Effective Coupling of Signal Timing Estimation Model and Dynamic Traffic Assignment in Feedback Loops: System Design and Case study

Milan Zlatkovic
Research Assistant Professor
Department of Civil and Environmental Engineering
University of Utah
110 Central Campus Dr., Rm. 2000B
Salt Lake City, Utah 84112
Phone: (801) 819-5925
Fax: (801) 585-5860
E-mail: milan@trafficlab.utah.edu

Xuesong Zhou*
Associate Professor
School of Sustainable Engineering and the Built Environment
Arizona State University
Tempe, AZ 85287, USA
Email: xzhou74@asu.edu

*Corresponding Author

Word Count: 5739 + 1750 (5 Figures + 2 Tables) = 7489

Submitted for publication and presentation considerations at the 93rd meeting of the Transportation Research Board 2014
ABSTRACT

This paper presents an integrated framework for effective coupling of a signal timing estimation model and Dynamic Traffic Assignment (DTA) in feedback loops. There are many challenges in effectively integrating signal timing tools with DTA software systems, such as data availability, exchange format, and system coupling. In this research, a tight coupling between a DTA model with various queue-based simulation models, and a Quick Estimation Method (QEM) Excel-based signal control tool is achieved and tested. The presented framework design offers an automated solution for providing realistic signal timing parameters and intersection movement capacity allocation, especially for future year scenarios. The framework was used to design an open-source data-hub for multi-resolution modeling in Analysis, Modeling and Simulation (AMS) applications, in which a typical regional planning model can be quickly converted to microscopic traffic simulation and signal optimization models. The coupling design and feedback loops are demonstrated on a simple network, and we examine the theoretically important questions on the number of iterations required for reaching stable solutions in feedback loops. As shown in our experiment, the current coupled application becomes stable after about 30 iterations, when the capacity and signal timing parameters can quickly converge, while DTA’s route switching model predominately determines and typically requires more iterations to reach a stable condition. A real-world work zone case study illustrates how this application can successfully be used to assess impacts of road construction or traffic incident events that disrupt normal traffic operations and cause route switching on multiple analysis levels.
1. INTRODUCTION

Dynamic traffic assignment (DTA) models have been increasingly recognized as an important tool for assessing operational performances at different spatial, and across various analysis temporal regimes. The advances of DTA in this aspect are built upon the capabilities of DTA models in describing the formation, propagation and dissipation of traffic congestion in a transportation network. Planning practitioners have recognized the full potential of DTA modeling methodologies that describe within-day and day-to-day network-level and corridor-level congestion in a transportation network with Advanced Traffic Management and Advanced Traveler Information provision strategies (1).

The coupling of traffic assignment and signal control has long been a focus of numerous research efforts. However, due to the complexity of the problem, usually one of two approaches is used: (a) assuming fixed signal control (capacities) and using DTA to perform network loading; (b) assuming fixed traffic assignment and performing signal control evaluation and optimization (2, 3). Model coupling can be achieved on multiple levels, such as full integration, close coupling, loose coupling and isolated approach (4, 5).

While leading-edge research is underway that will significantly advance transportation modeling practices, current signal timing estimation tools (e.g. Synchro, TRANSYT-7F, PASSER) have not been designed to facilitate integrated modeling and practices in combining signal timing and DTA, so these processes are largely ad-hoc and relatively inefficient. Many planners have found numerous challenges in effectively integrating signal timing tools with DTA software systems, such as data availability, exchange format, and system coupling. Determining accurate signal timing data requires precise inputs on intersection geometry and turning volumes. The types of data required for inputs into models are not available, or are difficult to obtain from existing regional planning data, which typically only have the number of through lanes as link level attributes. External data sources, particularly movement-specific attributes, can be coded in signal timing package’s format, such as Synchro’s Uniform Traffic Data Format (UTDF).

Although these data are becoming increasingly available for use in transportation planning applications, the data sources are not yet easily integrated into standard regional planning practices and still require significant labor resources to map signal locations, and convert the data to a useful format for existing DTA packages.

In the current planning practice, the lack of standardized signal data formats calls for manual manipulation or customization of utility tools to interface between planning models and signal timing estimation models. One solution to this challenge is to develop an open data format that allows software vendors and planners to implement data conversion utilities from their own proprietary format. For maximum effectiveness, this research develops open data format through a publicly visible, vendor-independent process to ensure interoperability among diverse traffic modeling and simulation tools.

Although a full integration approach that builds both signal timing estimation and DTA on same design is desirable for multi-resolution modeling, this approach makes the extensions of individual models very difficult. Therefore, a tight coupling approach that can use well defined data format to transfer data files between different resolutions of common models is more desirable in order to achieve scalability, modularity, interoperability, and extendibility. Along this line, the individual models can be developed on the modular basis, where each module can perform its function and can be easily connected and extended to meet future modeling needs.

One of the most significant gaps in the functionality of current multi-resolution modeling systems is the lack of behavioral response associated with (1) capacity allocation at signalized
intersections, (2) traffic congestion due to route diversion and (3) other demand management strategies such as road pricing. Transportation models in many cases require significant manipulation and calibration to produce reasonable results for networks that are over saturated and/or include demand management strategies that affect travelers’ choice of mode, departure time, and route. However, by only focusing on the “demand” side of congestion modeling, many transportation analyses do not reflect the true effects of operating conditions, particularly when it comes to examining signalized intersection performance over multiple days for the purposes of estimating reliability. Therefore many practitioners have observed a critical need for a seamless feedback loop between DTA and signalized intersections capacity allocation.

The demonstration prototype of this study aims to integrate an open-source DTA system, called DTALite, with an Excel-based signal timing estimation tool, called the Quick Estimation Method (QEM). Specifically, a tight coupling between the two models has been developed, enabling a simultaneous assessment of traffic assignment and traffic signals, through an open-source Analysis, Modeling and Simulation (AMS) data-hub, namely Network Explorer for Traffic Analysis (NEXTA) GUI. The integration of the two models and the feedback loop procedure are described on a simple two-route network, and their full capabilities are systematically tested on a real-world network with a work zone application example. The developed QEM tool can be also easily coupled with other DTA packages, such as DYNASMART and TRANSIMS, as AMS data hub can serve as a data exchange utility for a number of simulation packages.

Majority of research in the combined traffic assignment-traffic control environment has been performed for either DTA and fixed signal control, or static assignment and signal control optimization. One mathematical formulation that uses incremental assignment logic is presented in (6) and (7). This is basically an extension of signal optimization under the static assignment, where the signals were repeatedly adjusted for small added portions of the demand until the total demand was loaded. Another conceptual solution framework consisting of two interactive loops for the combined optimization problem is described in (8). However, very limited real-world implementation or numerical application with widely used signal timing standard (e.g. HCM methodology) was performed within the proposed framework. A mathematical formulation and heuristic simulation-based procedure to obtain time-varying assignment of OD vehicles to paths jointly with time-varying signal settings that seek to optimize the total travel time in the network is given in (2). This framework was tested on the Dallas-Fort Worth network with three congestion levels. The implementation of the designed algorithm decreased average travel times in the network, and showed promising results related to the convergence. Another integrated framework that uses the game theory to model the combined DTA-traffic signal problem is explored in (9). In their research, the dynamic traffic control problem and DTA are integrated as a non-cooperative game between a traffic authority and highway users.

Bi-level programming and optimization has been widely used to examine the combined DTA-traffic control problem. Sun et al. (3) developed a bi-level programming formulation and heuristic solution approach for dynamic traffic signal optimization in networks with time-dependent demand and stochastic route choice. In their solution approach, Genetic Algorithm is used to find the upper level signal control variables, Incremental Logit Assignment is developed to find the user optimal flow pattern in the lower level problem, and Cell Transmission Simulation is implemented to propagate traffic and collect real-time traffic information. The study showed that the heuristic solution approach can provide a realistic representation of traffic operation in both network and link level operational performance. Another bi-level formulation for the combined dynamic equilibrium and signal control, presented in (10), defines a signal
control operator in the upper level that optimizes the signal setting to minimize the system travel
time, while the road users in the lower level minimize their own costs, leading to dynamic user
equilibrium. The projection algorithm was used to solve the lower level, and the mixed integer
programming solver to solve the upper level. The algorithm was successfully tested on
hypothetical networks with one and multiple signalized intersections.

Focusing on effective software system integration for transportation planning
applications, our paper presents a data exchange and software coupling framework in which two
models, DTALite (as a representative mesoscopic DTA package) and QEM signal timing
estimation, are successfully linked to address practical needs for providing realistic signal timing
parameters and intersection movement capacity allocation, especially for future year scenarios.
The next section provides descriptions of the general DTA assignment-simulation model and
computational engine. The QEM signal timing methodology and application are described in
section 3. Section 4 describes the coupling of the two models and demonstrates the feedback
loop and their convergence on a simple two-route network. The framework is tested on a real-
world network with a work zone implementation in section 5.

2. MODELING COMPONENTS OF MESOCOPIC DTA PACKAGES

As a newly developed open source DTA package, DTALite uses a number of computationally
simple but theoretically rigorous traffic queuing models in its lightweight mesoscopic simulation
engine. Similar to many existing mesoscopic DTA packages, the four major modeling
components in the DTALite system architecture are summarized as follows: (i) Time-dependent
shortest path finding based on a node-link network structure from regional planning models; (ii)
vehicle/agent attribute generation that converts OD demand matrices in conjunction with
additional time-of-day departure time profile, as well as possible variable value of time
distribution. (iii) DTA method that considers major factors affecting agents’ route choice, such
as different types of traveler information and road pricing strategies, and (iv) a wide range of
traffic flow models, namely point queue, spatial queue with jam density constraint and simplified
kinematic wave models, are available to model essential road capacity reduction or enhancement
measures.

The traffic assignment and simulation modules are integrated and iterated to either
capture day-to-day user response, or find steady-state equilibrium conditions. Within this
simulation-assignment framework, the rich set of output data include traffic MOEs at different
spatial and temporal scales, ranging from network, corridor-level, links and movements. Typical
speed, volume and density measures and agent-based trajectories can be visualized and
processed through the NeXTA user interface. Overall, DTALite uses a link-based and
movement-based simulation architecture with capacity constraints at both link and movement
levels. The capacity constraints make calibrating network capacities a very important step in the
process of building our models. On the other hand, the open-source code base, accessible at
https://code.google.com/p/nexta/, can be flexibly used to enhance to meet specific modeling
needs.

3. QEM METHODOLOGY AND EXCEL-BASED APPLICATION

By adapting the methodology from Highway Capacity Manual (HCM) 2010 (11), QEM is a
simplified method for determining the signal timing parameters, critical intersection volume-to-
capacity (V/C) ratio, and delay for a signalized intersection. To reduce the data input efforts for
DTA and other planning applications, this methodology requires only minimal intersection
group and traffic data, such as the approach lane configuration and turning volumes. It
assumes certain default values for some of the variables which are integrated into the calculation
methodology. QEM provides approximate intersection performance results for the known signal
timing parameters, or calculates the signal timing using a simplified calculation procedure based
on the critical movement methodology.

The QEM Excel-based calculation engine can perform signal timing calculation and
capacity assessment of signalized intersections based on the user-defined inputs, or inputs passed
to it through AMS applications. The QEM Excel tool also uses some additional signal timing
calculation, given in the Signal Timing Manual (12) and Utah Department of Transportation
(UDOT) standards (13). It can work as a stand-alone application, or integrated with NeXTA. The
QEM software package can be downloaded at
https://drive.google.com/?tab=wo&authuser=0#folders/0Bw8gtHCvOm7WTjg0a2FvdlhrRjQ.
The QEM calculation steps are given in Figure 1.

Based on the given inputs, this step first determines the intersection major and minor
approaches, and then sets the initial phase numbers to through and left turns. Once the initial
phasing is selected, the spreadsheet determines the left turn treatment as protected only,
protected plus permitted, or permitted only for each intersection approach. After the left turn

FIGURE 1  QEM calculation steps.
treatment is selected, the dual ring-barrier phase structure is defined, with all phases which are
used in the given case.

The second step uses the given intersection geometry and traffic inputs, and applies the
HCM methodology to compute lane volumes for each lane group. This procedure takes into
consideration the number of lanes for each approach, lane assignment, right-turn-on-red
volumes, and left turn treatment. These volumes are then used in the follow-up calculations.

Once the phasing plan, left turn treatment and lane volumes are known, signal timing
parameters can be computed in the third step as follows:

a) Compute saturation flow rates for protected, permitted and protected plus permitted
movements
b) Compute critical phase volume and lost time
c) Compute critical sum and cycle lost time
d) Determine cycle length
e) Compute phase green times and splits

The critical phase volume ($CV$) is the volume for the movement that requires the longest
green time during the phase. A phase plan summary table, which includes movement codes and
sequences from previous steps, is used to determine the phase plan summary and identify the
critical sequence of movements within the signal cycle. This table also determines the lost time $l$
assigned to each phase. When all phases have been completed, the critical sum ($CS$) of the
critical phase volumes is computed. The cycle lost time $L$ represents the sum of the phase lost
time for each of the critical phases.

Once the critical phases have been determined and $CV$, $CS$ and $L$ computed, the cycle
length is calculated as:

$$C = \frac{L}{\frac{CS}{RS}}, \text{with } C_{\text{min}} \leq C \leq C_{\text{max}} \text{, and } C = C_{\text{max}} \text{ for } CS \geq RS$$

(1)

Where:

- $C$ – Cycle length (s)
- $C_{\text{min}}$ – Minimum cycle length (s)
- $C_{\text{max}}$ – Maximum cycle length (s)
- $L$ – Sum of cycle lost time (s)
- $CS$ – Critical sum (veh/h)
- $RS$ – Reference sum flow rate (veh/h), $RS = 1,530 \times PHF \times f_a$
  PHF – Peak Hour Factor
  $f_a$ – Area type adjustment factor (0.9 for Central Business District – CBD, 1.0
  otherwise)

When the cycle length is determined, the spreadsheet calculates phase split (inter-green)
times $g$ for the phases determined in the ring-barrier structure. These times are determined by the
following equation:

$$g = (C - L) \times \frac{CV}{CS}, \quad g \geq g_{\text{min}}$$

(2)
where $g$ is the phase split time (s), $CV$ is the critical phase flow rate (veh/h), and $g_{\text{min}}$ is the minimum split time, which is set to nine seconds in the algorithm.

Furthermore, the spreadsheet calculates additional phasing data, such as the effective phase green times and capacities based on the HCM methodology, and yellow, red clearance, pedestrian walk and flash-don’t-walk times, based on the methodology described in (12) for the given intersection geometry and traffic characteristics. When all the data are known, the spreadsheet produces the intersection MOEs, which include capacities and V/C ratios for all movements, as well as the critical intersection V/C, control delay per vehicle, intersection LOS and intersection status.

As an important level of service measure, the control delay in the spreadsheet is calculated as the sum of the uniform delay, which is the delay that occurs if arrival demand in the lane group is uniformly distributed over time, and the incremental delay, which accounts for some randomness in vehicle arrivals. The uniform and incremental delays for each lane group are calculated as follows:

\begin{align*}
\text{Uniform delay } d_1 &= \frac{0.5C(1 - g/C)^2}{1 - \text{[min(1, X) \cdot g/C]}} \quad [s] \\
\text{Incremental delay } d_2 &= 900T \left[ (X - 1) + \sqrt{(X - 1)^2 + \frac{4X}{cT}} \right] \quad [s] \tag{4}
\end{align*}

\[ \text{Control delay } d = d_1 + d_2 \quad [s] \tag{5} \]

where $X$ is the lane group V/C ratio, $c$ is the lane group capacity (veh/h), $T$ is the duration of the analysis period (h), and other variables as previously defined. The critical intersection V/C ratio $X_C$ is calculated as follows:

\[ X_C = \frac{CS}{1700PHF f_o \left(1 - \frac{L}{C}\right)} \tag{6} \]

The calculated intersection parameters and MOEs are communicated back to the user, and the interface depends whether the QEM spreadsheet is used as a stand-alone application, or integrated with NeXTA. If used as a stand-alone application, the user can view the results in a specially designed Summary Sheet, or can export them as a PDF report. In NeXTA, the user can select and view the intersection results in the node editor.

4. COUPLING QEM AND DTA SOFTWARE PACKAGES

Figure 2 shows the model coupling, information flow and feedback loop between simulation and signal timing estimation modules.
In order to fully integrate multiple stand-alone models or simulation programs, it is important to standardize data exchange between those tools at different levels of resolution. In addition, open standard will enhance the interoperability of simulations (or applications or systems) to coordinate and work together in a common (virtual) environment. This ability requires the AMS architecture to be built on a common understanding of data schema for traffic modeling and simulation, in particular, transportation network, traffic demand, traffic control devices as well as traffic sensor data. The data formats typically used in the system coupling consist of customized binary files, comma-separated values (CSV), Extensible Markup Language (XML), and database formats. In this research, we use CSV format, which allows a successful and rapid integration with other modeling tools, creating a feasible environment for cross-resolution modeling applications. In Table 1, we provide a summary of different model coupling approaches that can be used to integrate two software systems. In our software prototype, the tight system coupling between the two models is implemented through Excel-based data exchange interfaces, such as Excel Automation. Based on a subset of Component Object Model (COM), Excel Automation is an inter-process communication mechanism created by Microsoft. It has been supported by many programming languages such as C++ and Java.


<table>
<thead>
<tr>
<th>Approach</th>
<th>Key features</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>External Executables/</td>
<td>Use text file as input and output between programs</td>
<td>No change to existing programs</td>
<td>Ad-hoc one-to-one program interface</td>
</tr>
<tr>
<td>Stand-Alone programs</td>
<td>Configuration file or batch file to control iteration flow</td>
<td>Easy programming calling on the same machine</td>
<td>Extensive file exchanges</td>
</tr>
<tr>
<td></td>
<td>Call perform manual iterations using difficult programs</td>
<td>Easy to be debugged by planners/models</td>
<td>Loose coupling between programs</td>
</tr>
<tr>
<td>Static or dynamic</td>
<td>Users provide simulation-callable DLL</td>
<td>Unified programming interface</td>
<td>Relatively limited functionalities</td>
</tr>
<tr>
<td>programming library API, Lib, dynamic</td>
<td>Simulation programs will embed the user-defined library during the execution</td>
<td>Efficient integration for specific functionalities such as signal timing control or vehicle</td>
<td>Require familiarity of advanced programming languages (e.g. C or C++)</td>
</tr>
<tr>
<td>linking library (DLL), OCX, Active X, COM</td>
<td>of application</td>
<td>statistic calculation (triggered by vehicle movement functions)</td>
<td>Only supports data exchange on the local machine</td>
</tr>
<tr>
<td>Message passing and</td>
<td>Use Interface Definition Language (IDL) to specify a mechanism in software</td>
<td>Written in multiple programming languages, integrate large-scale simulation programs running on</td>
<td>Extremely complex programming interface</td>
</tr>
<tr>
<td>inter-program data sharing</td>
<td>for standardizing the method-call semantics between application objects</td>
<td>multiple computers/different platforms</td>
<td></td>
</tr>
<tr>
<td>Common Object</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Request Broker</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Architecture (CORBA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web service in cloud</td>
<td>Service provider pool the processing power of multiple remote computers in</td>
<td>Allow integrated data collection, data sharing and traffic modeling</td>
<td>Only suitable to routine tasks with extensive user debugging,</td>
</tr>
<tr>
<td>computing</td>
<td>a cloud to perform simulation/modeling tasks</td>
<td>Perform simulation tasks might normally be difficult, time consuming, or expensive for an</td>
<td>Difficult for modelers to debug and examine simulation results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>individual user or a small company to accomplish, especially with limited computing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>resources and funds</td>
<td></td>
</tr>
</tbody>
</table>
**TABLE 2** Comparison on QEM and DTA Modeling Assumptions and Components

<table>
<thead>
<tr>
<th>Level</th>
<th>QEM</th>
<th>Typical Mesoscopic DTA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link</td>
<td>Lane-by-lane intersection approach</td>
<td>Lane-by-lane link and intersection</td>
</tr>
<tr>
<td></td>
<td>representation; can represent shared</td>
<td>approach representation; does not represent</td>
</tr>
<tr>
<td></td>
<td>lanes</td>
<td>shared lanes</td>
</tr>
<tr>
<td>Lane</td>
<td>Lane group representation for specific</td>
<td>Lane group representation for links and</td>
</tr>
<tr>
<td></td>
<td>intersection movements</td>
<td>specific intersection movements</td>
</tr>
<tr>
<td>Demand</td>
<td>One hour only (usually peak hour)</td>
<td>Any time period within 24 hours, single or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>multiple days</td>
</tr>
<tr>
<td>Intersection Delay</td>
<td>Control delay (uniform plus incremental</td>
<td>Uniform delay only</td>
</tr>
<tr>
<td>Queue Spillback</td>
<td>delay)</td>
<td></td>
</tr>
</tbody>
</table>

Before performing system-level software integration, we need to systematically compare and understand the modeling assumptions and potential discrepancies between two applications. The detailed comparison on different modeling levels is shown in Table 2. As an intersection analysis tool, QEM requires correct geometric features to perform signal timing estimation. A typical mesoscopic DTA package uses the link-node network representation, so it requires geometric properties of both links and nodes. It can also represent irregular nodes, opposed to QEM that only represents three and four-legged intersections. It is important to recognize that the widely used HCM signal timing parameter estimation and optimization methodology is based on the peak hour in any time period during the day. Therefore, the QEM application only operates with one hour traffic volumes, so the DTA outputs (potentially across different hours of a day) have to be converted to one hour volumes in order to provide consistent input to QEM. In the proposed data flow, we can also use time-dependent loading factors to accurately map demands during the day to different peak periods (e.g. morning peak vs. afternoon peak).

Since QEM is a signal timing application, it can represent all signal timing parameters for the given inputs. A typical DTA package has the option to use QEM cycle lengths and effective green times and provide estimated delays. However, from each run of simulation it can calculate the uniform delay only, without the additional incremental delay caused by randomness in vehicle arrivals. As DTALite uses a deterministic queuing model, when the queue spillback is not considered, its simulated delay (from a point queue model) is equivalent to the commonly used uniform delay formulation defined as the following:

$$d = 0.5 \cdot \frac{\mu \cdot r^2}{(\mu - \lambda) \cdot C}$$  \hspace{1cm} (7)

Where:
- \(d\) – average delay per vehicle (s)
- \(\mu\) – maximum discharge rate (veh/s)
- \(r\) – phase (movement) red time (cycle minus effective green)
- \(\lambda\) – vehicle arrival rate (veh/s)
- \(C\) – intersection cycle length (s)
In the current form, QEM does not calculate queue lengths, and therefore it is not able to provide information on queue spillback and upstream intersection blockage. On arterial streets, DTALite can use spatial queue models that can determine queue lengths and spillback, so it can assess the impacts of lane blockage. From this comparison, it can be seen that QEM and DTALite supplement each other to provide a detailed analysis on any network level.

4.1. DTA and QEM Feedback Loop Convergence

Essentially, the feedback loop in the coupled system operates as the following sequence.

1) Initial traffic flow volumes are calculated from traffic assignment package based on capacity values associated with BPR functions from the regional planning model.

2) As a dynamic, traffic responsive system, QEM determines and changes intersection movement capacities based on the current traffic volumes.

3) With re-estimated movement capacity or signal timing from QEM, a DTA program simulates traffic flow along each path subject to the new capacity constraints at each signalized intersection, and then recalculates experienced travel time along each path. The route switching algorithm will use either MSA or proportional switch to adjust path flow volume at different departure time intervals, which further affect link or movement volumes coming through the QEM controlled intersections. Essentially, the new capacities from step 2 can impact DTA and cause traffic rerouting among available routes.

4) Increase iteration counter by 1. Check the convergence patterns in assignment travel flow volumes, if the link flow changes between two iterations then go to step 2 to estimate the signal timing parameters, and therefore movement capacities.

To test the feedback loop and the simultaneous convergence in the DTA-QEM framework, a simple network with two between one origin and one destination is tested. The network is shown in Figure 3 a), and the inputs are as follows:

- Analysis period = 60 min
- OD demand = 9000 veh/h
- Freeway distance = arterial distance = 3 miles
- Freeway free flow travel time = 3 min
- Arterial free flow travel time = 4.5 min
- Number of lanes freeway = 2
- Number of lanes arterial = 3
- Freeway capacity = 4000 veh/h

One signalized intersection is assumed to located in the middle of the arterial corridor. The cross-traffic at the intersection is set to be 450 veh/h, with no left or right turns. For simplicity, it is assumed that there is no flow in the opposite direction along the main corridor. The current cycle length, main movement effective green time, capacity, V/C ratio and delay are obtained through QEM for each iteration and for the current volume assigned to the arterial corridor, and that movement capacity is used in the next iteration. The feedback loop of assignment-signal timing adjustment is performed for 100 iterations. The convergence patterns are shown in Figure 3 b) – c), where Figure 3 b) shows a comparison of travel times, Figure 3 c)
shows the movement capacity at the intersection, and Figure 3 d) shows a comparison of demand-to-capacity (D/C) ratios.

a) Two-route network inputs

b) Freeway and arterial travel times from traffic assignment program
c) Estimated arterial capacity from QEM

![Graph showing arterial capacity and iteration number]

**FIGURE 3** Two-route network and DTA and QEM convergence.

d) Freeway and arterial Demand/Capacity ratios

![Graph showing D/C ratio and iteration number]
The intersection capacity is the best indicator of how QEM responds to changes in traffic patterns. The intersection capacity stabilizes in the area between 30 and 35 iterations. This is approximately where the other parameters converge, with minimal changes in the consecutive iterations. The full travel time convergence, however, is very sensitive to the route switching route used in typical mesoscopic DTA and the underlying traffic flow models. On the other hand, this is a much more realistic representation of real-world conditions, where small travel time variations are always observed due to the nature of random departure time of vehicles even for the same link capacity and the same path flow pattern. The experiment was repeated for different demand levels, geometric properties, freeway capacities and free flow travel times. The outcomes are very similar, with the signal timing parameters (in terms of capacity and assigned signal timing plan) achieving a good convergence in the area between 30 and 35 iterations. Therefore, to achieve the balance between a typical DTA package and QEM for signal timing estimation, we suggest that at least 30 iterations should be used.

5. WORK ZONE ANALYSIS APPLICATION

The coupling of the two modeling tools through a continuous feedback loop in each consecutive iteration provides a basis for a much more detailed traffic analysis for many practically important network capacity improvement or traffic impact scenarios, such as long-term or medium-term work zones or responsive traffic diversion under severe traffic incidents. Freeway work zones have significant impacts within their area of influence, both on the freeway and adjacent arterials. Majority of tools for work zone analysis can estimate these impacts on the freeway level, but not on the arterial level where the freeway traffic is diverted around the work zone. The integrated DTA and QEM tool can overcome this problem and provide impact analysis on both levels.

The coupled software system was tested on a real-world network in Salt Lake City, UT, shown in Figure 4 a), for a realistic freeway work zone scenario. The network includes a 4.8-mile stretch of I-15, the major arterials in the vicinity, and three freeway ramps. The I-15 corridor is split into 13 segments (S1 – S13) for further analysis. A total of 40 signalized intersections are represented in the network. This network is a subarea of the regional transportation planning network, and the current travel demand model and traffic data are used to create and validate the sub-network. The original OD demand matrices were used as input data, while NeXTA’s OD matrix estimation (ODME) feature was used to update the matrices and perform model validation (14). The current AADT values for the freeway and major arterials were obtained from UDOT, rescaled for the PM period directional traffic, and used in the validation process. A total of twenty-six sensor data points on freeways and arterials was used for this purpose. The model validation plot is given in Figure 4 b).
Three 3:00-6:00 PM models are analyzed and compared in the following three scenarios: (i) the Base model, which represents regular traffic operations; (ii) the Work Zone (WZ) model, where a work zone with a 50% reduction in capacity and a 45 mph speed limit is introduced on the southbound (SB) segment S6 of I-15, between 3900 S and 4500 S streets; and (iii) the Work Zone plus VMS (WZ + VMS) model, where a VMS with an estimated 15% driver response was introduced on the I-15 SB link before the 3300 S freeway ramp. All models were run in the integrated DTA-QEM environment for 50 iterations, to ensure the convergence or reaching reasonable stability between the models.

5.1. Results

By comparing scenarios (i), (ii) and (iii), the work zone and VMS impacts were analyzed for the freeway and signalized intersections. Figure 5 a) and b) shows comparisons of freeway volumes and speeds as a percentage of the free-flow speed, given for SB freeway segments. Figure 5 c) shows a comparison of the average signalized intersection delays.
a) 3hr freeway volumes

b) Percentage of free flow speeds
The work zone impacts are present on both levels. In both WZ scenarios, the freeway volumes changed along the whole stretch of the freeway. The reduced capacity and speed within the work zone area cause time-dependent changes in traffic patterns reported by the DTA model. The implementation of VMS further reduces the freeway volumes. The drop in freeway volumes starts at segment 2, which is located at the 3300 S ramp, where vehicles exit the freeway in search for alternate routes. The volumes are lower throughout the entire work zone, and they start to increase after the ramp at 4500 S. Once passed the 5300 S ramp, the volumes in the WZ scenarios return to regular levels, meaning that all the vehicles that used alternate routes returned to the freeway.

The speeds in the base scenario drop significantly at segments 4 and 5, which are the segments after the 3300 S on-ramp. In the WZ scenarios speeds at these segments are much higher, because of the significantly lower volumes after the 3300 S ramp. However, the speeds within the work zone affected area (within and around segment 6) drop significantly in both WZ scenarios. The speeds in the WZ + VMS scenario are slightly higher in the segments following the work zone. The speeds in the three scenarios converge to the same value after the 5300 S ramp. These results can be used to assess the work zone impacts on freeway traffic.

The analysis of delays at signalized intersections through QEM shows the impacts that a freeway work zone has on the adjacent arterial network. Signalized ramps experience the most increase in delays, which is higher in both WZ scenarios. The increase in delays at intersections just east and west of the freeway and along the adjacent arterials is still significant. On average, with QEM optimization for the new volumes, the delays for the 40 intersections were increased about 12% in the WZ scenario. Figure 5 c) also shows significantly higher delays at critical

**FIGURE 5** Work zone analysis results on freeway and intersection level.
intersections without the QEM optimization, increasing from 30% to almost 500% when compared to the Base scenario. The VMS implementation also has impacts on arterial operations, and the delays are in this case just slightly higher than in the Base scenario. QEM provides detailed information for all signalized intersections, allowing much better assessments of their performance.

6. CONCLUSIONS

Coupling DTA and traffic signal control models is a complex problem, and it has been a focus of numerous research efforts. This paper presents one solution to the problem, where a DTA model, for example DTALite used in our system implementation, is tightly coupled with a QEM Excel-based signal timing estimation tool. The two models operate in a continuous feedback loop in each consecutive iteration until the convergence threshold is achieved. In each iteration, the mesoscopic DTA package passes the intersection geometry and current paths to QEM, which performs signal timing estimation and determines movement capacities, and sends them back to the DTA package for next iteration. This framework provides detailed output data on multiple network levels. A simple network that demonstrates the operation of the feedback loop and the convergence of the two models is first presented, and then the framework is tested on a real-world network with a work zone implementation.

The results presented in this paper show promising results in coupling DTA and signal control through the designed framework. The current application becomes stable after about 30 iterations, when the network parameters converge, with minimal changes in consecutive iterations. The results from a real-world work zone example show that this application can successfully be used to assess impacts of any events that disrupt normal traffic operations and cause route switching on multiple levels, including signalized intersections. Future research efforts will be focused on improving the current models, as well as developing a user-friendly web-based GUI for easier data inputs and visualizations.

ACKNOWLEDGEMENT

The second author of this paper is partially supported through a FHWA research project titled “Effective Integration of Analysis Modeling and Simulation Tools”. Special thanks to our colleagues Brandon Nevers and other engineers at Kittelson & Associates, Inc. for their constructive comments. The authors express their gratitude to the Mountain Plains Consortium University Transportation Center that provided funds and assistance for this study. The work presented in this paper remains the sole responsibility of the authors.

REFERENCES


