World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016

Large-scale, high-fidelity dynamic traffic assignment: framework and real-world case studies

Qi Yang [a], Ramachandran Balakrishna [a,*], Daniel Morgan [a], Howard Slavin [a]

[a]Caliper Corporation, 1172 Beacon Street, Suite 300, Newton, MA 02461, USA

Abstract

We present a highly detailed, microscopic Dynamic Traffic Assignment (DTA) framework with sufficient fidelity to address emergent and future planning and operations applications. Congestion patterns are estimated at the lane level with explicit modeling of complex signal timing algorithms and their impacts on queues and spillbacks. A Geographic Information System (GIS) ensures the most accurate network representation. The flexible representation of travel demand at the resolution of individual vehicles facilitates the capturing of sufficient vehicle and driver classes, vehicle performance and driving behavior distributions, inter-vehicle interactions, and temporally fine trajectories. Model outputs are saved at any desired granularity, and may be used to assess entire distributions of performance metrics to support reliability studies. Applications of the framework include the study of connected vehicles, Intelligent Transportation Systems (ITS), advanced tolling systems, and emissions modeling, and safety analysis. The above features are implemented with unparalleled computational performance so that large-scale networks may be handled without the need for accuracy-running time tradeoffs. We describe four real-world projects that clearly demonstrate the advantages of the microscopic DTA in practice.

1. Introduction

Motivated by the pressing need to address transportation planning and operations problems characterized by fluctuating demand and supply phenomena, agencies are increasingly exploring Dynamic Traffic Assignment (DTA) as a necessary tool in practice. These problems, such as signal plan optimization, dynamic tolling, high-occupancy vehicle (HOV) lanes, high-occupancy toll (HOT) facilities, emissions monitoring and work zone scheduling, require a detailed handling of driver behavior and network performance that is not afforded by traditional static traffic assignment. Accurate rendering of the ground truth requires that individual drivers' heterogeneous behavior responses are captured reliably, and that the network's response to travel demand decisions respects basic constraints such as those imposed by the capacities of each roadway segment. DTA has demonstrated the potential to address these modeling requirements for some time now, with various research efforts beginning to crystallize into practical guidelines (Chiu

* Corresponding author. Tel.: +1-617-527-4700 ; fax: +1-617-527-5113.
E-mail address: rama@caliper.com

et al. (2011)). Various DTA tools such as DynaMIT (Ben-Akiva et al. (2010)), DYNASMART (Jayakrishnan et al. (1994)) and Dynameq (Mahut and Florian (2010)) have been reported in practice. The reader is referred to Sundaram et al. (2011) and Peeta and Ziliaskopoulos (2001) for a general description of DTA.

Traffic microsimulation has the potential to accurately capture real-world demand and supply phenomena that are inherently stochastic and heterogeneous. However theoretical DTA developments and their software implementations have long focused on mesoscopic network loading, primarily motivated by the perceived computational inefficiencies of traffic microsimulation as a network loading method within a DTA. While this argument was perhaps valid during the early phase of DTA research, current hardware and software technologies allow much faster microscopic loadings and hence render a microscopic DTA both feasible and practical. Mesoscopic loading, which attempts to approximate microscopic link performance phenomena through macroscopic relationships like the fundamental diagram (e.g. speed-density functions), represents a trade-off between modeling accuracy and running time that is neither well-studied nor quantified. It could be argued that calibrating a mesoscopic model is no easier than calibrating a microscopic model, especially since the approximations within the mesoscopic model may introduce inconsistencies as well as result in biased model parameter estimates. Moreover, the identification of time-varying origin-destination demand generally dominates the calibration problem, and is a challenge common to both microscopic and mesoscopic loading methods. It should also be noted that even a one-shot microsimulation relies on well-calibrated congested travel time estimates for all network links, which must generally be obtained from a DTA. Consistency necessitates that the calibration of a microsimulation tool be performed with a microscopic DTA. When the microscopic fidelity is thus required for a specific application, such as traffic signal optimization, traffic operations planning or managed lanes performance evaluation, microscopic DTA is the logical answer.

In this paper, we aim to show that microscopic DTA is feasible on very large datasets. We provide evidence from four large-scale networks from various regions of the US, many of which continue to be in use today as actively-maintained traffic operations/management tools. The rest of this paper is organized as follows: Section 2 outlines our study methodology in broad terms. Section 3 describes the four datasets on which the methodology was applied. Section 4 is a synthesis of salient numerical results from the four cases. Section 5 concludes the paper with general observations and potential next steps.

2. Study methodology

This study employs a DTA based on microscopic network loading principles. The DTA framework (Figure 1) solves for network travel times iteratively: drivers make route and mode choice decisions based on perceived congestion patterns. They then make their trips and experience the outcome of their choices. A learning model updates their perceptions from day to day as new routing options are explored and congestion patterns evolve collectively. The iterations continue in search of a stable Wardrop condition in the dynamic context: drivers in a given departure time interval cannot improve their travel times by switching to another route. Convergence is measured through the relative gap statistic computed and tracked for each departure time interval. The simulation-based network loading in each iteration is microscopic, fully capturing the dynamics and fidelity of real-world phenomena on both the demand and supply sides.

The DTA tool chosen for our tests is TransModeler (Caliper (2015)), which provides a platform for dynamic network loadings at the microscopic, mesoscopic and hybrid levels. Hybrid simulation offers the flexibility to model some network links under a mesoscopic regime while others may be microscopic, thus expanding the scope of simulation models to the regional scale. TransModeler features a native Geographic Information System (GIS) that models transportation networks in four dimensions: three spatial and one temporal. TransModeler also explicitly models the most complex traffic control systems and Intelligent Transportation System (ITS) infrastructure, along with the interaction of drivers to the same. TransModeler's detailed microscopic DTA has also been demonstrated on several large networks around the country.

The learning model in TransModeler follows a travel time averaging scheme:

$$X_{i+1} = (1 - \alpha_i)X_i + \alpha_i f(X_i) \tag{1}$$

where X_i represents the travel times input to iteration i; X_{i+1} are the output travel times from iteration i (and hence the input travel times to iteration $i + 1$); $f(\bullet)$ is the simulation model; and α_i is an averaging weight. The choice of
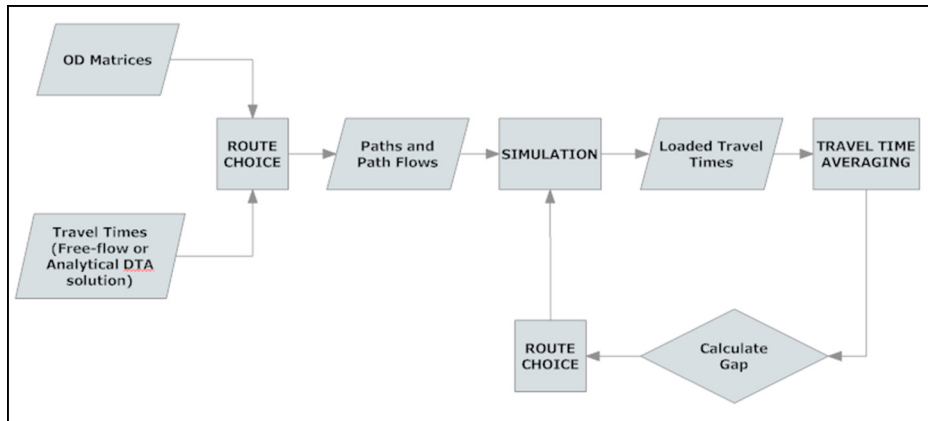
Fig. 1: Simulation-based DTA framework

α_i determines the nature of the learning model. TransModeler provides three options: method of successive averages (MSA), Polyak averaging and fixed-factor averaging.

The general study methodology follows the steps below:

1. Review the available demand data and assess its quality from the DTA perspective.
2. Collect available time-varying traffic data to support model calibration and validation.
3. Assemble a highly detailed network representation including accurate geography at the lane and intersection levels, consistent with the ground truth contained in aerial photographs and other data sources.
4. Prepare a complete input dataset and perform a one-shot traffic simulation to trouble-shoot network issues and assess initial model fit to observed data.
5. Calibrate the model on both the demand and supply sides to best fit the available data. This step may involve origin-destination (OD) trip table estimation on the demand side, and DTA on the supply side.
6. Validate the model with other data (when available) not used for calibration.

DTA model development in practice typically begins with a traditional (i.e. static) travel demand model. A sub-area analysis will then be run to crop out the geography of interest along with a consistent set of trip tables for the sub-area. This set of static trip tables will then be adjusted to change both its scale and structure so as to better match traffic (count) data collected at a much finer temporal resolution (say 15 minutes or hourly). We refer to this step as demand calibration. Each OD adjustment step is followed by a DTA step to update the link travel times to better reflect the new demand pattern. We denote this step as supply calibration.

It should be noted that demand and supply calibration are not independent of each other: one is always performed conditional on the parameters in the other. For instance, OD matrix estimation (ODME) adjusts demand levels and demand patterns to match observed traffic patterns, while fixing the parameters of the car-following,lane-changing and traffic control models that help determine network capacities. Model calibration itself is thus iterative (Figure 2), moving between ODME and DTA until reasonable consistency is achieved. A thorough review of the DTA model calibration problem, its structure, challenges and solution is presented in Balakrishna (2006).

Calibrated models must ideally be validated against data that were not part of the calibration step. The most common approach is to use a hold-out sample so that only a part of the available data are used for calibration. The remaining data are used to test if the calibrated models can predict well, a crucial test to minimize over-fitting. Alternatively, data of one type (say traffic counts) may be used for calibration while data from a different source (such as traffic speeds or travel times) may form the basis for model validation. Model calibration and validation involve numerous technical challenges, some being more familiar than others in practice. Prior work (see for example Antoniou et al. (2011) and Toledo and Koutsopoulos (2004)) throws light on several crucial issues that must be considered and addressed during this stage of model development.
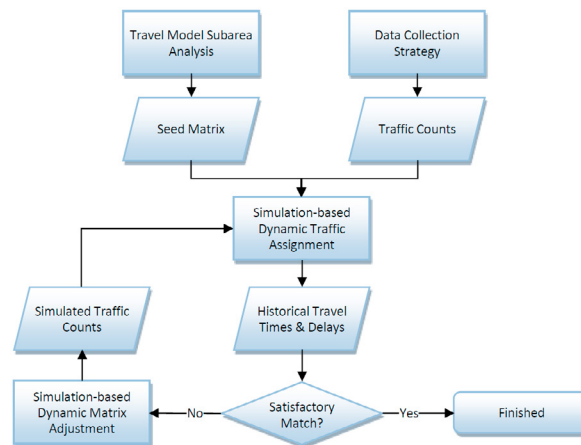
Fig. 2: Iterative demand-supply calibration

3. Description of datasets

We illustrate our approach through real-world projects that are necessarily large-scale and complex with modeling requirements that cannot be met by other approaches such as mesoscopic and macroscopic simulation. We present results from an array of networks covering different roadway structures, spatial demand aggregation levels and project objectives. Datasets are drawn from various parts of the US:

- Phoenix, Arizona
- Jacksonville, Florida
- Whatcom County, Washington
- Lake County, California

The Phoenix, Arizona model was developed for Maricopa Association of Governments, the Metropolitan Planning Organization (MPO) for the Greater Phoenix region. The network (Figure 3) was derived from the regional travel demand model via a sub-area analysis that yielded static OD matrices for an area of about 525 square miles. The network was augmented to include access to all major demand centers through an extensive coverage of driveways and cross-streets, represented by 17,333 nodes and 23,358 links. A total of 2,164 intersections feature traffic control ranging from pre-timed and actuated signals to ramp meters and stop signs. The static OD matrices were used as initial seeds to the ODME problem, which generated dynamic demand tables for DTA analysis.

The Jacksonville, Florida test involved the direct simulation of disaggregate trips output by an Activity-Based Model (ABM). Unlike traditional trip-based models, ABMs predict the movement of individual persons between various daily activities. ABMs can potentially capture the chaining of trips into tours, as well as model interactions among members of a household. Our Jacksonville test used trip outputs obtained exogenously from the DAYSIM ABM, which originate and terminate at one of 492,684 parcels spread across six counties. The geographic extent of the network and a sample of parcel locations is shown in Figure 4. External and truck trips are modeled on a system of 2,578 zones. The network is comprised of 73,260 nodes and 177,735 links. The AM peak period (5:00-9:00 AM) generated roughly 650,000 trips.

The Whatcom County model was developed for the Whatcom Council of Governments in Washington. Its spatial extent covers nearly 800 square miles covered by a mix of freeways and dense arterial streets (Figure 5). The network consists of 2,944 nodes and 3,798 links. 603 of the intersections featured traffic control systems, including numerous roundabouts.

The Lake County, California network (Figure 6) covers about 450 square miles and was represented by 3,300 nodes and 4,200 links. The detailed network covered 720 miles of roadway, of which 120 miles are on state routes. Three-hour AM and PM peaks were calibrated to capture the demand for auto and long-distance truck traffic in the region.
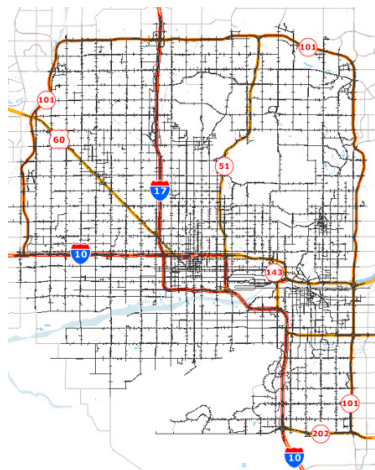
Fig. 3: Phoenix, Arizona network

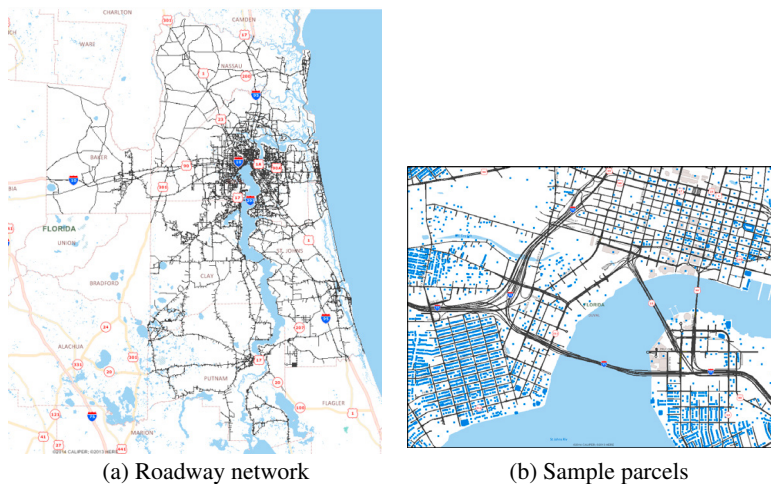(a) Roadway network   (b) Sample parcels

Fig. 4: Jacksonville, Florida

The available data included OD surveys, GPS travel time traces, directional vehicle counts and turning movement counts.

4. Results

All the case studies generally validate the enhanced modeling fidelity and scalability of the microscopic DTA framework. In addition, each example highlights specific modeling capabilities unique to that project. This section provides evidence of the levels of calibration achieved in each case.

The Phoenix model was calibrated against time-varying traffic counts. The calibrated model replicated the hourly link volumes with a maximum error of 15% across both the AM and PM peak periods on the freeways. The model was further validated against Inrix speed measurements. The fit between modeled and Inrix speeds is seen in Figure 7. It should be noted that matching speed or travel time observations is a much harder problem than matching only count data. It is now well-known that multiple OD solutions can result in the same (or similar) fit to traffic counts, while only a subset of these OD solutions may yield a realistic match on congestion patterns. The accurate fit to speed data therefore significantly enhances the quality of the model and the modeler's confidence in its outputs.
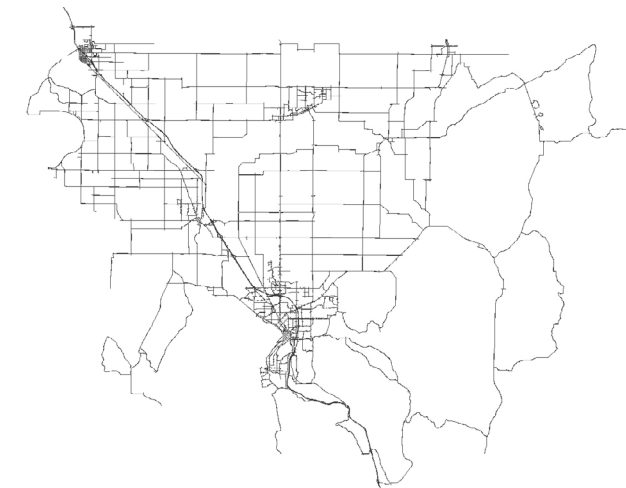
(a) Roadway network
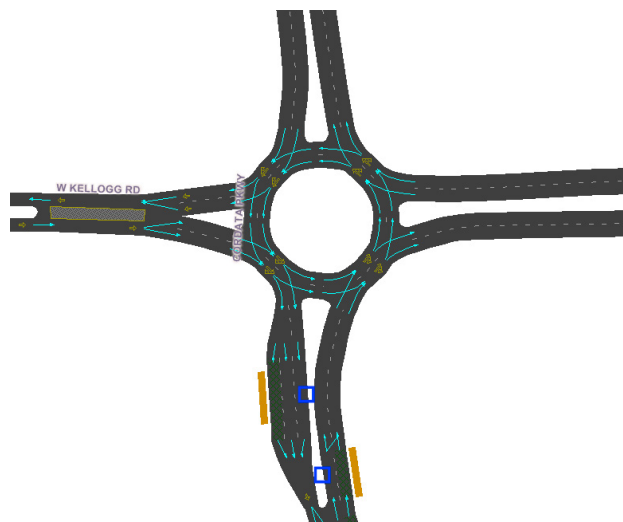
(b) Roundabout

Fig. 5: Whatcom County, Washington

The calibrated Phoenix model is being used in practice in a wide range of projects, and serves as a region-wide inventory of the latest signal timing plans,roadway modifications, and ITS infrastructure.

The Jacksonville model was calibrated against traffic counts provided by time period (AM peak, mid-day and PM peak). Figure 8 shows the fit to the observed counts during the AM and PM peak periods, across all locations for which count data were available. Figure 9 shows the ability to capture the high volumes on the freeways.

The Jacksonville case illustrates the simulation of disaggregate demand generated from an ABM, and represents the state of the art of integrating ABMs with traffic microsimulation (other integration efforts have largely relied on mesoscopic traffic models).

The Whatcom County model (developed for Whatcom Council of Governments) was calibrated to match a range of traffic and turning movement counts. The fit to the available hourly counts is summarized in Figures 10 and 11. A sub-area of the calibrated model was used to analyze several road diet scenarios targeting safety improvements to Alabama Street, a busy urban arterial in the region. These scenarios evaluated various options including the conversion of travel lanes to bicycle lanes, reduction in lane width, and signal timing plan re-configuration.
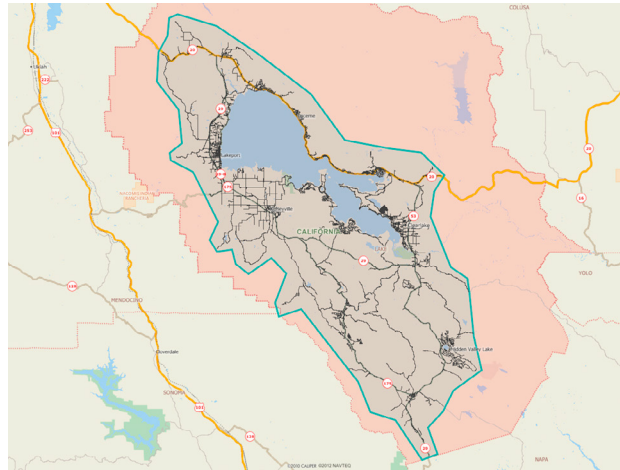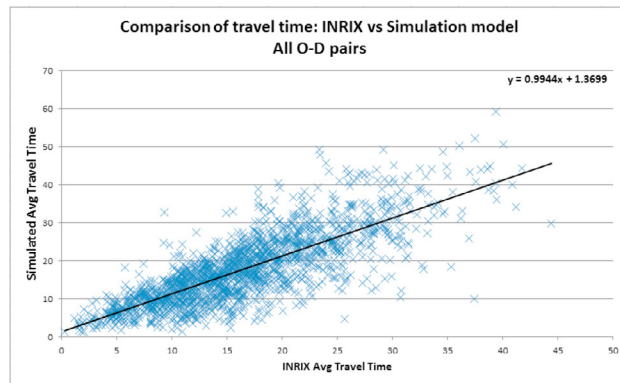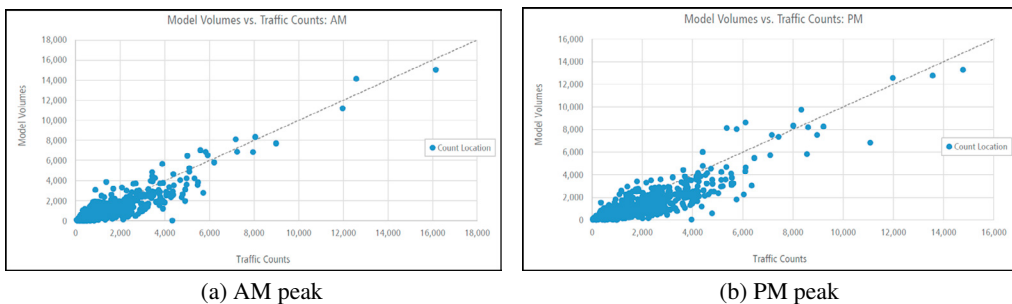
Fig. 6: Lake County network

Fig. 7: Comparison of DTA and Inrix speeds for Phoenix

| (a) AM peak | (b) PM peak |

Fig. 8: Jacksonville: fit to counts (all links)

The Lake County model was built to study the route choices of heavy vehicles and their impacts on the neighborhoods adjacent to the main freeways in the region. GPS travel time data were available to validate the model's predictions. Figure 12 illustrates the calibrated microscopic DTA's ability to replicate point-to-point travel time observations. In addition, a bootstrap procedure was adopted to estimate the variance of the model's travel time estimates and compare the same against the GPS data. The bootstrap is a statistical methodology to approximate a quantity's distribution through a random sampling of the quantity of interest with replacement. The technique allows the mod-
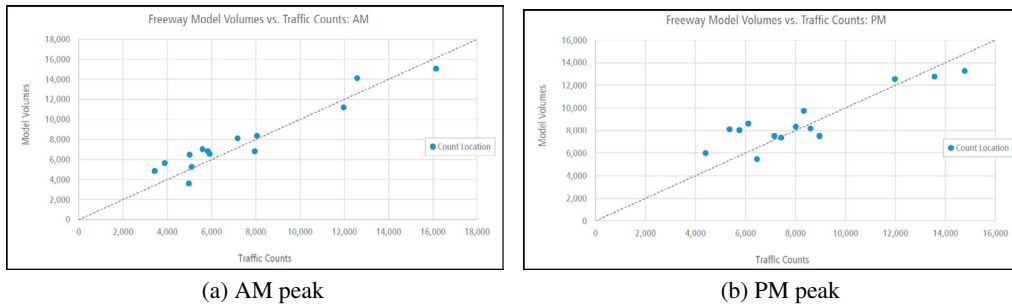
(a) AM peak

(b) PM peak

Fig. 9: Jacksonville: fit to counts (freeways)

(a) 2:00-3:00 PM
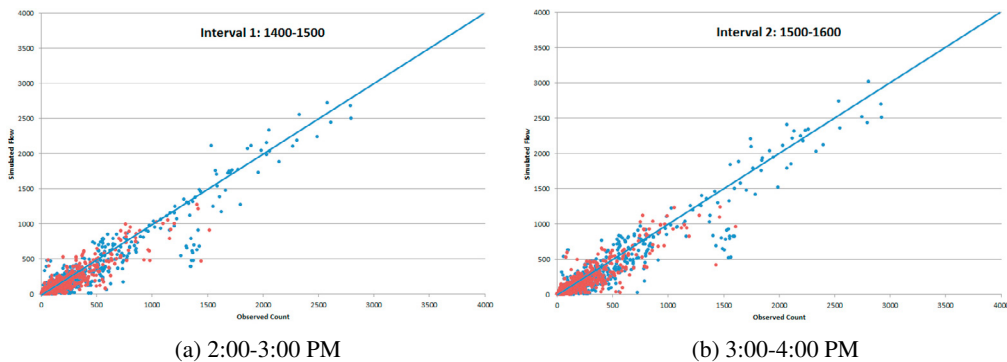
(b) 3:00-4:00 PM

Fig. 10: Whatcom County: fit to counts (2:00-4:00 PM)

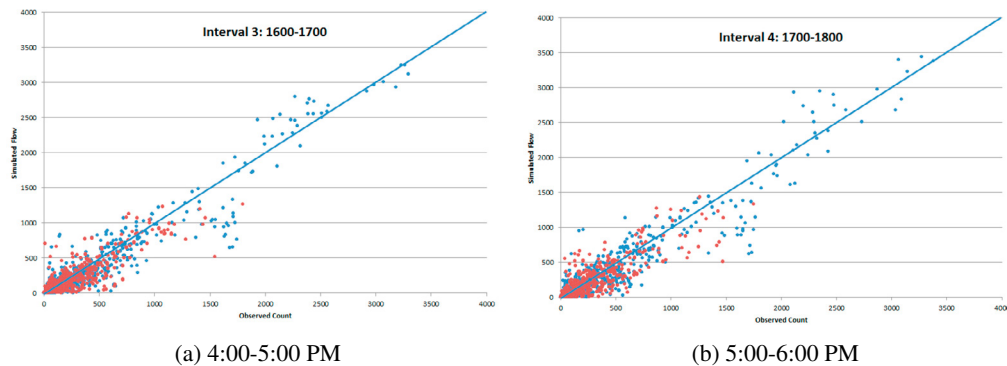(a) 4:00-5:00 PM

(b) 5:00-6:00 PM

Fig. 11: Whatcom County: fit to counts (4:00-6:00 PM)

eler to estimate the mean and variance of the measure of interest from a limited number of model runs. Table 13 summarizes the range of model outputs for each major travel segment, and its corresponding observed value.

5. Conclusion

The objective of this paper is to illustrate the concept of dynamic traffic assignment based on microscopic network loading. This approach marries the high fidelity of traffic microsimulation with the advantages of time-varying network performance estimation available in DTA, and forms the most logical modeling platform for several planning and operations applications such as workzone scheduling, traffic signal optimization, emissions modeling and managed lanes evaluation. We present the broad principles involved in microscopic DTA, and provide strong sup-
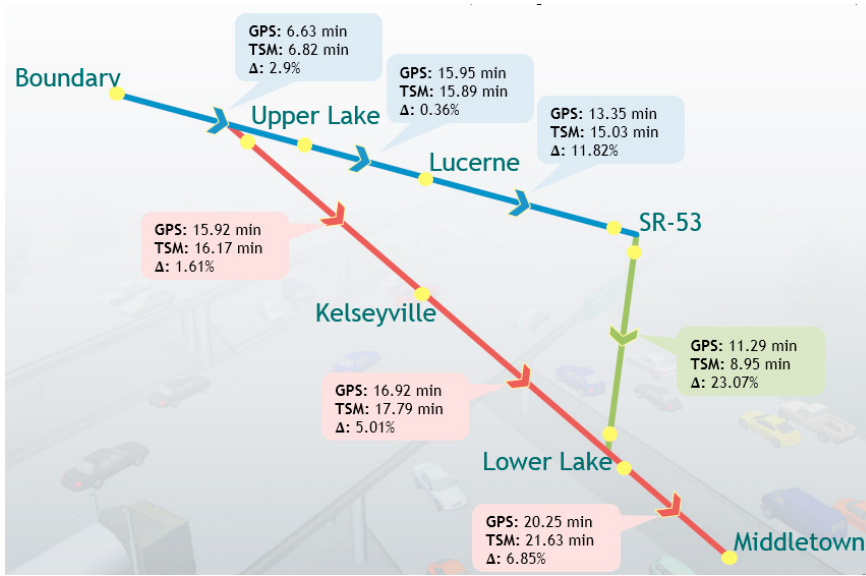
Fig. 12: Lake County travel time comparison

PM							
	From	To	LB TSM TT	Avg. TSM TT	UB TSM TT	Avg. GPS TT	% difference	TSM is...
Southbound	Boundary	Upper Lake	6.90	6.96	7.02	6.24	10.92%	Longer
	Upper Lake	Lucerne	16.15	16.38	16.61	15.58	5.01%	Longer
	Lucerne	SR 53	15.13	15.34	15.56	13.51	12.73%	Longer
	SR 53	Lower Lake	9.19	9.30	9.41	10.66	13.64%	Shorter
	Lower Lake	Middletown	21.42	21.62	21.82	20.52	5.21%	Longer
	Upper Lake	Kelseyville	16.34	16.61	16.89	16.03	3.56%	Longer
	Kelseyville	Lower Lake	17.66	17.95	18.25	16.86	6.26%	Longer
Northbound	Upper Lake	Boundary	6.92	6.99	7.05	6.55	6.45%	Longer
	Lucerne	Upper Lake	16.36	16.58	16.78	15.02	9.88%	Longer
	SR 53	Lucerne	14.84	15.03	15.20	13.12	13.53%	Longer
	Lower Lake	SR 53	9.10	9.21	9.31	10.14	9.60%	Shorter
	Middletown	Lower Lake	22.28	22.42	22.57	19.96	11.60%	Longer
	Kelseyville	Upper Lake	16.28	16.54	16.81	17.31	4.57%	Shorter
	Lower Lake	Kelseyville	18.15	18.55	19.02	16.82	9.81%	Longer

Fig. 13: Comparison of bootstrap estimates against GPS data for Lake County

porting evidence from four very large networks. The numerical results prove that microscopic DTA is both feasible and practical, running in reasonable time on desktop hardware that is affordable and readily available. These results have a strong bearing on the state of the practice, as modelers and planners now have a tool that does not compromise modeling accuracy based on running time constraints.

It is nevertheless important to continue demonstrations of microscopic DTA on even larger networks and in a wider range of applications. In the ABM context, a logical next step is to feed the microscopic DTA travel times back to assess its impacts on tour formation and scheduling. The ABM's sensitivity to dynamic skims (as opposed to static skims) should be of relevance to both theory and practice. A second problem of interest is the study of convergence behavior in such a feedback system.

Acknowledgements

References

Antoniou, C., Balakrishna, R., Koutsopoulos, H.N., Ben-Akiva, M., 2011. Calibration methods for simulation-based dynamic traffic assignment systems. International Journal of Modeling and Simulation 31, 227–233.

Balakrishna, R., 2006. Off-line Calibration of Dynamic Traffic Assignment Systems. Ph.D. thesis. Massachusetts Institute of Technology.

Ben-Akiva, M., Koutsopoulos, H.N., Antoniou, C., Balakrishna, R., 2010. Traffic simulation with dynamit, in: Barcelo, J. (Ed.), Fundamentals of Traffic Simulation. Springer. International Series in Operations Research and Management Science, pp. 363–398.

Caliper, 2015. TransModeler User's Guide.

Chiu, Y.C., Bottom, J., Mahut, M., Paz, A., Balakrishna, R., Waller, T., Hicks, J., 2011. Dynamic Traffic Assignment: A Primer. Transportation Research Circular E-C153, Transportation Research Board.

Jayakrishnan, R., Mahmassani, H., Hu, T.Y., 1994. An evaluation tool for advanced traffic information and management systems in urban networks. Transportation Research C 2, 129–147.

Mahut, M., Florian, M., 2010. Traffic simulation with dynameq, in: Barcelo, J. (Ed.), Fundamentals of Traffic Simulation. Springer. International Series in Operations Research and Management Science, pp. 323–362.

Peeta, S., Ziliaskopoulos, A., 2001. Foundations of dynamic traffic assignment: The past, the present and the future. Networks and Spatial Economics 1, 233–265.

Sundaram, S., Koutsopoulos, H.N., Ben-Akiva, M., Antoniou, C., Balakrishna, R., 2011. Simulation-based dynamic traffic assignment for short-term planning applications. Simulation Modelling Practice and Theory 19, 450–462.

Toledo, T., Koutsopoulos, H.N., 2004. Statistical validation of traffic simulation models. Transportation Research Record , 142–150.