Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 130 (2018) 844–849

Procedia Computer Science

www.elsevier.com/locate/procedia

The 7th International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications (ABMTrans 2018)

Calibration and validation of large-scale traffic simulation networks: a case study

Bekir Bartin[a]*, Kaan Ozbay[b], Jingqin Gao[b], Abdullah Kurkcu[b]

a Altinbas University, Civl Engineering Dept., Mahmutbey Dilmenler Cad No:26 Bagcilar, Istanbul, Turkey
b New York University, Tandon School of Engineering, Civil and Urban Engg. Dept. 15 Metrotech Center 6th Floor, Brooklyn, NY, USA

Abstract

The availability, accuracy and relevance of real world input data are essential for developing a reliable traffic simulation model. Large-scale traffic simulation models, in particular, require data from many sources and in great detail. Though it is now possible to obtain detailed field data with the advent of new technologies such as GPS, cellular phones, RFIDs, it is still a challenge to gather all available data, especially traffic flow data, in the required spatial and temporal accuracy. The central theme of this paper is the calibration and validation (C&V) development of a large-scale traffic simulation model using data from multiple sources.

Peer-review under responsibility of the Conference Program Chairs.

Keywords: Large-scale Models; microscopic traffic simulation; calibration and validation

1. Introduction

Microscopic traffic simulation tools aid to a great extent in estimating the impacts of various operational strategies in complex transportation networks with a fair degree of accuracy when simulation models are calibrated and validated correctly. Even with today's enhanced computational power, the calibration and validation (C&V) process of traffic simulation models is a major challenge because of the high level of uncertainty in the modeled systems[1]. The majority of this uncertainty stems from the necessity of large amount input data that are not always available or observable, thus leaving the analysts with a large set of parameters for calibration. What makes the matter more

* Corresponding author. Tel.: +905326808646;
E-mail address: bekir.bartin@altinbas.edu.tr

1877-0509
Peer-review under responsibility of the Conference Program Chairs.
10.1016/j.procs.2018.04.076

complicated is the fact that, more often than not, simulation tools are used to evaluate future design or operational alternatives either at the network level or at a critical location that is likely have a network-wide impact. As the scope of a simulation model grows, 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 in order to be parsimonious with the time and cost associated with data collection.

To that end the main objective of this paper is to present the process of building a large-scale traffic simulation model with the use of multi-source data for its C&V process via a real-world case study.

2. Literature Review

Even though, each available study on the C&V process suggests the use of rigorous methods, there is an apparent gap between the theory and practice of using such methods, especially when the network scope is large. The relevant studies in the literature vary from small scale traffic simulation models[2, 3-8] to medium to large scale models[9-18]. Most relevant studies in the available literature that developed medium to large-scale models generally utilize the traffic flow data as the only output data type for the C&V process. For example Toledo et al. [12] developed the simulation model of a medium size urban-freeway network north of Stockholm central business district. The data used for this study included traffic count data obtained from loop detectors and motorway control system, travel times from probe vehicles and queue lengths obtained both from probe vehicles and aerial photographs. However, the study only used the traffic counts data for the C&V process. Similarly, in Toledo et al. [14] a framework for the calibration of traffic simulation models using aggregate data was presented, in which the study focused on the estimation of an OD matrix using traffic counts observed from sensors. Zhao and Sadek [18] developed a large-scale traffic simulation model of the Buffalo-Niagara metropolitan area, and performed the C&V process using traffic volume data from 162 traffic sensors. Smith et al. [15] developed a large-scale simulation model of a transportation network in Vermont. Similar to the previously mentioned studies, the authors based their C&V process on traffic flow data obtained from nine signalized intersections.

As evidenced from the brief review of the relevant studies, as the network size increases so does the size of output data required for the C&V process. This leads to using only traffic flow as the primary data for the C&V process with minimal number of travel time or queue length data. A similar trend was also highlighted in Brackstone et al [19], where detector counts and travel time were found to be the most frequently used data type in the C&V process.

3. The case study and simulation model development

The following is a succinct description of the real-world case study used in this paper. New Jersey Turnpike Authority (NJTA) has been considering a long-term deck reconstruction work on one of its bridges, located on the Newark Bay Hudson County Extension (NBHCE) of New Jersey Turnpike (NJTPK), a tolled highway. The project corridor is located in a very critical area of the Northern New Jersey highway network and within one-mile of the Holland Tunnel, one of the three available and heavily used connections between New Jersey (NJ) and Manhattan, New York City (NYC). As part of the construction work, a long-term one-lane closure is considered on the eastbound direction of the NBHCE.

Identifying the impacts of a work zone, such as the one in this case study, is often difficult due to the complexity of highway traffic networks. A long-term lane closure can lead to major delays on the specific roadway where the work zone takes place, yet its impact on the rest of the network is often unclear without a comprehensive model. The question is then how to accurately evaluate the impact of a specific lane closure scenario on network traffic. The available methods are the use of macroscopic, mesoscopic or microscopic models. The use of macroscopic or mesoscopic modeling tools, such as using a transportation planning models with static or dynamic traffic assignment and reducing the capacity of links along the project corridor, might lead to mixed and counter-intuitive results. The experience of the authors with macroscopic transportation models has proved that even minor changes in the roadway infrastructure might lead to unexpected network-wide shifts in traffic [20,21]. These shortcomings of mesoscopic or macroscopic modeling tools for analyzing the problem in hand warrant the use of microscopic traffic simulation, where the impact of capacity reduction can be more accurately observed and analyzed.

The simulation model of the study area, shown in Fig. 1, was developed using PARAMICS microscopic traffic simulation software. The model consists of 3,784 links, 2,393 nodes, 133 zones and 106 traffic signals. The model was constructed in stages over the years, designed as a test-bed for various traffic impact analyses for the NJTA, and modified for this specific analysis to include various potential alternative routes. The C&V process included error-checking, demand estimation, capacity calibration, route choice calibration and system performance calibration. These are performed via an iterative process: each time a component calibration is performed, it impacts another component that is already calibrated; therefore, when all the components of the C&V is finished, then the system-wide performance is checked by comparing the selected simulation model outputs with the observed values.

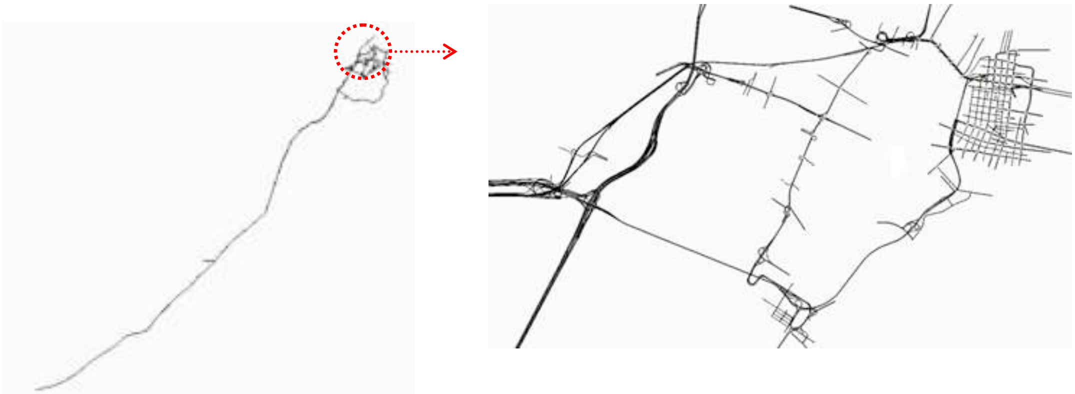

Fig. 1. Simulation Model Developed in PARAMICS

4. System-wide calibration and validation results

This section presents the final system-wide C&V results for the base case, i.e. do-nothing scenario. Note that the traffic simulation model of the study area is quite large. In addition to the amount of data and modeling effort required for the C&V process, it also takes considerable time to complete a single simulation run. For example, it takes between 50 and 60 minutes to finish one simulation run using a PC with a 3.10 GHz processor with 32 GB memory. Variables used for this process are (1) Throughput volume, (2) Queue lengths and (3) Travel times at selected key locations in the study network. Base-case simulation network was run iteratively based on the following formula [22]:

$$n^*(\gamma) = \min\left\{ i : \frac{t_{i-1,1-\alpha/2}\sqrt{S^2(i)/i}}{\left|\bar{X}(n)\right|} \le \gamma' \right\} \qquad (1)$$

Where, $n^*(\gamma)$, minimum number of simulation runs is determined to obtain a relative error γ, by iteratively increasing the number of simulation runs, *i,* by 1 until the proportion of the half-length of the confidence interval to the sample mean $\bar{X}(n)$, i.e. this less than or equal to γ'. Here, γ is the relative error of the sample mean, calculated as $\gamma = |\bar{X}(n) - \mu|/|\mu|$ and $\gamma'$is the adjusted relative error, given as $\gamma' = \gamma/(1-\gamma)$. A relative error $\gamma' = 0.10$ and 95 percent confidence level are used in simulation analyses of this study. In other words, the simulation runs are iteratively increased until condition shown in equation (3) is satisfied for all output variables simultaneously. It should be noted that the Bonferroni correction is not used here since most output variables, such as traffic volume, travel time and queue lengths can be assumed correlated.

4.1. Throughput volume comparison

Table 1 shows a portion of the comparison between the observed versus simulated throughput counts at select locations in the study network. The complete list of select locations was based on their importance and the availability of count data. The results indicate that the validated and calibrated simulation model generates traffic volumes that are in agreement with the observed counts. A GEH value that is less than 5.0 for more than 85% of links is recommended as acceptable by FHWA [22]. It is seen that in 86.2 percent of the cases, the model has GEH values smaller than 5.0. Note that the locations with a GEH value above 5.0 correspond to locations with limited number of volume counts. Therefore, it can be claimed that the simulation model is calibrated and validated for throughput volumes at an acceptable level of statistical accuracy after running the model with multiple random number seeds.

Table 1. C&V Results of Base Case

Throughput Volume - Intersection Level (vph)			
Location	Observed	95% C.I.**	GEH
I-78 Approach at Jersey Ave./12th St.	1,307	[1275,1356]	0.2
RT 139 Approach at Jersey Ave./12th St.	2,794	[2837,2961]	2.0
Erie Street/12th Street Intersection	2,803	[2822,2901]	1.1
Manila Street/12th Street Intersection	2,699	[2630,2725]	0.4
Marin Boulevard/12th Street Intersection	3,195	[3269,3325]	1.8
Holland Tunnel Toll Plaza	2,595	[2705,2756]	2.6
I-78 Grand Street Exit	1,355	[1248,1276]	2.6
I-78 LSC Exit	450	[520,549]	3.8
I-78 Interchange 14 - 14A	3,677	[3660,3726]	0.3
I-78 Interchange 14A – 14B	3,638	[3546,3601]	1.1
I-78 Interchange 14B – 14C	3,489	[3387,3430]	1.4
HT Feed to Route 139 Lower Level	720	[785,820]	3.0
Pulaski Bridge EB	2,700	[2975,3047]	5.8*
GEH < 5.0 86.21 %			

4.2. Travel time comparisons

Nine key routes were selected. Fig. 2 presents the comparison of observed and simulated travel times on the selected paths for the morning rush hours. Travel time comparison results indicate that the simulation model generates travel times that are in acceptable agreement with the observed ones. One of noticeable discrepancies is observed at path 5, between Interchange 14 and NYC via NBHCE, which is approximately 4 minutes less than the observed travel time. The probable reason for this discrepancy is the fact that travel times along this path include the NYC side. In the simulation model, when vehicles traverse the tunnel they reach their destination zone. However, in reality, vehicles are exposed to traffic after they cross the tunnel link. Since the simulation model does not include the street level details of the NYC side, these traffic conditions cannot be observed.

4.3. Queue length comparison

Aerial images were used to determine the queue lengths in the eastbound direction of NBHCE and Route 139, two main routes leading to the Holland Tunnel. As shown in Fig. 3, the observed queue lengths on NBHCE varied between 1.04 and 2.7 miles starting from the Holland Tunnel toll booths. Note that this range is by no means precise because it is calculated based on aerial images that are 15 minutes apart. However, even though there is a time gap between each aerial image they are indicative of how queues form on NBHCE eastbound. Using the base case simulation model, the 95 percentile of maximum queue lengths were calculated for the NBHCE and Route 139 eastbound direction, also shown in Fig. 3. The results indicate that queues on NBHCE eastbound reach the Grand Street off-ramp 60 percent of the time (1.30 miles from Holland Tunnel tolls), and reach Interchange 14C 20 percent of the time (2.07 miles from Holland Tunnel tolls). Considering that observed queues extend between 1.04 and 2.7 miles, it can be claimed that queue lengths on NBHCE eastbound in the base case simulation model are in close agreement with the observed values. As to the queue lengths on Route 139 it was found out that the observed queue length ranges between 0.19 and 1.32 miles with an average of 1.21 miles. The simulation results indicate that queues on Route 139 reach 0.44 miles length 80 percent of the time, and reach 1.32 miles 10 percent of the time. Thus it can be claimed that these simulated queue lengths match the range obtained from the aerial images.

Route	Observed TT Source	Observed TT (minutes)	[95% C.I.]	Base Model TT (minutes)	[95% C.I.]
1	ETC	7.15	[5.43, 8.87]	7.27	[7.2, 7.34]
2	ETC	3.32	[2.83, 3.79]	3.30	[3.24, 3.35]
3	BlueTOAD	16.30	[15.45, 17.14]	22.77	[20.5, 25.04]
4	TRANSCOM	37.05	[36.1, 37.98]	35.00	[33.25, 36.76]
5	TRANSCOM	42.55	[41.22, 43.9]	38.35	[36.08, 40.61]
6	BlueTOAD	12.03	[11.42, 12.64]	14.21	[12.85, 15.58]
7	BlueTOAD	3.85	N/A	4.45	[3.64, 5.26]
8	BlueTOAD	5.07	N/A	5.01	[4.97, 5.05]
9	BlueTOAD	3.18	N/A	3.02	[2.86, 3.17]

Fig. 2. Comparison of Observed and Simulated Travel Times on Selected Paths

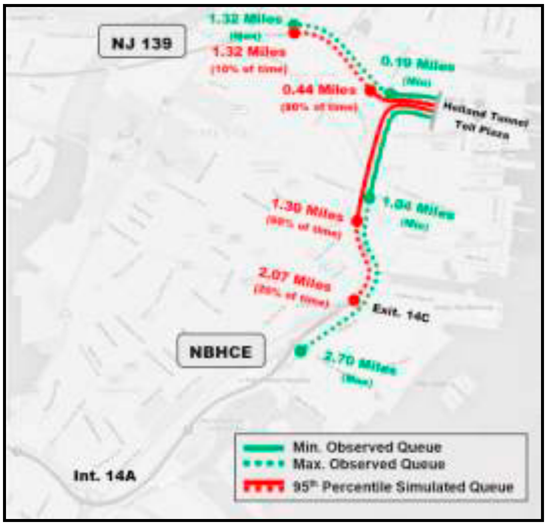

Fig. 3. 95 Percentile Maximum Simulated Queue Lengths in Base Case

5. Summary

There are inherent difficulties in building and validating and calibrating large-scale microscopic traffic simulation models, As described throughout this paper and also deliberated in the literature[13,15] in detail, these challenges include constructing the network in the correct scale using aerial images, inputting the details of link geometry and capacity, adding of various traffic control signs and devices and the details of their turning movement priorities, selecting the number and the location of demand zones and their connections to the traffic network, estimating / converting an O-D demand matrix, acquiring the necessary data for validation / calibration process and the amount of computational time. This paper presents a case study to demonstrate how the data from multiple sources can assist in dealing with the C&V of large-scale traffic simulation models. In specific, it details the modelling effort required to build a large-scale traffic simulation model, including the available data requirements, generating an O-D demand matrix and the results of the C&V process.

Acknowledgement

This work was supported by the New Jersey Turnpike Authority. The authors are indebted to the Authority for their continuing support in research and their generosity in data provision.

References

1. Barcelo, J. "Models, Traffic Models, Simulation, and Traffic Simulation" Chapter 1 in *Fundamentals of Traffic Simulation*. International Series in Operations Research and Management Science Volume 145, 2010.
2. Iyer, S., K. Ozbay, and B. Bartin, "Ex Post Evaluation of Calibrated Simulation Models of Significantly Different Future Systems," *Transportation Research Record: Journal of the Transportation Research Board*, No. 2161, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 49–56.
3. Abdulhai, B., J. B. Sheu, and W. Recker. (1999) "Simulation of ITS on the Irvine FOT Area Using "Paramics 1.5" Scalable Microscopic Traffic Simulator: Phase I: Model Calibration and Validation." California PATH Research Report UCB-ITS-PRR-99-12. University of California, Berkeley.
4. Lee, D. H., X. Yang, and P. Chandrasekar. 2001. "Parameter Calibration for PARAMICS Using Genetic Algorithm." Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C.
5. Park, B., and J. D. Schneeberger. (2003) "Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System." *Transportation Research Record: Journal of the Transportation Research Board, No. 1856,* TRB, National Research Council, Washington, D.C., pp. 185–192.
6. Ma, T., and B. Abdulhai. (2002) "Genetic Algorithm-Based Optimization Approach and Generic Tool for Calibrating Traffic Microscopic Simulation Parameters." *Transportation Research Record: Journal of the Transportation Research Board, No. 1800,* TRB, National Research Council, Washington, D.C., pp. 6–15.
7. Bartin, B., Ozbay. K, Yanmaz-Tuzel, O. and List, G. (2006) "Modeling and Simulation of Unconventional Traffic Circles." *Transportation Research Record: Journal of the Transportation Research Board.* Vol. 1965, pp. 201-209.
8. Lee, J.B. and Ozbay, K. (2009) "New calibration methodology for microscopic traffic simulation using enhanced simultaneous perturbation stochastic approximation approach" *Transportation Research Record: Journal of the Transportation Research Board* 2124, pp. 233-240.
9. Ozbay. K, Mudigonda, S and Bartin, B. "Microscopic Simulation and Calibration of an Integrated Freeway and Toll Plaza Model." Presented at the 85th TRB Annual Conference, 2006, Washington, D.C.
10. Bartin, B., Mudigonda, S. and Ozbay, K. "Impact of Electronic Toll Collection on Air Pollution Levels: Estimation Using Microscopic Simulation Model of Large-Scale Transportation Network" *Transportation Research Record: Journal of Transportation Research Board. No. 2011*. 2008. pp.68-77.
11. Rakha, H., Van Aerde, M., Bloomberg, L., and Huang, X. (1998). "Construction and calibration of a large-scale microsimulation model of the Salt Lake area." *Transportation Research Record: Journal of the Transportation Research Board, No. 1644*, TRB, National Research Council, Washington, D.C., pp. 93–102.
12. Toledo, T., H. N. Koutsopoulos, A. Davol, M. E. Ben-Akiva, W. Burghout, I. Andréasson, T. Johansson, and C. Lundin. (2003). "Calibration and Validation of Microscopic Traffic Simulation Tools: Stockholm Case Study." *Transportation Research Record: Journal of the Transportation Research Board, No. 1831,* TRB, National Research Council, Washington, D.C., pp. 65–75.
13. Jha, M., Gopalan, G., Garns, A., Mahanti, P., Toledo, T., and Ben-Akavi, M. E. (2004). "Development and calibration of a large-scale microscopic traffic simulation model." *Transportation Research Record: Journal of the Transportation Research Board, No:1876*, TRB, National Research Council, Washington, D.C., pp. 121–131.
14. Toledo, T., Ben-Akiva, M., Darda, D., Jha, M., and Koutsopoulos, H. (2005). "Calibration of microscopic traffic simulation models with aggregate data." *Transportation Research Record: Journal of the Transportation Research Board, No. 1886*, TRB, National Research Council, Washington, D.C., pp 76–84.
15. Smith, M., Sadek, A. W., and Huang, S. (2008). "Large-scale microscopic simulation: Toward an increased resolution of transportation models." *Journal of Transportation Engineering*., 10.1061/(ASCE)0733-947X(2008)134:7(273), 273–281.
16. Praveen, E., Sharma, S. and McGhee, C. (2010). "Development of a Large-Scale Traffic Simulation Model for Hurricane Evacuation - Methodology and Lessons Learned." *Natural Hazards Review*, Vol. 11, No.4.
17. Xiong, C., Zhu, Z, He, X, Chen, X, Zhu, S., Mahapatra, S., Chang, G-L and Zhang, L. (2005). "Developing a 24-Hour Large-Scale Microscopic Traffic Simulation Model for the Before-and-After Study of a New Tolled Freeway in the Washington, DC–Baltimore Region." *Journal of Transportation Engineering*. DOI: 10.1061/(ASCE)TE.1943-5436.0000767.
18. Zhao, Y. and Sadek, A. W. (2012). "Large-scale agent-based traffic micro-simulation: Experiences with model refinement, calibration, validation and application." The 1st International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications. *Procedia Computer Science*. Vol 10. pp. 812-820.
19. Brackstone, M, Montanino, M., Daamen, W., Buisson, C. and Punzo, V. (2012). "Use, Calibration and Validation of Traffic Simulation Models in Practice: Results of a Web based Survey" Presented at the 91st Annual Meeting of Transportation Research Board, Washington, D.C.
20. Ozbay, K., Yanmaz-Tuzel, O., Mudigonda, S., Bartin, B. (2008) "Evaluating Highway Capacity Investments using a GIS-based tool - Trip-based Full Marginal Cost Approach." *Transportation Research Record: Journal of the Transportation Research Board*. Vol. 2024, pp. 44-53.
21. Ozbay, K., Bartin, B. Mudigonda, S., Iyer, S. (2013) "ASSIST-ME Post-processing Tool for Transportation Planning Model Output" *Transportation Research Record: Journal of the Transportation Research Board*. Vol. 2399. pp. 63-73.
22. Law, A. M. and Kelton, W. D. Simulation Modeling and Analysis, 2nd Edition, McGraw-Hill, Inc., 1991.
23. Dowling, R., Skabardonis, A. and Alexiadis, V. (2004). Traffic Analysis Toolbox. Volume III: Guidelines for Applying Traffic Microsimulation Software. Publication FHWA-HOP-04-040. FHWA, U.S. Department of Transportation.