Large-scale dynamic traffic assignment: practical lessons from an application in Sydney, Australia

Melissa Duell (corresponding), Nima Amini, Sai Chand, Hanna Grzybowska, Neeraj Saxena, S. Travis Waller

School of Civil and Environmental Engineering
University of New South Wales, Sydney, Australia

Corresponding author: m.duell@unsw.edu.au

***Abstract*—Traditional static traffic assignment models no longer meet the strategic planning needs of most major metropolitan areas, especially in regard to evaluating major infrastructure projects. One promising possibility is dynamic traffic assignment (DTA), which has been receiving greater attention in the research community for the last ten years. This work describes the ongoing experience of building the first large-scale DTA model in Australia. We divide our experiences into categories regarding data, implementation, and visualization, and we discuss the challenges faced as well as our methods for overcoming those challenges. Finally, we discuss initial model results and the calibration process. In the future, the DTA model described here could aid in evaluating important policy decisions and infrastructural development in the context of the macro/meso-scale network operation. This project serves as a proof of concept for the Australia region and may provide valuable insight to other practitioners interested in emerging areas of transport planning and traffic modeling.**

Keywords*—dynamic traffic assignment; large scale; Australian application;

I. INTRODUCTION

Dynamic traffic assignment (DTA) is one of the next steps in the continuing evolution of practical traffic models. While static transport planning models have a rich history in research and in practice, it is well acknowledged that they do not meet all needs and offer numerous opportunities for improvement. Additionally, traffic micro-simulation, which is another useful prediction tool, is impractical for most large-scale applications due to the intensive data and computational requirements.

DTA is able to capture time-dependent phenomena such as queue spillback, bottlenecks, and temporal congestion. In addition, DTA is an active field of research, which makes it an appealing alternative to traditional static traffic assignment models. However, DTA applications remain relatively scarce in practical settings, possibly due to model complexity and general confusion regarding the practicalities of large-scale implementation and calibration. Thus, this work intends to share insight and offer practical knowledge about building and implementing a large-scale DTA model.

This paper presents the development of the metropolitan area dynamic assignment model (MADAM) for Sydney, Australia. This application consists of a two-hour AM-peak network consisting of 58583 links, 20730 nodes, 2282 zones, 1262930 vehicle demand, 490 signalized intersections, and 1159 bus routes. A number of the project challenges involved processing data, particularly using the available static planning data, such as origin-destination trip matrix, to generate the data necessary for a dynamic model, such as time-dependent vehicle demand. The team reports the presence of a learning curve regarding the handling of relatively large datasets that was ultimately overcome by employing more visualization techniques in GIS software. Model deployment and calibration is currently an ongoing process, but current results are presented here.

First, this work discusses the background of DTA and other deployed DTA models. Next, we summarize a number of aspects regarding model development, and then we present a sample of our results, experiences with model calibration, and the limitations of our modeling approach. This paper concludes with a brief discussion of future model extensions.

II. BACKGROUND

A. Dynamic traffic assignment

The static traffic assignment models that serve as the route choice component of a traditional four step planning model are generally based on the Wardropian principle of user equilibrium (UE) [1]. In UE, users choose the shortest path to minimize their own travel time, which collectively results in a state of network equilibrium. These models have been enduring due to numerous favorable properties, such as uniqueness, stability, and computational tractability. Static models are generally based on link performance functions and output average measure of network performance. They are not well-suited to capturing time-dependent network phenomena such as bottlenecks, reliability, or the "peak" effect of travel demand [3].

However, of course it is well established that time dynamics play a vital role in traffic conditions. This application seeks to capture the impact of different traveler

978-1-4673-6596-3/15 $31.00 © 2015 IEEE
DOI 10.1109/ITSC.2015.282

departure times on network conditions (although it does not account for departure time choice). Fig. 1 illustrates trips over a twenty-four hour period on an average weekday in Sydney as reported in the Household Travel Survey [5]. While the peak variation shown in Fig. 1 is not unique to Sydney, it does make a convincing case as to the importance of accounting for time dynamics in traffic models.

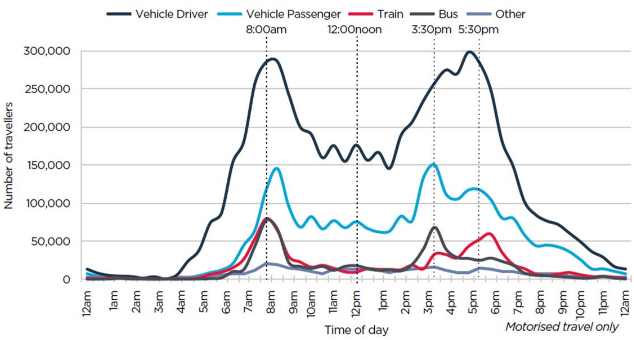

Fig 1. Illustration of time-variance in weekday trips from the New South Wales Household Travel Survey (BTS, 2013)

Dynamic traffic assignment is an established method that is well described by Chiu et al [5]. In this work, we refer to simulation-based DTA that follows a general procedure such as in Fig. 2. Based on an input demand for a chosen time period, DTA seeks the network conditions of *dynamic user equilibrium*, in which the travel time on all paths between an origin-destination at a departure time (ODT) is equal. DTA is generally solved using an iterative procedure in which time-dependent minimum cost paths for each ODT are found, then the updated network conditions are simulated, and the distribution of vehicles on each path between an ODT are adjusted in order to minimize a relative gap.

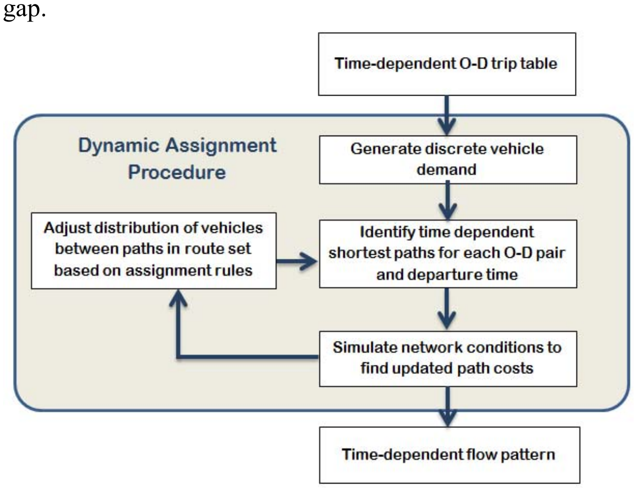

Fig. 2. Overview of the DTA Procedure

The DTA platform utilized in this application is known as VISTA [11] and has also been used in a number of other deployed DTA approaches. A variety of other DTA platforms based on different methodologies are also commercially available.

B. *Previously deployed DTA models*

There are four primary DTA models that served as valuable knowledge bases for the Sydney application presented here:

- Chicago, Illinois, USA [6]
- Dallas/Fort Worth, Texas, USA [7]
- Austin, Texas, USA [8]
- San Francisco, California, USA [9], [10]

A comprehensive review of each of these projects is beyond the scope of this work. However it is important to note that this project builds upon a foundation established by the previously deployed models that experienced many of the same issues we encountered during the MADAM development. Nevertheless, the Sydney application is a large-scale effort and the first in Australia, and of course, the city has its own unique network structure that make this DTA model different from models in other cities. We believe our experiences offer a valuable addition to the existing community.

C. *Study area*

This project recounts the development of a DTA model for the area of Sydney, Australia. Sydney is built primarily on several large bays and has a population of approximately 4.8 million. An average weekday features 16,670,000 trips as of 2013, of which 69% are by vehicle (BTS, 2013). Traffic congestion is considered a significant problem, especially during peak hours. Sydney ranked 21st in the world in the TomTom 2014 Congestion Index, which was a drop from previous years [13]. While traffic microsimulation models have been developed for various projects areas around the region, this is the first attempt at a strategic level DTA model.

III. An Overview of Practical Considerations

This work describes the experiences of building a large-scale DTA model in Australia. Fig. 2 illustrates the four major steps of the workflow in this project, as well as a brief description of the primary concerns for each step. The four steps included acquiring the data, processing the data, implementing the model, and finally, calibrating the model. There is significant overlap between the different tasks, but Fig.3 nevertheless represents a general overview of the steps for model development. An overview of the skills the team needed to work on MADAM include: a thorough understanding of DTA methodology, SQL scripting for large dataset processing, familiarity with Linux and the command line, ArcGIS, and python for scripting with ArcGIS. Of course, proficiency with spreadsheets is useful as well.

The remainder of this section describes the important aspects of “lessons learned” in terms of data, implementation of the model, and the importance of visualization.

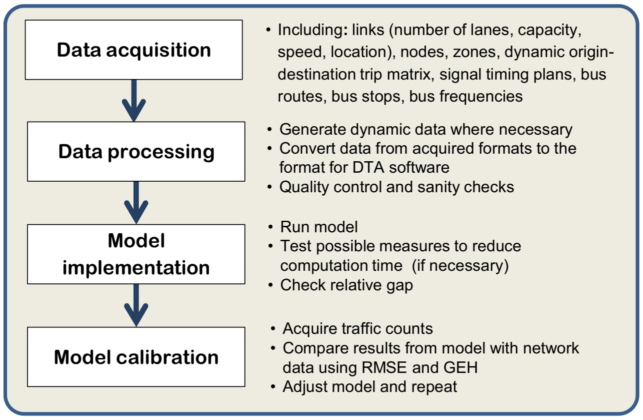

Fig. 3. Summary of project workflow

A. Data

A significant portion of building a DTA model involves developing the most appropriate set of input data possible. This is a particularly challenging task for a large-scale DTA implementation, although there are still fewer requirements than for alternate traffic modeling approaches such as microsimulation. In the context of this project, some datasets were available, while other datasets had to be generated based on static planning data or other readily available data sources.

Knowledge of the area of interest is vital for quality control at this stage. A representative example in our experience was the impact of parking on the number of available lanes on a link, which affects the capacity, which impacts route choice. This was an important consideration for the Sydney region. A number of major routes in the Sydney network are comprised of arterial roads that allow parking on curb-side lanes. However, some of these routes prohibit parking during AM peak hours. After consulting with experienced traffic modelers in Sydney, the team determined that during the AM peak, inbound corridors have 3 lanes available, while the outbound corridors generally have 2 lanes available.

For this application, the primary source of data was the Sydney Strategic Travel Model (STM3) [4] that was developed by the Bureau of Transport Statistics (BTS) and the Roads Network Model that was developed by the Roads and Maritime Services (RMS). Each of these models were in the EMME2 software format and included approximately 80,000 links, covering an area from Newcastle about 160 km north of Sydney down to Illawarra 100 km south of Sydney to the Blue Mountains region 75 km to the west. The main differences between the two models was reflected in their intended purposes. The STM is primarily a travel demand model and included more data regarding transit, disaggregated travel zones and data that impacted mode choice, while the Roads model had slightly more detailed data regarding the road network.

Considering the size of the dataset, comparing the two models proved challenging until the appropriate routines were developed using SQL databases and ArcGIS. The comparison showed that both models were almost identical in terms of the network structure, links, and nodes. For the MADAM project, the network structure including nodes, connectors, links, and the capacities and free flow speed of all links was obtained from the STM3, and later adjusted for calibration based on the Roads model.

The modeling approach used in the DTA simulation platform requires specific input data including: the origin zone, destination zone, vehicle type, and the demand between the zones. This information was extracted from the STM3 for the AM peak. There are 2,722 travel zones representing both origins and destinations of the potential trips, with a total of 7,409,284 OD pairs with nonnegative demand (although many of these had a demand less than one). The total demand of 1,665,440 equivalent passenger car units, represents the AM peak period adjusted to account for cars and trucks.

In order to capture the impact of a time-varying travel demand, the departure time profiles characterizing the vehicle trips during each departure time intervals must be input to the model. Based on a literature review, the team developed a methodology to utilize travel survey data to create a departure time profile specific to statistical sub-divisions (i.e., vehicles from outer regions may depart earlier than vehicles from inner regions). The dataset the team used to develop the departure time profiles for each OD was a result of five waves (08/2009 to 12/2013) of the Household Travel Survey (HTS) performed by BTS. The length of the selected time intervals 15 minutes, which was considered to provide a good balance between having sample size per interval and an appropriate length of time during which departure time choices could occur. Due to the limited number of samples some of the OD pairs were aggregated to Statistical Subdivision (SSD). The dynamic demand matrix was obtained by applying the OD profile to the static demand obtained from the STM3 model and is shown in Fig. 4.

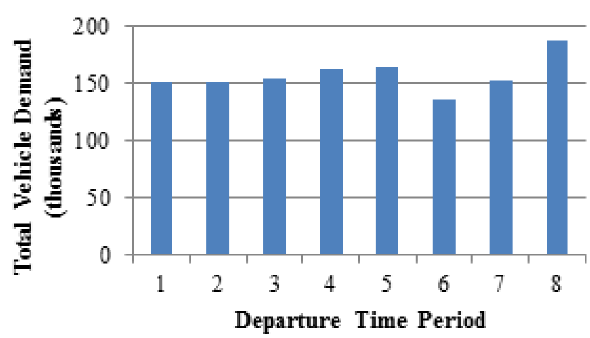

Fig. 4. Vehicle demand per departure time period

To a greater degree than previously deployed DTA models, buses have a significant presence in the Sydney network, especially during AM peak and in the Sydney CBD. While this project does not include transit assignment, it can include the presence of buses in the simulation procedure that determines network conditions. Therefore, the team consulted various data sources and included 1159 bus routes, around 25,000 bus stops, and an assumed dwell time of 25 seconds in the model area

The data required for modeling transit for DTA included: *i*) the links comprising the bus route, *ii*) bus frequencies or bus schedules, *iii*) the location of stops, and *iv*) the dwell time at each stop. Two sources of transit data were explored in the course of the project namely: the General Transit Feed Specification (GTFS) data from New South Wales Transport Data Exchange Program, and the transit data that was included in the STM3 model.

The use of explicit signals data in DTA may facilitate more realistic flow representation. The STM3 model stored green to cycle time ratio and the cycle times for each leg of a signalized intersection, based on data from aggregated Sydney Coordinated Adaptive Traffic System (SCATS) reports for each peak period. However, the signals data was intended to supplement link performance functions and was incomplete for the purposes of DTA. Therefore, the team developed a methodology to deduce dynamic signals movements based on available data.

B. *Implementation*

After all the data is acquired, processed, and translated into the DTA format, the model was ready to be implemented. Model implementation presented another set of challenges. One of the biggest questions for implementation is the computation time for very large scale models such as Sydney. Ultimately, the team decided to explore possibilities to decrease the size of the model and the corresponding run time.

Firstly, the team decided to focus on the Sydney city area, as shown in Fig. 5. This involved eliminating the model areas in Newcastle, Wollongong, and the Blue Mountains. The demand for these areas was aggregated and three new travel zones were created. Ultimately, this decreased the vehicle demand from 1.8M to 1.2M. However, it should not impact the route choice in the city model because the appropriate demand was still included and there is essentially only a single route to access these locations. Fig. 5 shows the model (not including centroid connectors) that was ultimately selected to represent the Sydney area.

Another significant source of computation time is the number of origin-destination pairs. The bottleneck in the DTA software is the time-dependent all-to-one shortest path algorithm for each destination. Therefore, a decrease in the number of destinations may have a significant impact of computational performance. However, too much aggregation will result in a loss of traffic interaction and a model that is too difficult to calibrate.

The demand data used for MADAM was acquired from the STM3, which as a travel demand model, required disaggregate zones in order to capture realistic walking distances to bus stops for a mode split, particularly in the CBD area. Additionally, the demand was based on a destination choice that doesn't include parking, which may not be realistic in a downtown area where parking is not readily available at a disaggregate destination.

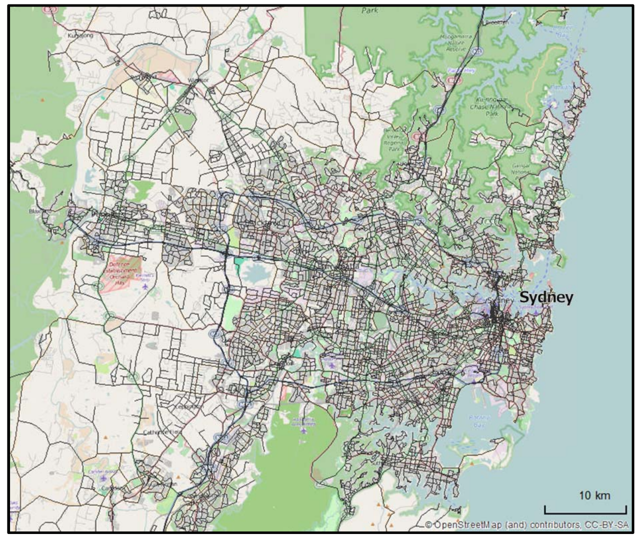

Fig. 5. The Sydney network

Therefore, the team decided to aggregate travel zones to improve computational performance. While an automated aggregation process was devised, ultimately it needed to be adjusted manually through visual inspection to ensure that the aggregation was appropriate. With the reduced number of zones, the run time to evaluate a change to the model varied but was generally in the neighborhood of forty-eight to seventy-two hours.

C. *Visualization*

Ultimately, visualization was one of the most challenging but also valuable aspects of the MADAM application. Visualization allowed the team to identify discrepancies in the datasets. The visualization aspect of this project was achieved using ArcGIS.

Of the many lessons learned during the MADAM project, the value of visualization was one of the most important. Visual inspection was important for comparing the sets of network data, processing the signals and the transit data. It was helpful for ensuring that zones were aggregated appropriately. It was essential for evaluating MADAM results and for the calibration process, where it allowed us to identify which routes may be overloaded with demand and to find nearby corridors that may be underloaded. Additionally, spatial analysis tools in ArcGIS allowed us to make connections between unrelated datasets and significantly reduced the manual workload.

IV. Results and Calibration

This section describes the initial results and calibration process from the MADAM project.

A. Results

Ultimately, the DTA model included 58419 links, 19967 nodes, 1131 travel zones, 1262930 vehicle demand, and 490 signals. The team first ran MADAM using a uniform departure profile, then added the dynamic departure time profile, signals, and transit data, in order to evaluate the changes of including different aspects of the model data. In terms of the results, the team examined measures such as the relative gap, the total system travel time, the cost gap per vehicle, the average travel time on links for specified periods, and the volume of links for specified time periods.

Fig. 6 illustrates the MADAM output from a model version that is midway into the calibration process, meaning that various adjustments to the original link data have been made. Fig. 6 shows a two hour estimated traffic volume, where the darker, thicker lines indicate a higher volume. The model displays a significant amount of congestion, and almost five hours of simulation time after the network loading period ends is necessary to exit all vehicles from the network.

Ultimately, examining the results from the MADAM application is an ongoing process that is closely tied to calibration, which we discuss briefly in the next section.

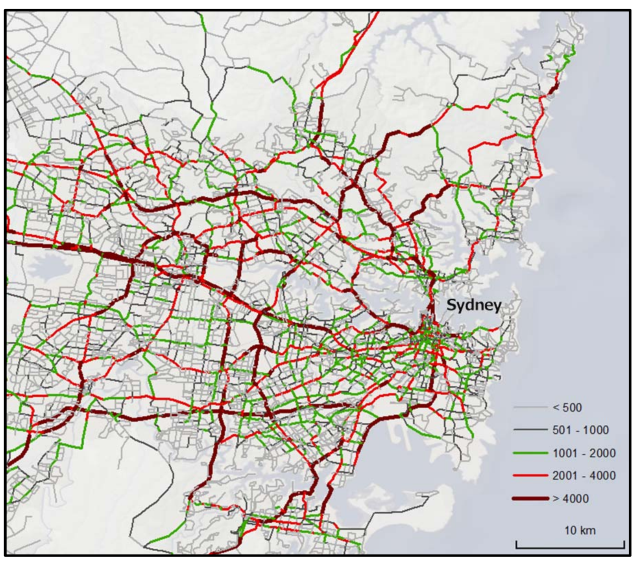

Fig. 6. Demonstration of DTA model output

B. Calibration

A primary purpose of a DTA model is to evaluate potential infrastructure changes in the network relative to a base case scenario. Thus, to ensure the model's predictive capabilities are as accurate as possible, the model output needs to be compared with real data such as link traffic counts or route travel time estimations in order to establish the base case model. When discrepancies are identified, the model needs to be adjusted so that it represents the reality as accurately as possible. This process is known as calibration. It is iterative and lasts until the similarity between the model and the reality is considered suitable, while preserving the favorable model attributes, which in this case is the route choice mechanism.

Thus, the calibration process calls for a variety of real-life data measures to compare with model output. The best calibration data is traffic counts, which can be more or less available depending on the location of interest. However, the MADAM application was able to obtain data that provided a thorough coverage of the entire Sydney network. In the final stage of data acquired, RMS provided 327 average traffic counts for two one-hour-long time intervals of morning peak (7am-8am, 8am-9am) collected at 165 locations as shown in Fig. 7. Fig. 7 also displays the relative magnitude of the two hour count, with the blue bar representing the inbound station and the orange bar representing the outbound station. The team compared these counts with the output from the MADAM.

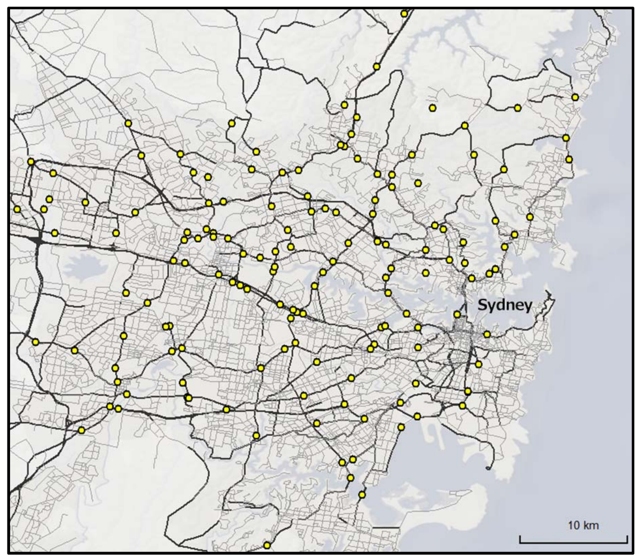

Fig 7. Locations of calibration data counts

The calibration process is arguably one of the most important parts of modeling and it requires skill and knowledge to execute. Ideally a model requires little calibration or a methodological calibration, keeping in mind the intended purpose of the model. However, as practitioners familiar with traffic models are probably aware, the calibration for a traffic assignment model is not straightforward, even in the static case. The process primarily involves making adjustments in the network to link capacity and link speed in order to impact the route choice mechanism and reduce or increase vehicle flows in the appropriate location. Obviously, determining the appropriate locations is the challenge. In some cases the travel demand matrix can also be adjusted, although this project focused on link and link corridor characteristics.

For the MADAM application, the calibration metrics included visual inspection, the absolute and relative difference between the count and model output, root mean square error (RMSE), and the GEH statistic. RMSE measures the differences between the values observed in the real-world and the values predicted by a model. The GEH statistic is commonly used in traffic modeling and it captures the relative nature of a count, i.e., if a predicted volume is 300 vehicles off from a real count, it matters if that count is 400 or 4000.

Fig. 8 presents a summary of the current calibration results. Based on the team's judgement, a two-hour AM peak GEH less than 20 is considered acceptable (although a GEH less than 10 is preferred). This allows for a sufficient margin of error which may stem from various discrepancies between the count data, how the counts were collected, and the model data. Of the 322 calibration points dispersed around the network, 211 have a GEH less than 20. Sydney is a highly corridor-based network, and many of the calibration adjustments have been changes to the speed and capacity along various corridors of links based on specific areas in the network (e.g., adjusting to get the proper flow dispersed between the five places to connect between North and South Sydney).

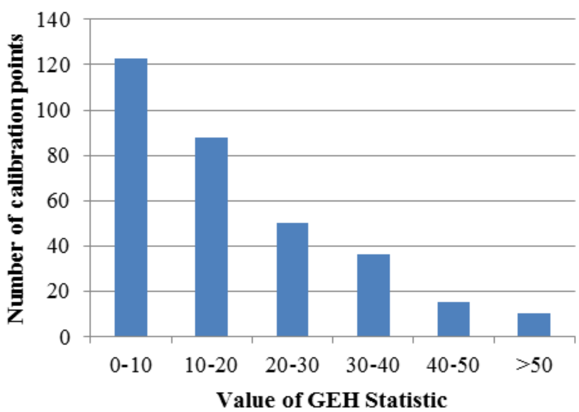

Fig 8. Current calibration results

V. CONCLUSION AND RECOMMENDATIONS

This paper describes the experiences of building the first large scale DTA model in Australia, applied to the Sydney metropolitan area. The project gathered and synthesized numerous data sources including the Sydney Strategic Travel Model (STM3), the Roads Network Model (RMS), the household travel survey, the Sydney GTFS data, Sydney SCATS signals data, and traffic count data from permanent stations acquired from the RMS journey information division. The team implemented the model and devised various techniques to address computation time. Currently, the run time is about 48 hours to evaluate updates in the model.

A calibrated DTA model presents numerous opportunities for future extension, particularly in regard to applications such as environmental impact evaluations. Of course, traffic assignment serves as an important component of a four-step transport planning model, so it would be interesting to incorporate the travel demand aspects and see if predictions change versus the static case. More detailed transit data or even transit assignment could be included. Measures to address the computational challenges will also be necessary. Finally, in order to evaluate the effects of reliability, the team intends to extend the deterministic DTA model to account for volatility in day-to-day traffic flows.

ACKNOWLEDGMENTS

The authors gratefully acknowledge project funding from Transport for New South Wales. In addition, significant assistance and support in the form of time and data came from the Bureau of Transport Statistics (BTS) and the Roads and Maritime Services (RMS). In particular, the authors would like to thank Chris Zito from RMS, Malcolm Bradley and Matthew Jones from BTS, and James Sloan from UNSW.

REFERENCES

[1] Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. In *ICE Proceedings: Engineering Divisions* (Vol. 1, No. 3, pp. 325-362). Thomas Telford.

[2] Chiu, Y.-C., Bottom, J., Mahut, M., Paz, A., Balakrishna, R., Waller, T., and J. Hicks. (2011). "Dynamic Traffic Assignment: A Primer", *Transportation Research Circular E-C153*. Transportation Research Board, Washington, DC.

[3] Peeta S. and A.K. Ziliaskopoulos. (2001). "Foundations of dynamic traffic assignment: The past, the present and the future." *Networks Spatial Econom*, 1:233–265.

[4] Bureau of Transport Statistics, 2011, *Sydney Strategic Travel Model (STM): Modelling Future Travel Patterns,* NSW Government. Available: <http://www.bts.nsw.gov.au/ArticleDocuments/84/TR2011-02-STM-modelling-future-travel-patterns.pdf.aspx> [Accessed 24 March 2015]

[5] Bureau of Transport Statistics, 2013, Household Travel Survey Report: Sydney 2012/13, NSW Government. Available: < http://www.bts.nsw.gov.au/Publications/Reports/Reports/default.aspx> [Accessed 24 March 2015]

[6] Chang, E & Ziliaskopoulos, A, 2003, *Data Challenges in Development of a Regional Assignment: Simulation Model to Evaluate Transit Signal Priority in Chicago,* Transportation Research Record, Vol. 1841, pp. 12-22

[7] Boyles, S, Ukkusuri, S, Waller, T, Kockelman, K, 2006, *A Comparison of Static and Dynamic Traffic Assignment Under Tolls: A Study of the Dallas-Fort Worth Network,* paper presented at the 85th annual meeting of the Transportation Research Board, Washington DC

[8] Duthie, J, Nezamuddin, N, Ruiz Juri, N, Rambha, T, Melson, C, Pool, M, Boyles, S, Waller, T, Kumar R, 2013, *Investigating Regional Dynamic Traffic Assignment Modeling for Improved Bottleneck Analysis: Final Report*, Center for Transportation Research at The University of Texas at Austin

[9] Parsons Brinckerhoff (PB), 2012, *San Francisco Dynamic Traffic Assignment Project "DTA Anyway": Final Methodology Report*, San Francisco County Transportation Authority

[10] Xyntarakis, M, Sall, E, Hicks, J, Charlton, B, 2009, Regional Dynamic Traffic Assignment for Real World Travel Demand Models – Trying it out in San Francisco, San Francisco County Transportation Authority

[11] Ziliaskopoulos, A. K., and S.T. Waller. (2000). An Internet-based geographic information system that integrates data, models and users for transportation applications. *Transportation Research Part C: Emerging Technologies*, 8(1), 427-444.

[12] Ziliaskopoulos, A, Waller, T, Li, Y, Byram, M 2004, *Large-Scale Dynamic Traffic Assignment: Implementation Issues and Computational Analysis,* Journal of Transportation Engineering, Vol. 130, No. 5, pp. 585-593

[13] TomTom, 2014. TomTom Traffic Index, Measuring Congestion Worldwide. Available: http://www.tomtom.com/en_au/trafficindex/#/list