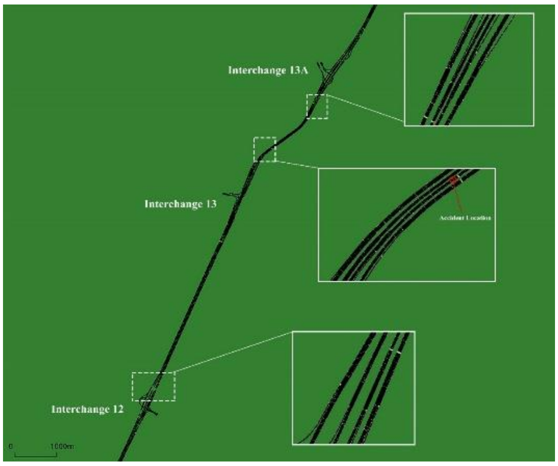

Figure 1 – *The Simulation Network of Study Area*

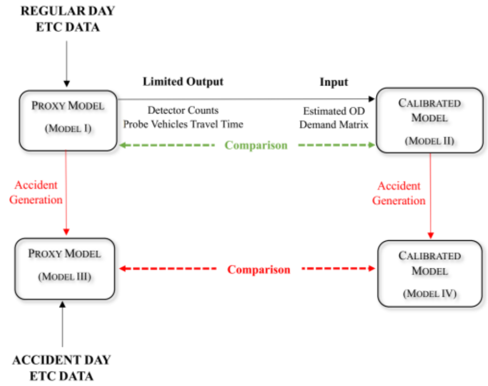

Figure 2 – *The Simulation Analyses Models*

3 RESULTS AND DISCUSSION

Below are some of the key results from the simulation analyses conducted in this study.

Figure 3 compares the travel time histograms for all Models with their respective ground truth values between OD pairs. Model I and Model II are closely aligned with the observed values, while Model III and Model VI are considerably different from their observed values. Additionally, Figure 4 plots individual vehicle travel times against departure time, categorizing vehicles by type and route choice. Each data point in Figure 4 represents the observed travel time, and the line plots represent the simulated average travel time calculated at a 10-minute time interval. The shaded regions around each line indicate one standard deviation range from the average. The results show that Model I and II could successfully reproduce the observed travel times. However, under the accident scenario, there is a considerable difference between the observed and simulated travel times, both in Model III and IV. Figure 4(c) and (d) show that Model III and IV are successfully able to capture the peak induced by the accident. Nevertheless, they fail to capture the higher travel times once the crash was cleared. Further investigation of the NJTPK TMC's incident log reveals that significant delays were recorded two miles north of the study network in the morning period during the same accident day, which propagated down to the study area. This delay that occurred outside the study network clearly had a significant impact on the travel times within the study network, as also evidenced in Figure 4. It should be noted that the fact that Model IV is also able to capture the delays due to the accident despite the limited amount of ground truth data is expected, given that the accident occurred on the mainline. In other words, the estimation of the OD demand matrix was conducted to correspond to the mainline traffic flows, despite a significant disparity between the estimated OD demand matrix and the actual one.

It is also worth mentioning that through the available incident logs and accident database that covered one-and-a-half years, it was cumbersome to find an accident case that was straightforward to simulate its impact on traffic. Because the selected study area is part of the larger NJTPK, any incident that occurred outside impacted the study area. For example, in a case where an accident occurred on cars-only lane, travel times on cars/trucks lanes would exhibit higher travel times, as calculated via ETC data, even though these routes do not directly affect each other. These instances underscore the interconnected nature of traffic networks, where disruptions to traffic flow in one area can have repercussions for others. Investigation of such accident cases is deferred to future work.

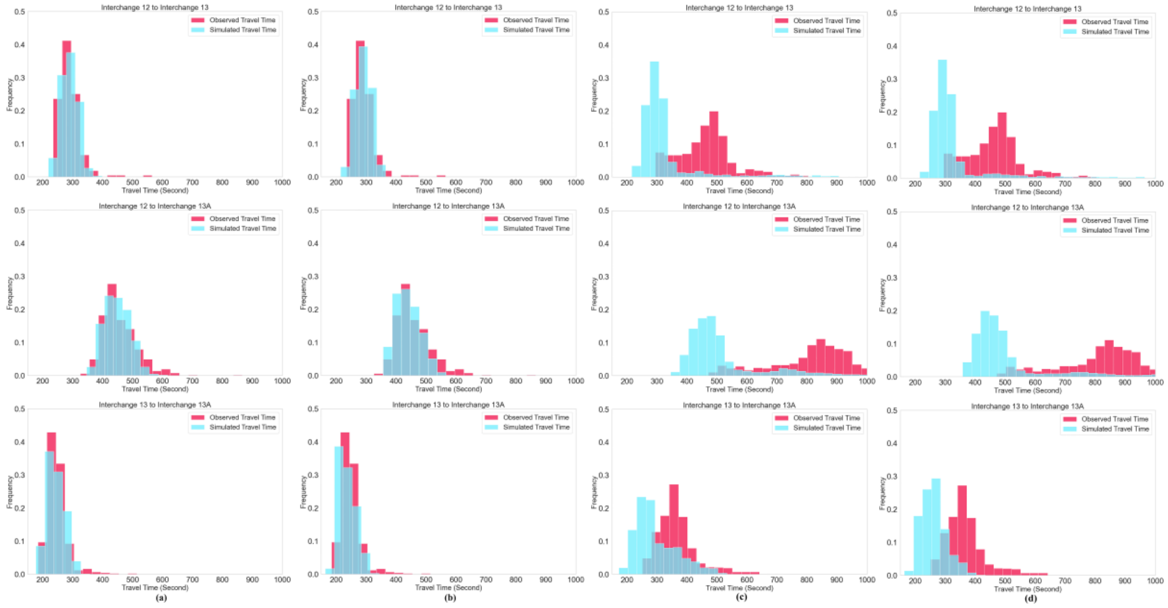

Figure 3 – *The Observed and Simulated Travel Time Histogram for each OD pair: (a) Model I, (b) Model II, (c) Model III, (d) Model VI.*

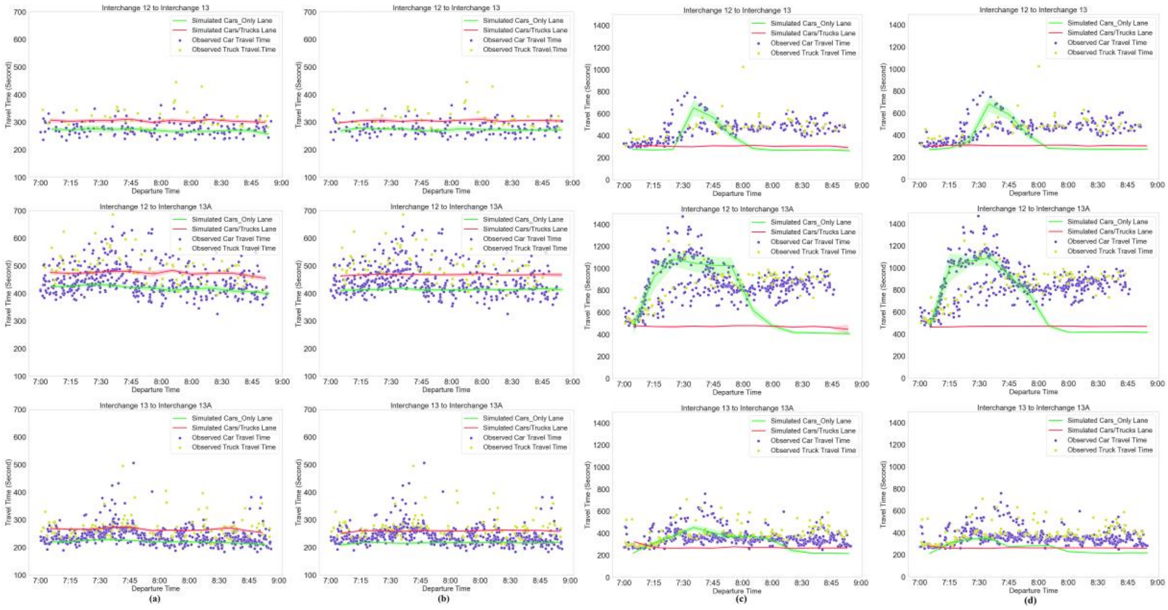

Figure 4 – *The Observed and Simulated Travel Time versus Departure Time: (a) Model I, (b) Model II, (c) Model III, (d) Model IV.*

References

Bartin, B., Ozbay, K., Gao, J. and Kurkcu, A., 2018. Calibration and validation of large-scale traffic simulation networks: a case study. Procedia computer science, 130, pp.844-849.

Cascetta, E. and Nguyen, S., 1988. A unified framework for estimating or updating origin/destination matrices from traffic counts. Transportation Research Part B: Methodological, 22(6), pp.437-455.