

Fiscal Year 2010 Task Reports

final report

prepared for

National Capital Region Transportation Planning Board (TPB)

prepared by

Cambridge Systematics, Inc.

with

Gallop Corporation

This page intentionally left blank.

Table of Contents

1.0	Further Investigation of Convergence in User Equilibrium Traffic Assignment and Speed Feedback.....	1-1
2.0	Short-Term Model Enhancements: Trip Purpose and Special Generators	2-1
3.0	Short-Term Model Enhancements: Time-of-Day Model, Queue Delay Function, and Multi-Run Assignment	3-1
4.0	Short-Term Model Enhancements: Transit-Related Enhancements	4-1

This page intentionally left blank.

Further Investigation of Convergence in User Equilibrium Traffic Assignment and Speed Feedback

■ 1.0 Introduction

As follow-up to the Fiscal Year 2009 Task Reports, the National Capital Region Transportation Planning Board (TPB) tasked Cambridge Systematics (CS) with further investigation into traffic assignment convergence methodologies, with a special focus on the Origin User Equilibrium (OUE) method implemented in Caliper Corporation's TransCAD software. Specifically, TPB has requested information regarding the following topics:

- Discussion and review of the “advanced assignment algorithms” in commercial travel demand forecasting software packages by each of four major vendors: Caliper, Citilabs, INRO, and PTV;
- Consideration of the suitability of using route flows resulting from any user equilibrium assignment methodology;
- Discussion of the adoption of advanced assignment algorithms in regional travel demand forecasting models through contacts with MPOs and planning agencies;
- Description of the use of a hybrid assignment approach as described in the Fiscal Year 2009 Task Reports; and
- Further description of speed convergence metrics, including comments on the suitability of the metric and threshold mentioned in a 2007 presentation by Dr. Howard Slavin.

This report documents CS' reviews and findings on these issues. Section 2 includes a discussion of the suitability of using route flows from user equilibrium assignments. CS found that debate on this topic continues in research circles and no firm conclusion has yet been reached. Bar Gera's proportionality condition appears to have some validity as a means to calculate unique route flows using path- and origin-based assignment algorithms.

Section 3 discusses the assignment algorithms that are currently available or under development. CS surveyed four software vendors in the field about the advanced assignment capabilities of their respective software packages. With completed surveys from all four vendors, CS was able to summarize the available assignment algorithms in each package, along with their capabilities with regards to route flow analysis. The full responses from each vendor are provided in the Appendix. As detailed in Section 3.1, each of the vendors includes both a higher-speed version of the traditional Frank-Wolfe

algorithm in addition to either a path-based or origin-based algorithm. The capabilities and computation speeds of these algorithms vary, and are presented as provided by the vendors. The methods implemented by both Caliper and INRO maintain full route flow capabilities, including select link, select zone, and subarea analysis. Section 3.2 details the use and experiences with non-Frank-Wolfe assignment algorithms at five agencies using a variety of software platforms.

At the request of TPB, further discussion of the hybrid assignment approach is provided in Section 4. This approach has not to our knowledge been implemented at other MPOs or planning agencies, but is based on the practice of using incremental assignment at the Baltimore Metropolitan Council (BMC). CS does not recommend further exploration of the approach for TPB due in part to the recent performance advances in equilibrium assignment approaches.

Finally, a discussion of possible metrics that could be used to measure the convergence of trip tables between feedback loops is provided in Section 5. To the best of our knowledge, the metric recommended by Slavin in 2007 has not yet been adopted by any MPOs, although Caliper has conducted some further testing. However, it is not yet possible to determine the appropriate threshold level for this measure, although it is likely to be dependent on the scale of the travel skims.

■ 2.0 User Equilibrium and Route Flow Analysis

One of the many output data from the traffic assignment process has traditionally included link flows and route flows. Link flows are defined as the total volume on each link in a network. Link flows are insufficient for some common types of analyses that use assignment results, including turning movements at intersections, subarea windowing, and selected link analysis. For these analyses, route flows are required. Route flows identify volumes on a link by origin and destination, and identify all the routes used by each origin-destination pair and the volumes associated with them. However, in static equilibrium highway assignment, only the uniqueness of link flows is guaranteed; there may be multiple sets of route flows associated with the optimal link flow solution.

Recent research into user equilibrium and traffic assignment methods has called into question the nature of route flows in a user equilibrium traffic assignment. Testing has indicated that assignment results may be dependent on the method used to calculate them and potentially the order in which zones are assigned to the network¹.

The issue of the uniqueness of route flows appears to be more of concern with the new families of path- and origin-based assignment algorithms, but as of this writing,

¹ A quick test of reordering the zones in a network reveals if this is the case with a particular algorithm.

discussion amongst researchers and practitioners continues on the issue of the uniqueness of route flow results taken from any user equilibrium assignment algorithm, including the standard Frank-Wolfe (FW) assignment methods. As stated in a recent presentation by Florian (2009), “Regardless of the algorithm used, the only unique results are the total flows and the class impedances.” Recent research as summarized by Caliper (2010) has indicated that:

Unadjusted route-based and origin-based assignments lead to biased select link analysis, and conventional methods may also do so, but to a lesser extent. The reason is that iterative cost updates are made by origin for origin-based methods, and by origin or by origin-destination pair for route-based methods that enumerate routes. This leads to order dependence of the critical link assignment results because the specific new paths added are influenced by selection of prior ones.

Select link analysis based upon the FW or conjugate descent methods is more democratic in that no origin or route is given priority in terms of order of computation. Its defect is that in congested situations, the solutions reflect some unreasonable routes that are added in early iterations after the first set of shortest paths is computed and that are not dropped. This is mitigated, but only somewhat, by achieving good convergence.

The research indicates that all algorithms may be affected by the problem of non-uniqueness, although according to Dr. David Boyce, the linear solutions of the Frank-Wolfe algorithms may be affected least. Arguments continue as to whether all assignment algorithms are affected by this problem (and to what degree), or only specific types. Solutions to this problem, specifically the imposition of the condition of proportionality proposed by Bar Gera (2009), provide a possible method for ensuring the uniqueness of route flows. As the link-based Frank-Wolfe algorithms provide results that are very close to unique, the proportionality condition may not need to be applied to this family of algorithms, but results from origin- and path-based algorithms are still suspect.

■ 3.0 Advanced Assignment Algorithms

Recent advances in computing power have made possible the development of new methods for traffic assignment that reach user equilibrium faster and more tightly than the standard link-based Frank-Wolfe link-based algorithm that has traditionally been used. Throughout this report, the term “advanced assignment algorithms” is used to denote all recently developed algorithms that converge faster and/or to a higher degree than the Frank-Wolfe algorithm. These advanced assignment algorithms, under a range of different names, provide tighter convergence in a shorter time period primarily based on the theory of saving acyclic subnetworks, which leads to more efficient shortest path calculations. These new algorithms fall into two broader families which represent the state-of-the-art in traffic assignment: path-based and origin-based algorithms. Each vendor of travel demand forecasting software has developed a quick-convergence

assignment method in one of these families. In addition, the software vendors have implemented various means of improving the convergence speed of the link-based Frank-Wolfe algorithm. These types of advanced assignment algorithms have only been adopted by a small number of agencies. This section details the availability of advanced assignment algorithms in the available software packages and the use of these algorithms in the marketplace.

As discussed in Section 2, the uniqueness of route flows in user equilibrium results can affect the reliability of many common types of travel demand modeling analyses, most particularly the select link, select zone, and subarea analyses. These tools are important to transportation planners throughout the country, and are used frequently in the Washington region. The functionality of these types of analysis in an assignment algorithm is an important consideration in comparing and analyzing the available options. While the software vendors have developed and released new assignment algorithms with varying levels of select link and select zone functionality, the research community is still uncertain about the accuracy of these claims. This section presents the functionality of each software package as described by the vendor.

Similarly, many tests and studies have been published regarding the processing speed of assignment algorithms when compared with the Frank-Wolfe method. These speed improvements are presented in this report as provided by the vendors. In addition, due to differences in computing power, network size, trip tables, congestion levels, and other test conditions, the test results should not be used to compare the speeds between algorithms tested by different software vendors.

3.1 Availability and Capabilities of Advanced Assignment Algorithms

Four major vendors of travel demand forecasting software (Caliper, Citilabs, INRO, and PTV) were sent a short questionnaire regarding the availability of advanced assignment algorithms in their software packages. In addition, training courses, published documentation, and presentations from recent conferences were used to outline the capabilities of these packages. All four vendors have been working on implementing some form of both a quick-convergence assignment algorithm and a faster version of Frank-Wolfe as shown in Table 1. In addition to the assignment methods shown in Table 1, each of the software packages offers a range of other assignment methods as detailed in the vendor survey responses provided in Appendix B and the respective user guides. According to recent research by Bar Gera (2009), several of the advanced methods are unable to produce unique route flow results, which indicate that these solutions cannot be used to perform a range of route flow based analysis, including select link, select zone, and subarea analysis.

The functionality of the select link/zone analysis tools is still a subject of considerable debate in research circles. Some of the software vendors, including Citilabs and PTV, have opted to disable these tools when using the new assignment algorithms because their accuracy has been questioned. Caliper and INRO, on the other hand, have developed

methods to counteract the proportionality and uniqueness concerns voiced by many researchers and academics which have been incorporated into their software packages.

Level of convergence as measured by “relative gap” is also an important issue in the utility of traffic assignment algorithms. As detailed in this section, almost all of the algorithms converge at least to the recommended level (relative gap of 10^{-5}). However, it should be noted that the level of convergence achieved by an algorithm in a given time period is highly dependent on the size of the network including the trip tables, the level of congestion, and the number of zones. Convergence speed is further affected by the computing environment.

Table 1. Advanced Assignment Methods Available

Vendor	Algorithms Available	Select Link/Zone Functionality*
Caliper TransCAD 5.0 Release 3	N-Conjugate Frank-Wolfe user equilibrium assignment	Fully functional
	Origin user equilibrium (OUE) based on Algorithm B	Fully functional
	Path-based user equilibrium	Fully functional
Citilabs Cube Voyager 5.1.1	Bi-Conjugate Frank-Wolfe user equilibrium assignment	Fully functional
	Path-based assignment using gradient projection method	Disabled
INRO	Parallel standard traffic assignment	Fully functional
Emme 3.3	Path-based traffic assignment	Fully functional
PTV VISUM 11.0	Linear user cost equilibrium (LUCE), an origin-based gradient method based on the origin-based assignment algorithm	Not available
	Path-based equilibrium	Fully functional
	Equilibrium Lohse, a variant of Frank-Wolfe	Fully functional

*As described by the software vendor.

Table 2. Future Improvements

Vendor	Next Release	Expected Date	Included Improvements	Areas of Continued Research
Caliper	TransCAD 6	Beta: Summer 2010 Full release: End of 2010	<ul style="list-style-type: none"> • TransCAD 6 will come in a 64-bit version that will relax memory limits for the largest problems • Additional speed improvements 	<ul style="list-style-type: none"> • Active, but not specified
Citilabs	Cube Voyager 5.2	Fall 2010	<ul style="list-style-type: none"> • A “warm-start” mode for static equilibrium traffic assignment • Performance improvements to path-based assignment, and • Some improvements for Cube Avenue mesoscopic simulation tool 	<ul style="list-style-type: none"> • A path-size logit route choice algorithm has been prototyped • Continue to improve implementation of the path-based assignment algorithm and will continue to collaborate with the academic research community to identify methods of resolving the zone order dependence and proportionality issues • Researching other non-link-based approaches, such as origin-based assignment (OBA) and traffic assignment by paired alternative segments (TAPAS)
INRO	Emme 3.4	Beta: Summer 2010 Full release: Fall 2010	<ul style="list-style-type: none"> • No additional improvements 	<ul style="list-style-type: none"> • Bush-based assignment algorithms
PTV	VISUM 11.5	Summer 2010	<ul style="list-style-type: none"> • Updated specialized assignment variants for toll assignment and for assignment with detailed intersection delays • Warm start implementation of LUCE • LUCE will have an option to turn off bush storage and save memory • All of the post-processing analysis methods (including select link/zone) will be re-implemented to work directly with LUCE 	<ul style="list-style-type: none"> • Plans to refine the method so that the results come closer to overall proportionality

In addition, the survey also collected information about improvements to be included in imminent releases and areas for continued research in the area of traffic assignment. While many of the specifics of software development and strategic research and development directions were not disclosed for business reasons, the survey responses indicate that further improvements in the field of traffic assignment are expected in 2010, and further into the future. Summaries of the improvements expected in the next releases of each software package are included in Table 2.

As assignment algorithms and computing power continue to advance, the next wave of regional models may take advantage of mesoscopic modeling capabilities. Mesoscopic modeling provides many of the benefits of both macroscopic regional modeling and microscopic simulation tools. As shown in Table 3, all of the software vendors already have mesoscopic modeling capabilities through various tools and add-ons.

Table 3. Availability of Mesoscopic Modeling

Vendor	Response
Caliper	Currently available in TransModeler software which is integrated into TransCAD. Dynamic Traffic assignment (DTA) models are also built in to TransCAD. Other mesoscopic models are under development.
Citilabs	Currently exist through the Cube Avenue add-on.
INRO	Dynameq is currently available for dynamic traffic assignment and mesoscopic modeling.
PTV	Plans to integrate mesoscopic simulation that incorporates junction detail in the near future; most likely to include Mezzo mesoscopic simulation with VISUM.

Full results of the vendor surveys can be found in Appendix B of this report. Included below are summaries of the survey results, outlining the available assignment algorithms and their capabilities as provided by their vendors. The responses are presented in alphabetical order by vendor.

3.1.1 *Caliper*

Caliper offers a range of assignment methodologies in its TransCAD software. The Origin User Equilibrium (OUE) method is based on Algorithm B and has been “evolved to handle all of the MMA² procedure features and options.” Very high levels of convergence

² MMA or multimodal, multi-class assignment refers to TransCAD’s Frank-Wolfe based assignment algorithm.

are possible with the use of OUE. The OUE method also allows for the use of a warm-start which can compute a solution based on a previously achieved equilibrium assignment. This method can offer substantial time savings for running multiple model runs or performing a series of speed feedback loops. For some very large and complex problems, OUE can require more than 2 gigabytes (GB) of memory, which is not available in the current TransCAD release for 32-bit Windows. A 64-bit Windows version will be able to access the larger amounts of memory necessary to solve assignments that have very large networks, a dense zone structure, and/or a large number of assignment classes.

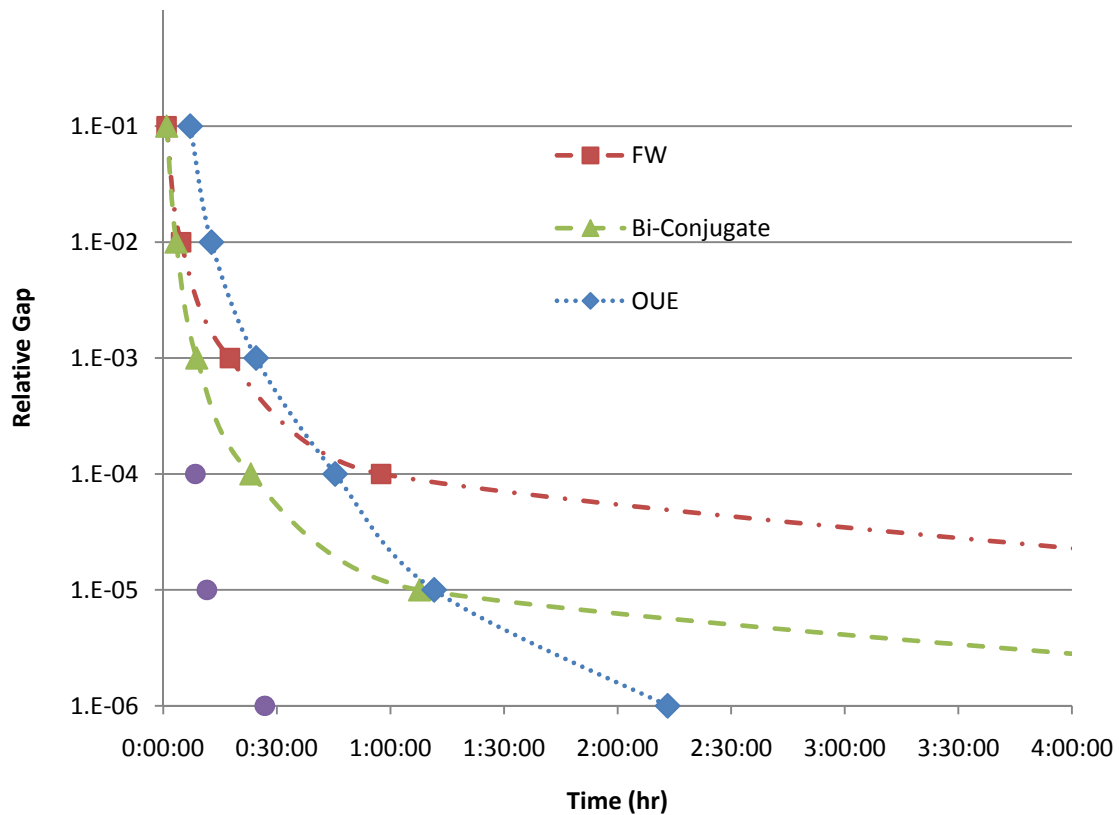
The OUE algorithm as implemented in TransCAD 5.0, release 3 allows for a full range of post-assignment analyses, including select link and select zone. TransCAD incorporates the “maximum entropy” solution to calculate the most likely route flows as a means for ensuring the uniqueness and proportionality of the route flows in an OUE assignment.

TransCAD also includes the N-Conjugate Frank-Wolfe algorithm, which is a higher speed version of the traditional Frank-Wolfe assignment method. This method speeds the convergence process significantly according to tests documented by Caliper (2010) – by a factor of two for low levels of convergence and by a factor of six for a relative gap of 10^{-5} . Multi-threading of the Frank-Wolfe assignment algorithms is available in TransCAD, which can substantially speed convergence.

Performance tests conducted by Caliper on their available assignment algorithms indicate that the major benefit of OUE is the ability to achieve substantially higher levels of convergence in a shorter time than using other methods. As shown in Figure 1, both the standard and Bi-Conjugate Frank-Wolfe algorithms tail out after a certain level of convergence, while OUE does not. For a cold start, the time savings offered by OUE are really only significant past a relative gap of 10^{-5} . However, as shown in Figure 1, for subsequent assignments, OUE warm starts offer the possibility of even shorter run times.

Caliper indicates that research continues in a number of areas related to traffic assignment. However, the biggest improvement expected for the next release of TransCAD is the advent of a version for 64-bit Windows to relax memory constraints and allow OUE to be used on larger and more complex problems.

Figure 1. Comparison of Available Assignment Algorithms in TransCAD 5.0



Note: Model runs include five user classes. Four threads are used for the Frank-Wolfe and Bi-Conjugate methods.

3.1.2 Citilabs

Citilabs offers a range of assignment algorithms in its Cube platforms. Currently, Citilabs recommends the use of the Bi-Conjugate Frank-Wolfe method because it offers fast convergence and maintains all the functionality and proportionality of the standard assignment procedures. Using this algorithm does not require changes to any inputs, including data and network files. According to Citilabs:

This option also enables additional computations during the ADJUST phase to more accurately solve for LAMBDA, resulting in fewer path-building iterations required to reach higher-precision convergence. Bi-Conjugate performs better than Conjugate Frank-Wolfe and is recommended by Citilabs as the best method for achieving high precision user equilibrium assignments without loss of proportionality in results.

Citilabs has also implemented a path-based gradient projection assignment algorithm which provides quick results to a high level convergence, but is not yet recommended for

use in most modeling applications. This is primarily based on the uniqueness issues previously mentioned that have not yet been satisfactorily solved. According to Citilabs:

This algorithm is among the fastest user equilibrium assignment methods implemented in Cube Voyager. However, due to the zone order dependence inherent in this method... intra-step distributed processing of path-based assignments using Cube Cluster is not possible at the current time. Furthermore, academic research has shown that non-link-based traffic assignment methods... do not preserve proportionality in results, suggesting that although the same link flow solution is obtained by these algorithms, the path flow solution is not unique and may in fact be faulty. It is partly for this reason that select link/zone analysis and turning movement as well as path file outputs are disabled in our implementation of path-based assignment.

Citilabs questions the proportionality associated with all of the non-link-based assignment algorithms including their path-based algorithm and origin-based algorithms. Primarily, these issues relate to the zone-based incremental loading process used in the algorithm, which makes the results highly dependent on the size, structure, and numerical order of the zone system.

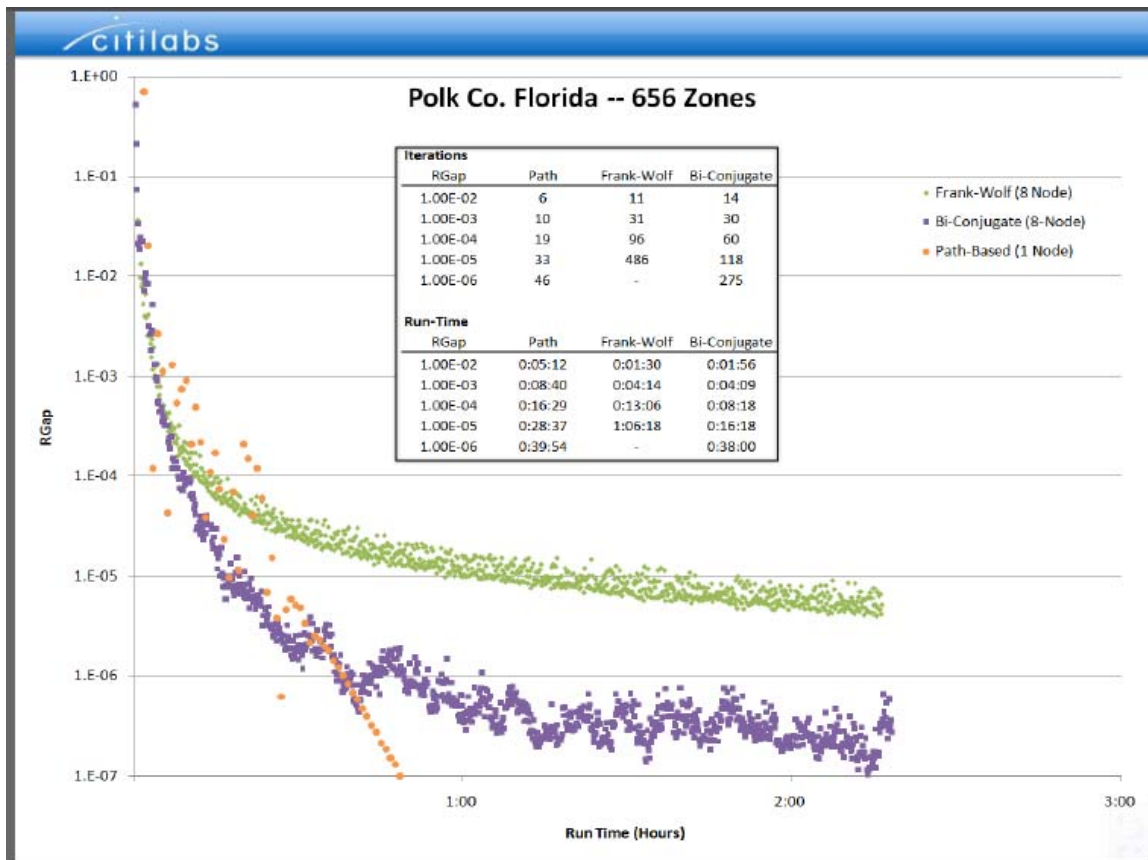
Testing was conducted by Citilabs comparing the performance of three assignment algorithms in Cube Voyager when applied to the model for Polk County in Florida. The results of these tests are shown in Figure 1 using an eight-core processor. (The path-based algorithm cannot use multiple cores at the current time.) As shown, the Bi-Conjugate algorithm consistently outperforms the standard Frank-Wolfe method for relative gaps less than 10^{-3} . The speed advantage of the path-based algorithm is only visible at even higher levels of convergence.

The path-based algorithm currently available in Cube does not include the possibility for warm-starts, which may further improve performance. Research into this possibility continues, and is expected to be complete for the release of Cube Voyager 5.2 in the Fall of 2010.

Cube Cluster is an add-on to Cube Voyager that enables intra-step and multistep distributed processing of highway assignment on multiple computer processors. Cube Cluster allows the user to completely control the allocation of processing workload to available processors on a local area network. However, currently the path-based algorithm cannot make use of the distributed processing capabilities of Cube Cluster. Citilabs does not currently plan to support multi-threading, since hardware for distributed processing is currently cheaper and easier to obtain. The distributed processing of Cube Cluster can be used to provide speed improvements to the Frank-Wolfe based assignment algorithms, including the Bi-conjugate method.

According to Citilabs, Cube Voyager will run on any modern Windows-based machine, including 64-bit as well as 32-bit XP, Vista, and Windows 7 operating system versions. Other hardware/software requirements were not indicated, except that it was noted that multiple computers may be taken advantage of with Cube Cluster.

Figure 2. Citilabs Test Results of Available Assignment Algorithms



Source: Citilabs, 2010.

3.1.3 INRO

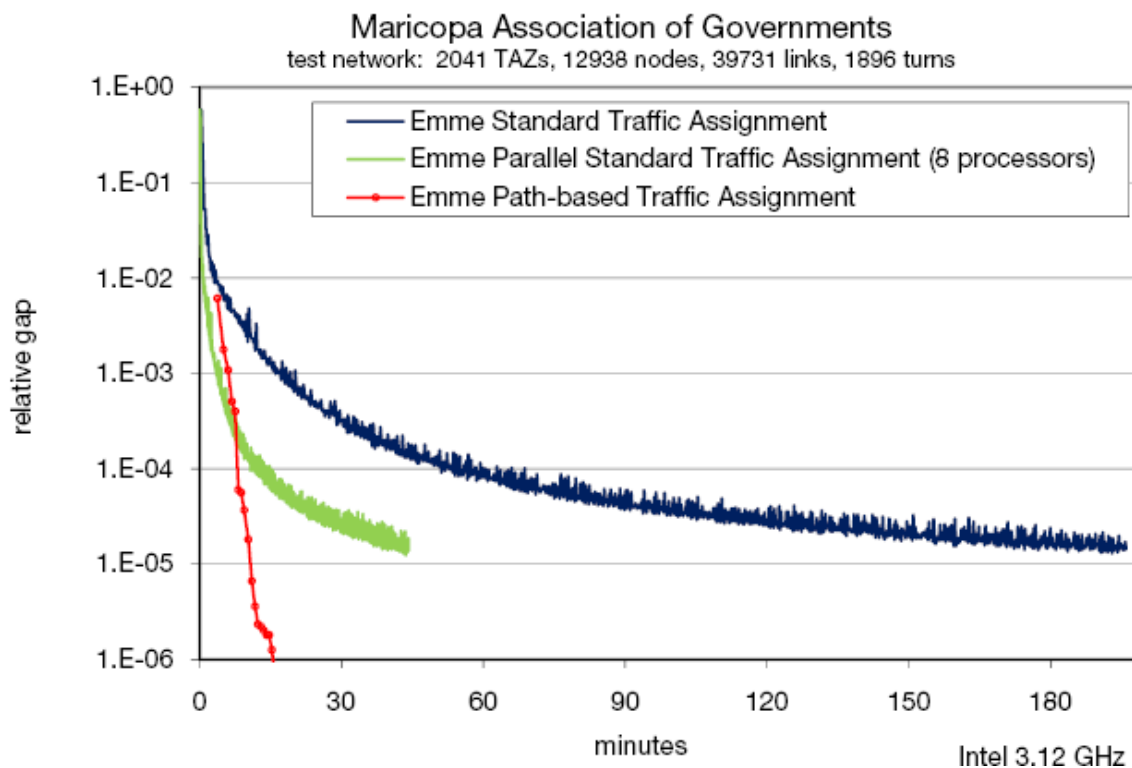
Emme offers several static assignment methodologies including a version of the Frank-Wolfe algorithm that uses parallel processing and a new path-based assignment algorithm. The parallel processing algorithm offers speed improvements over standard Frank-Wolfe algorithms through the use of multiple computer processors and converges to approximately 10^{-4} before tailing off. Path storage enables high speed path analysis such as select link/zone, as well as warm starts for further performance improvements.

The path-based algorithm provides substantial speed and convergence improvements over the link-based methods. Testing results of the assignment algorithms available in Emme show that the path-based algorithm converges faster than other methods available in the software package as shown in Figure 3. Testing also indicates that for modest levels of convergence (relative gap of 10^{-2} or less) the Frank-Wolfe algorithms are still faster, although this level of convergence is not recommended by INRO.

Processor speed and input/output speed will have the greatest affect on assignment speed for all of the assignment algorithms in Emme. The parallel Frank-Wolfe algorithm requires multiple processors to achieve the greatest effect and large applications typically

require a few hundred megabytes of memory. However, the path-based algorithm functions best with enough memory to store all the generated paths in physical memory during the assignment process. For very large complex network, this can require several GB of memory.

Figure 3. Example Comparison of Convergence Times for Assignment Algorithms Available in Emme



3.1.4 PTV

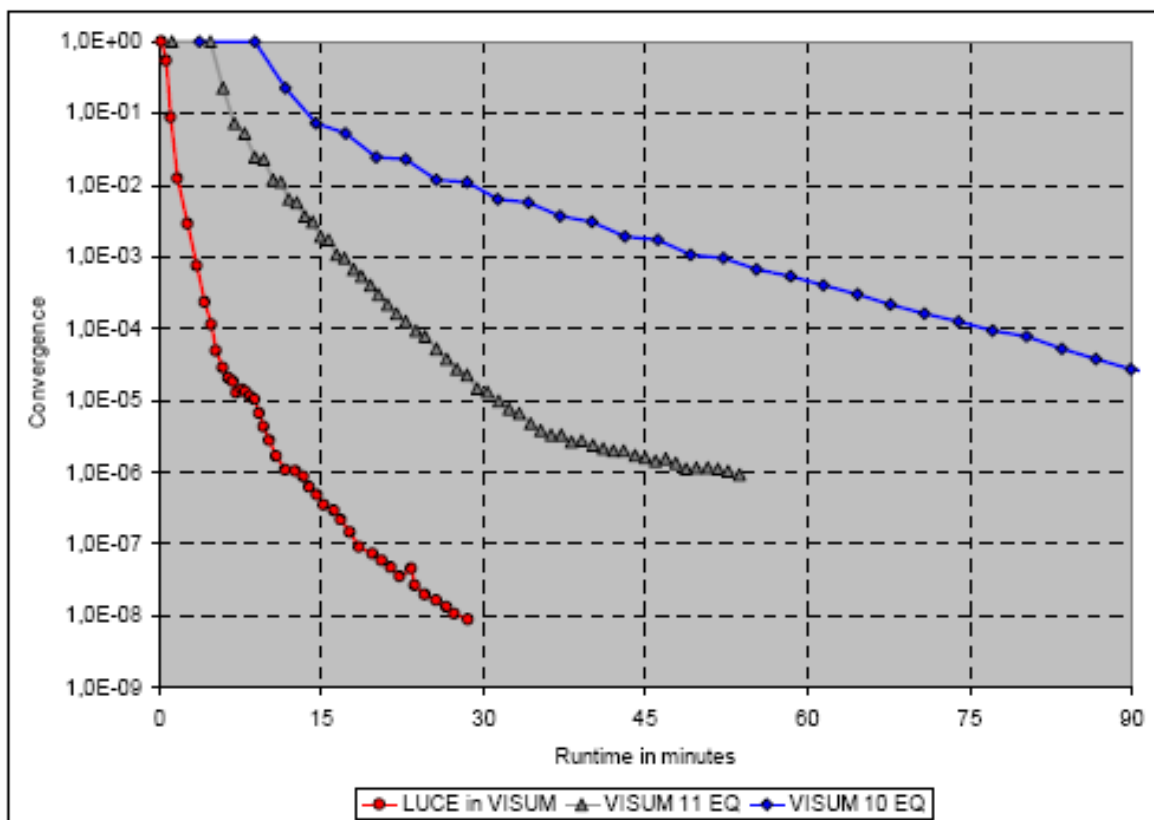
PTV offers a range of both static and dynamic traffic assignment methods in its VISUM 11.0 platform. The static assignment method currently recommended by PTV is the Linear User Cost Equilibrium (LUCE), an origin-based assignment (OBA) technique developed as a further improvement to OBA. The LUCE algorithm is structured so that path data is stored in bushes, ensuring “that at least all of the paths for one origin satisfy the proportionality condition” (PTV, 2010). According to PTV:

At this stage PTV provides LUCE as a prototype in VISUM 11, intended mainly for evaluation purposes. It does run on realistic networks, but it currently has some technical limitation: most importantly post-assignment analysis (skimming, path listing, matrix estimation,...) is not possible yet. Currently all these operations post-process paths, and are therefore not available in LUCE, which – through the improved proportionality – loads too many paths to be stored in memory.

PTV is working to lift most of these limitations for the next release of the VISUM software so that the necessary post-processing analyses will be possible. VISUM will also be implementing a warm start for LUCE as of the summer of 2010.

PTV has conducted testing on the LUCE algorithm in comparison with other assignment methods. As shown in Figure 4 with a test on the Chicago regional model, LUCE converges to high levels (10^{-7} or less) substantially faster than Frank-Wolfe based algorithms available in VISUM.

Figure 4. Comparison of VISUM Assignment Algorithm Speed Performance



Source: PTV, 2010.

VISUM also offers a path-based equilibrium assignment method that does not converge as quickly as LUCE but maintains all of the functionality of the Frank-Wolfe assignment method, including select zone and select link. In addition, VISUM includes a link-based improvement to the Frank-Wolfe method called Equilibrium Lohse. According to PTV:

For applications where relatively modest relative gaps must be reached very quickly, VISUM contains a variant of Frank & Wolfe (called Equilibrium Lohse after the original researcher). It shares the properties of F&W observed generally,

i.e., it can be multithreaded easily, loads a rich path set, and achieves good convergence speed down to $\sim 10^{-3}$, but tails off after that.

PTV makes the following recommendations regarding hardware/system requirements:

- 64-bit Windows with 8-16 GB of memory for large-scale models; and
- Multi-core machines can cut run times since many of the computationally intensive calculations are multi-threaded.

3.2 Advanced Assignment Algorithms in Regional Planning Models

Based on experience working with regional planning models throughout the country, CS staff identified a number of agencies and MPOs that use or have investigated the use of selected advanced assignment algorithms. These agencies were contacted to help catalog national experience to date. Five agencies responded and were interviewed by CS staff. While the focus was on agencies utilizing TransCAD’s OUE algorithm, CS attempted to contact agencies using advanced assignment methods in each of the available platforms. As shown in Table 4, five agencies using three software vendors responded to our interview requests and agreed to answer questions about their assignment processes. The experiences of each of these agencies are summarized in this section.

Table 4. Agencies Interviewed by Modeling Software Platform

Caliper TransCAD	Citilabs Cube/TRANPLAN	INRO Emme	PTV VISUM
SANDAG: MMA Prince George’s County, MD: OUE NCTCOG: MMA	LA Metro (TRANPLAN): Bi-Conjugate Testing	PSRC: Gradient Projection	None interviewed

3.2.1 Caliper Users

Caliper provided CS with the names of three agencies that use the TransCAD OUE assignment method in their regional model: Prince George’s County, MD; Victoria, British Columbia, Canada; and the Whatcom Council of Governments in Bellingham, WA. Only the Prince George’s County model is of a size comparable to TPB, as the others are of substantially smaller regions. CS contacted many large agencies which use TransCAD and was able to interview three large agencies which are TransCAD users, although only one is currently using OUE for highway assignment. In addition to the agencies listed in

Table 4, the Southern California Association of Governments (SCAG) in the Los Angeles Metropolitan area indicated that significant testing had been completed by Caliper using their model and the OUE algorithm although we were not able to collect any additional information.

Prince George's County is currently using TransCAD 5.0 and has been using the OUE algorithm for traffic assignment for approximately three years. During this time, staff has reported no problems or issues with the assignment methodology. As discussed in Slavin et al. (2006), much of the work to implement the OUE algorithm, including the calibration and validation, was conducted by Caliper directly, and Prince George's County still works closely with consultants to run and update the model. Because Prince George's County is modeling the whole TPB area using a version of the TPB model, this example provides an especially valid comparison to TPB's current model operations. The model includes over 2,500 zones, five assignment classes, and three time periods.

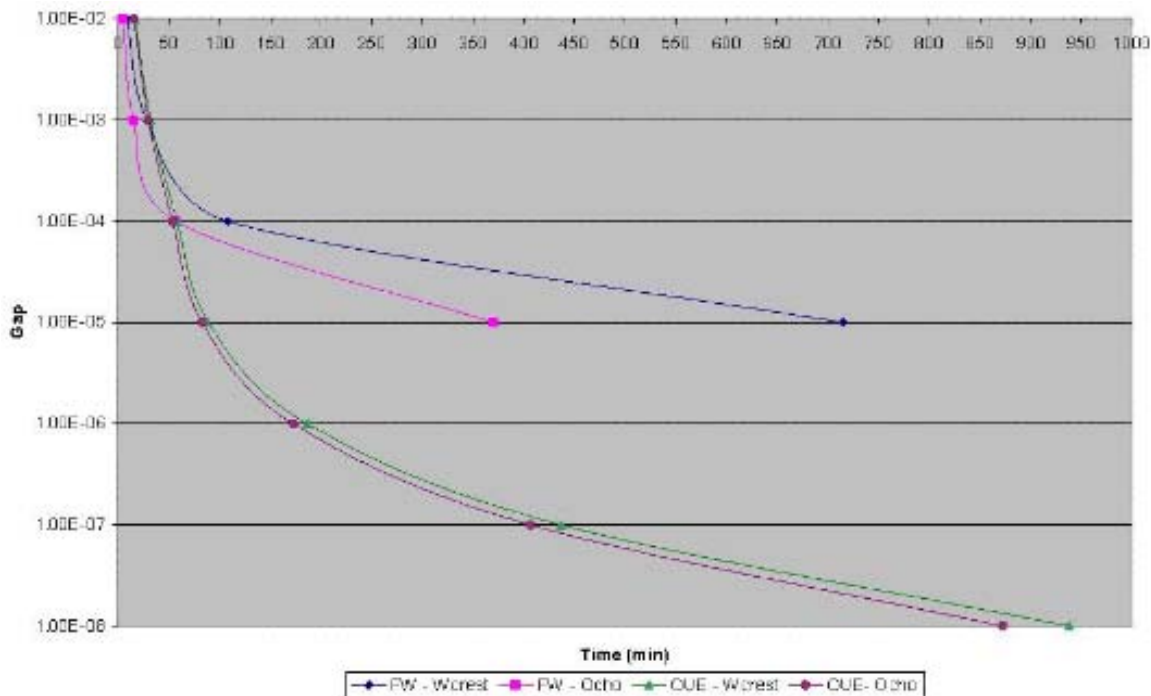
Prior to implementation of OUE, Prince George's County was using the standard Frank-Wolfe multi-modal multi-class assignment (MMA) in TransCAD. As documented by Slavin et al. (2009), the switch to OUE produced assignment results that are very well calibrated and match observed data throughout the region very well. Small network changes result in fewer unrelated changes to the assignment solution (or "noise"); this is likely related to the higher convergence achieved by OUE.

According to Prince George's County staff, the current model reaches convergence of 10^{-5} in 10-12 hours requiring approximately 150 iterations. Test results from Caliper shown in Figure 5 for an evening peak assignment illustrate the speed and convergence benefits of OUE when compared to Frank-Wolfe assignment methods. The blue and pink lines represent Frank-Wolfe assignments (on four- and eight-core machines respectively), while the green and purple lines represent OUE assignments (on four- and eight-core machines respectively). As shown, additional processing power can substantially improve the performance of the Frank-Wolfe algorithm, although similar gains are not available when using OUE. The specifics on the user class(es) being portrayed in the figure are not indicated, but it is assumed that the results are typical.

While a warm start would be possible using OUE, Prince George's County has not used this option recently. Recent test results from Caliper – shown in Table 5 – indicate that substantial time savings upwards of 90 percent can be realized using a warm start.

Prince George's County staff use a post-processor program developed by Caliper to run select link and select zone analysis. The post-processor allows these types of analysis to be done without rerunning the whole assignment and typically takes only a few minutes to complete.

Figure 5. Prince George’s County Run Time Comparisons



Source: Slavin, H., J. Brandon, A. Rabinowicz, and S. Sundaram, 2009.

Table 5. Prince George’s County Warm Start Test Results

Run Description	Time To Converge (h:m:s)	Percent Improvement
Cold Start	1:28:02	--
+/- 5% perturbation run 1	0:08:51	89.9%
+/- 5% perturbation run 2	0:08:53	89.9%
+/- 5% perturbation run 3	0:08:45	90.1%
+/- 10% perturbation run 1	0:11:10	87.3%
+/- 10% perturbation run 2	0:11:18	87.2%
+/- 10% perturbation run 3	0:10:00	88.6%

Source: Slavin, H., J. Brandon, A. Rabinowicz, and S. Sundaram, 2009.

The San Diego Association of Governments (SANDAG) currently uses the Frank-Wolfe MMA assignment algorithm in TransCAD 5.0. The current model uses a convergence criterion that requires a relative gap of 0.001 before assignment stops. Table 6 highlights the number of iterations currently required to achieve this level of convergence for a 2050

model run. A full model run takes approximately 16 hours; 2-3 hours are required just for the assignment step.

Table 6. SANDAG Model – Iterations to Convergence (2050)

Time Period	Iterations
Morning Peak	79
Evening Peak	120
Off-Peak	31

Source: Data provided by SANDAG.

SANDAG has discussed possible use of the OUE assignment method with Caliper and determined that the switch would be unfeasible at the current time. The SANDAG model, which includes almost 4,700 zones and 8-11 assignment classes, cannot run on the 32-bit TransCAD 5.0 platform due to memory allocation issues. Once the 64-bit TransCAD 6.0 is released, SANDAG is likely to investigate further the possibility of using OUE. SANDAG was able to test OUE using five sample mode tables, but was not able to use any more assignment classes successfully.

North Central Texas Council of Governments (NCTCOG) also uses TransCAD 5.0, but has done no formal testing of the OUE assignment methodology. Currently, after significant testing of various volume delay functions, the NCTCOG model runs in approximately 16 hours, which has been determined by MPO staff to be an appropriate time period as it allows for an overnight model run. The model currently converges to a relative gap of 10^{-4} in between 100-600 iterations (depending on the time period); each iteration typically takes on average 20 seconds. There is some interest in pursuing OUE in the future if the algorithm can substantially improve run times, however it is not on the “critical path” for NCTCOG model improvements.

3.2.2 *Citilabs Users*

The Los Angeles County Metropolitan Transportation Authority (LA Metro) is an agency CS contacted which is using Citilabs products and is investigating the use of advanced assignment algorithms. LA Metro uses TRANPLAN for their travel demand forecasting. According to LA Metro staff, the agency has been very focused on FTA New Starts modeling in the recent past, and has not made much use of the highway assignment capabilities of their model. The agency’s official travel demand forecasting model uses the standard Frank-Wolfe assignment method and there are no immediate plans for switching. However, LA Metro staff indicated a need to analyze tolling policies in the future, which will require a new focus on highway assignment. LA Metro staff did not identify a timeline for this need.

LA Metro is implementing the Frank-Wolfe Bi-Conjugate assignment method recommended by Citilabs in the model being used for their Nexus Study. This study is investigating the effects of a countywide development impact fee program, and is not using the official LA Metro model. The study model is not yet complete and so results and experiences using the Bi-Conjugate algorithm are not yet available.

3.2.3 *INRO Users*

The Puget Sound Regional Council (PSRC) in the Seattle region has been using the path-based traffic assignment methodology available in the Emme software platform for approximately 12-18 months. The PSRC model includes over 1,100 zones, five time periods, and 11 assignment classes and reaches convergence when the normalized gap or the relative gap are less than 0.02. Using six feedback loops, the gradient projection assignment methodology has improved model run times in the base year by more than 50 percent compared to the Frank-Wolfe algorithm previously used, from 20.4 hours to 10 hours. Convergence for each time period is reached within 25-50 iterations. It should also be noted that PSRC has elected to have their model only calculate the gap measures every five iterations; this results in the model being run for up to four additional iterations once the convergence criteria have been reached. While the gradient projection algorithm has the capability for a warm start, PSRC has not yet made use of this feature.

In addition to convergence and run time statistics, PSRC also analyzed other aggregate measures of regional travel calculated by the model and compared the new algorithm to the Frank-Wolfe algorithm used previously. As shown in Table 7, vehicle miles traveled (VMT), vehicle hours traveled (VHT), and average speed measures throughout the network change only minimally with the new assignment methodology. Table 8 shows the observed changes in regional mode shares for work trips in the PSRC region; non work mode shares showed no change. Additional comparisons show that total link flows throughout the network are similar between the two methods and reveal only slight changes in the regional travel time distribution. These measures, developed and supplied by PSRC staff, indicate that results developed using the gradient projection assignment algorithm are very similar to those developed using Frank-Wolfe in a fraction of the time.

Because Emme saves all of the path results of the assignment process, PSRC is able to quickly and conveniently perform select link and select zone analysis without rerunning the assignment. In addition, these saved results allow staff quick access to skims on any attribute for each of the assignment classes. Based on these capabilities and the results, the substantial time savings realized, and testing results, PSRC indicated that they are happy with the new assignment method.

Table 7. Percent Change in PSRC Model Equilibration Measures

Dimension	VMT	VHT	Average Speed
Morning Peak	-0.2%	0.0%	-0.3%
Midday	0.3%	0.5%	-0.3%
Evening Peak	0.3%	1.2%	-1.0%
Evening	-0.5%	-0.7%	0.3%
Night	-0.1%	-0.1%	0.0%
Daily	0.0%	0.4%	-0.3%
Freeways	0.1%	0.8%	-0.7%
Arterials	0.0%	0.3%	-0.3%
Connectors	0.0%	0.0%	0.0%
All Facilities	0.0%	0.4%	-0.3%

Source: PSRC and Florian, 2009.

Table 8. Change in Mode Share in PSRC Model

Mode	Frank-Wolfe	Gradient Projection
Single Occupancy Vehicle	80.1%	80.3%
Carpool	7.2%	7.2%
Transit	8.1%	8.0%
Transit-Walk	6.6%	6.6%
Transit-Auto	1.5%	1.3%
Bike	1.6%	1.6%
Walk	3.0%	3.0%

Source: PSRC and Florian, 2009.

■ 4.0 Hybrid Assignment Approach

The hybrid assignment approach outlined in the Fiscal Year 2009 Task Reports compilation was reviewed further at the request of TPB. The proposed hybrid approach would use equilibrium assignment to a high level of convergence for the first and last iterations of the speed-feedback loop process and use incremental assignment for the intermediate iterations. To our knowledge, the approach as described is not used by other MPOs. However, the Baltimore Metropolitan Council (BMC) uses an assignment

approach which mixes aspects from incremental and equilibrium assignment techniques with the goal of obtaining more stability in assignment results.

In testing done with an earlier version of the TPB travel demand forecasting model, the equilibrium assignment (even without reaching a stable converged state) produced more accurate traffic volumes than an incremental assignment. The test confirmed that although incremental assignment is a faster algorithm to apply, it should not be used for a final assignment process.

The ideas of the hybrid assignment approach were as follows:

- Use incremental assignment only for the intermediate assignment loops to save processing time by decreasing the number of iterations used in these loops;
- Provide stable skim times from these intermediate iterations for inclusion in the final iteration;
- Do not use the link volume averages from prior iterations for the final iteration, instead the final trip table would be assigned directly using the final equilibrium assignment; and
- Apply a tight convergence criterion for the final assignment to ensure stability across the network.

This approach was originally postulated as a stop-gap method to be considered as a means for reducing the model run time. However, the theoretical issues with incremental assignment coupled with the recently enhanced ability to reduce the run time of equilibrium assignment methods with a high level of convergence through hardware, software, and algorithm advances suggests that no further consideration should be given to the hybrid assignment approach. Better options exist for achieving the goal of shortening the model run time and improving the accuracy of the assignment.

■ 5.0 Speed Feedback Convergence Metrics

As indicated in the Fiscal Year 2009 Task Reports, there is no measure used as state of the practice to show convergence between speed feedback loops. This type of convergence would be measured in addition to network volume convergence, which is typically measured using relative gap or some related measure. Only the Denver Regional Council of Governments (DRCOG) was found to measure this type of convergence at all, using the criterion of achieving one percent or less of links with a greater than 10 percent change in link volume. Another possible measure of convergence between feedback loops has been proposed by Slavin (2007) as the “skim matrix root mean square error.” This metric measures the difference between skim matrices in adjacent feedback loops. As convergence is reached, the difference between the skim matrices should decrease, indicating increasing stability between loops. As stated by Dr. Slavin, the use of both this

metric and the relative gap convergence method for traffic assignment creates a fixed point solution for the travel demand forecasting problem.

To our knowledge, since its introduction in 2007, no agency or MPO has adopted this measure for feedback convergence. CS was able to find no evidence of additional testing of the metric beyond that initially conducted by Caliper for the original presentation. However, according to Dr. Slavin, Caliper has been using this metric in all projects since 2007, although no further papers have been written on the subject. The appropriate threshold level for this criterion, whether one percent, 0.1 percent, or some other value has not been determined, although Caliper's work has set the value at 0.1 percent. The appropriate threshold level for this measure may be different for different model sets. The appropriate value may be highly dependent on the value of the average skim time. For example, while a one percent difference may be acceptable on a matrix with an average travel time of 100 minutes, it may be unacceptable if the average travel time were only ten minutes. Since experience with this measure is limited, it is not possible to apply a standard threshold to all models. Testing within the framework of the TPB model would be necessary to determine an appropriate threshold for this model. Dr. Slavin also notes that not all models will converge and some may converge to an incorrect solution. These affects also need to be considered when determining whether this measure should be used with the TPB model, and what threshold level should be implemented.

■ 6.0 Acknowledgments

We would especially like to thank all the agencies and companies who took the time to speak with us, including:

- Los Angeles County Metropolitan Transportation Authority (LA Metro);
- North Central Texas Council of Governments (NCTCOG);
- Prince George's County Maryland National Capital Park and Planning Commission;
- Puget Sound Regional Council (PSRC);
- San Diego Association of Governments (SANDAG);
- Caliper Corporation;
- Citilabs;
- INRO; and
- PTV.

■ 7.0 References

Bar Gera, H. (2009). Practical Implications of Finding Consistent Route Flows. Presentation at the 11th TRB National Transportation Planning Applications Conference. <http://www.trb-appcon.org/TRB2009presentations/s14/01_bar_gera_nie_and_boyce_finding_consistent_route_flows.ppt>

Caliper Corporation (2007). Travel Demand Modeling with TransCAD 5.0: User's Guide.

Caliper Corporation (2010). What TransCAD Users Should Know about new Static Traffic Assignment Methods.

Citilabs (2010). New Assignment Methods in Cube Voyager. Webinar presented March 11. <<https://www1.gotomeeting.com/register/363881992>>

Dial, R. (2009). Comparative Performance Strategy and Tactics Genealogy. Presentation at the 11th TRB National Transportation Planning Applications Conference. <<http://teachamerica.com/APP09/APP09S14Dial/player.html>>

Florian, M. (2009). Changing Assignment Algorithms-- The Price of Better Convergence. Presentation at the 11th TRB National Transportation Planning Applications Conference. <<http://teachamerica.com/APP09/APP09S14Florian/player.html>>

Paschai, B., K. Yu and A. Mirzaei (2009). Increasing Precision in Highway Volume through Adjustment of Stopping Criteria in Traffic Assignment and Number of Feedbacks. Presentation at the 11th TRB National Transportation Planning Applications Conference. <<http://teachamerica.com/APP09/APP09S14Yu/player.html>>

PTV (2009). LUCE in VISUM.

Slavin, H. (2007). Achieving Planning Model Convergence. Presented at the 11th TRB National Transportation Planning Applications Conference. Daytona Beach, Florida. <<http://teachamerica.com/App2007/02Slavin/index.htm>>

Slavin, H., J. Brandon, A. Rabinowicz, and S. Sundaram (2009). Application of Accelerated UE Traffic Assignments. Presentation at the 11th TRB National Transportation Planning Applications Conference. <<http://teachamerica.com/APP09/APP09S14Slavin/player.html>>

Slavin, H., S. Sundaram, and A. Rabinowicz (2006). TransForM- A New Regional Travel Demand Model Developed for Prince George's County. Caliper Corporation. March, 2006. <<http://www.caliper.com/Library/PGCModel.pdf>>

Zhou, Z. and M. Martimo (2009). New Empirical Study of Alternative Traffic Equilibrium Algorithms. Presentation at the 11th TRB National Transportation Planning Applications Conference. <<http://teachamerica.com/APP09/APP09S14Zhou/player.html>>

■ **Appendices**

Appendix A: Glossary

Appendix B: Vendor Survey Responses

This page intentionally left blank.

Appendix A

Glossary

This page intentionally left blank.

■ Appendix A: Glossary

Frank-Wolfe Assignment Algorithm: An algorithm developed in 1956 to assign traffic to a congested network in order to achieve a state of user equilibrium in which no traveler can improve their travel time by selecting a different route. The Frank-Wolfe algorithm is link-based, in that it seeks to minimize the travel times on each link in the network to the extent possible. This algorithm is iterative and serves as the basis for traffic assignment in travel demand forecasting software.

Link Flows: Total traffic volume on any link in a network.

Multi-modal Multi-Class Assignment (MMA): A generalized cost assignment routine in TransCAD that allows for individual modes or user classes to be assigned to the network simultaneously. The assignment routine allows each mode or class to have different characteristics. The MMA assignment in TransCAD is based on the Frank-Wolfe algorithm.

Relative Gap: A metric used to measure the convergence of a traffic assignment. Relative gap measures the relative difference between the total travel time in the network and the total travel time if all travelers were using the shortest path.

Route Flows: Route flows identify volumes on a link by origin and destination, and identify all the routes used by each origin-destination pair and the volumes associated with them.

Warm Start: A method for using the results of a previous traffic assignment to inform the solution of another traffic assignment. The warm start essentially uses the paths and/or volumes previously calculated as a starting point for the next assignment. This can save a substantial amount of computation and therefore time.

This page intentionally left blank.

Appendix B

Vendor Survey Responses

This page intentionally left blank.

■ Appendix B: Vendor Survey Responses

B.1 Caliper

The attached document which we have just prepared for TransCAD users will provide a great deal of information on some of the static assignment methods in the released TransCAD 5. Other responses to your questions can be found below.

1. What is the latest version of your software and what assignment algorithms are available in it?

TransCAD 5 Release 3 is current and Release 4 is forthcoming in May. The accompanying document and the user manual provide the answer to this question.

2. Do the algorithms converge to a relative gap of 10^{-5} ?

Of course, but this is not a good question. If your question is would they do so for the MWCOG network and trip table, the answer is most assuredly yes. Of course, CS can easily try out the different software options with MWCOG data with TransCAD and other commercial software. Another key point is that it is important to have methods that go to much lower relative gaps. This makes it possible to identify the appropriate level of convergence for different assignment problems and different model applications.

3. What, if any, are the hardware requirements to run these algorithms?

There are no special requirements, but the performance of the algorithms is improved by increasing the number of cores available.

4. When is your next release expected? Will it include any additions or enhancements to the assignment methodology?

TransCAD 6 will be in beta this summer and released before the end of the year. TransCAD 6 will come in a 64-bit version that will relax memory limits for the largest problems. There are likely to be additional speed improvements in TransCAD 6. There will definitely be many other enhancements to the assignment methodology, but we do not wish to make the details public at this time.

5. Are any additional enhancements or additions to the assignment methodologies currently planned for research and/or development?

Yes. Caliper has a very deep R&D pipeline but we do not usually comment on future developments until products are nearly ready for release.

6. Are there plans for developing mesoscopic modeling capabilities?

These have existed for quite some time in our TransModeler software which is integrated with TransCAD. We have other meso models under development and other DTA models that are practical for large planning networks.

7. Is there other information surrounding this topic and your software that you believe we should be aware of?

Yes, and some of it is discussed in the writeup for TransCAD Users and the references listed there. Also, we have many DTA options that are available with varying degrees of fidelity and computational complexity. Our hybrid traffic simulation technology is also applicable to planning models, which should be known to you as CS has used our hybrid simulation in a number of projects.

What TransCAD Users Should Know about New Static Traffic Assignment Methods

Recent research by Caliper and others has led to a variety of new and improved methods for calculating user equilibrium traffic assignments. New options provide much faster computing and also the achievement of much tighter convergence, resulting in more accurate impact assessments and select link analysis.

Even before any of the improvements discussed below, TransCAD was significantly faster in computing user equilibrium than any other commercial software. We believe that the further improvements have widened TransCAD's lead in computing speed considerably.

Greater speed is needed because of the desire to do larger and more complex problems, and greater accuracy is needed because, without it, there can be large errors in estimates of the impacts of plans and projects. The discussion below purposely limits technical detail so as to be accessible to a wide audience of users and managers, but there are numerous references that provide a full discussion of prior research and the topics discussed. Also, we at Caliper would be happy to discuss any questions that you might have about these matters.

Multi-threading of the MMA Assignment

Most large regional models use the MMA (multi-mode, multi-class assignment) routine in TransCAD. This model accommodates HOV lanes, toll roads including those with entrance-to-exit tolls, multiple user classes and class prohibitions, trucks of different sizes, and varying values of time, and is appropriate for problems that need one or more of these features.

For the conventional UE assignment, based upon a well-tuned Caliper implementation of the Frank-Wolfe (FW) algorithm, multi-threading in TransCAD results in a nearly proportional speedup in computation that is a function of the number of physical cores in a computer's CPU. Figure 1 illustrates the speed of convergence for a rather large and congested network assignment problem for the metropolitan Washington, D.C. region with 2500 zones, over 36,000 links, and 5 user classes. We use this particular traffic assignment problem as a test case because it is from a real, deployed model and matches ground counts closely. It is much more challenging from a computational point of view than all the test networks used in the research literature. The presence of 5 user classes results in much longer running times than would be experienced for a single class assignment.

We use the relative gap (RG) measure of convergence, which is a common and reasonable figure of merit, for the presentation of assignment convergence results in the discussion here. Most models in the U.S. have traditionally used a relative gap of .01, but that is insufficient for impact assessment, and as this has become more widely known, increasingly tighter assignments are being computed. Nevertheless, gaps below .001 are rarely encountered for large models due to concerns about computing time.

The computer used to generate these results has two quad-core Xeon CPUs that run at 2.93 GHz. It was purchased more than a year ago and does not have the hyper-threading featured on newer chips from Intel. At that time, it was one of the faster options available. For some of the tests, we disabled one of the CPUs and/or some of the cores.

As one can see from Figure 1 below, there is a significant improvement due to multi-threading, yielding a nearly proportional speedup in the FW MMA assignment as a function of the number of CPU cores. Doubling the number of cores generally halves the amount of time it takes to reach a given level of convergence.

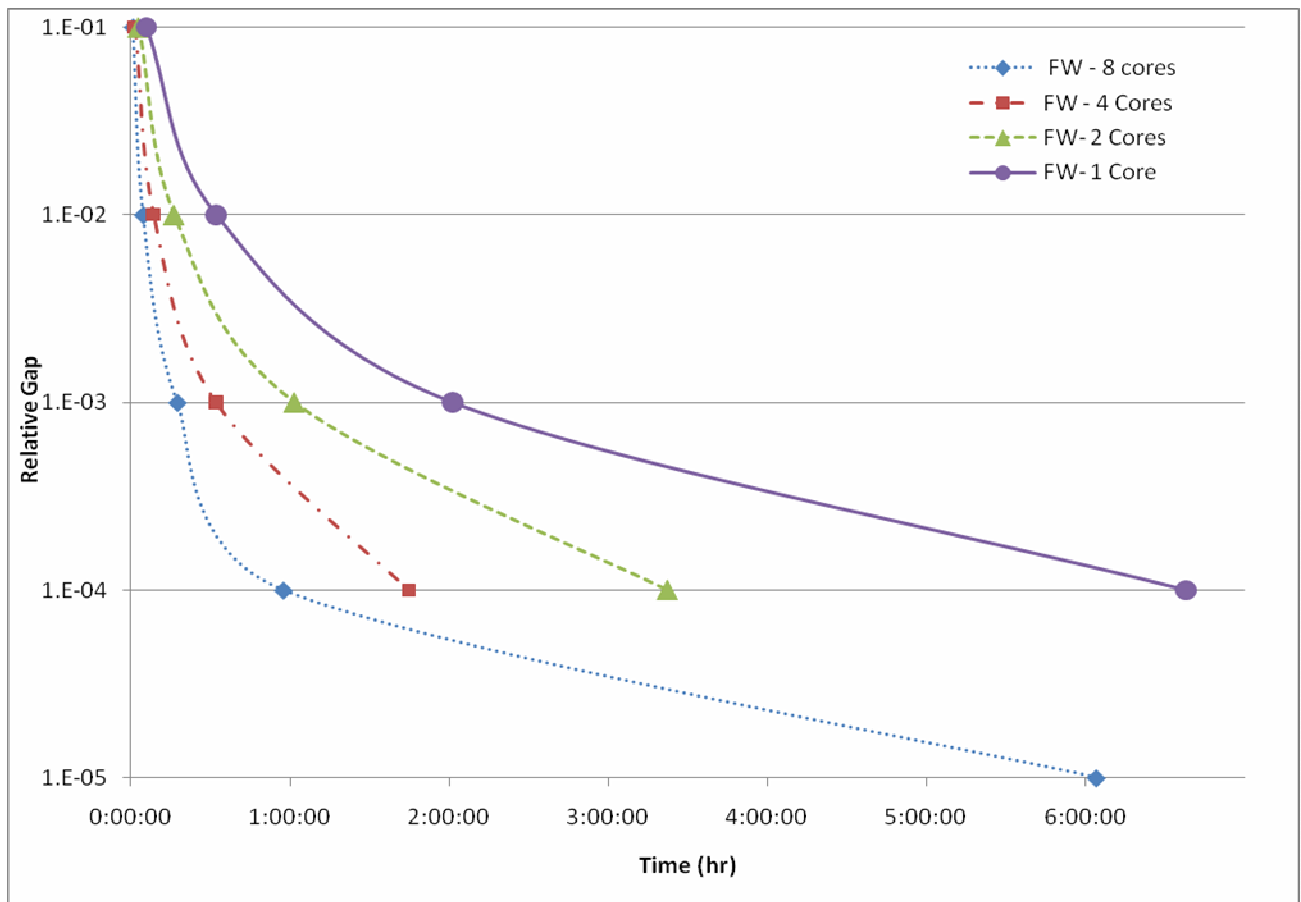


Figure 1: Convergence graph for the MMA FW assignment with 2, 4, and 8 computer cores

With the new 6 and 12 core processors coming on the market, even further gains will be easily achievable, and there is always the possibility of using computers with 2 to 4 or more CPUs. Also, newer chips from Intel have more effective hyper-threading, which adds additional improvement, although not as much as additional CPUs or cores.

About five years ago, we implemented distributed processing for the MMA assignment. We have since dropped it, as it is much less efficient than multi-threading due to data communication overhead.

Irrespective of the number of cores or the speed of the assignment, you can also see that the FW assignment method's rate of convergence tails at a certain point, an observation well-known to both theoreticians and practitioners. This means that this method will not be able to achieve orders of magnitude lower convergence. You may find it interesting to know that it takes the FW method 650 and 4145 iterations to reach relative gaps of .0001 and .00001, respectively.

Conjugate Descent Options Added to MMA

Proposed by Daneva and Lindberg (2003), the bi-conjugate descent FW (BFW) method uses a little more memory than the conventional FW assignment, but not so much that it would typically be a concern today. FW holds two link flow vectors in memory where conjugate descent methods keep 3 or more link flow vectors in memory, which are used in choosing a more effective search direction than FW. Conjugate descent methods are easily multi-threaded, so there is no tradeoff in using them. These options were added in Release 3 of TransCAD 5. Figure 2 shows the running times for the test network with the bi-conjugate FW method with 1, 2, 4 and 8 computer cores.

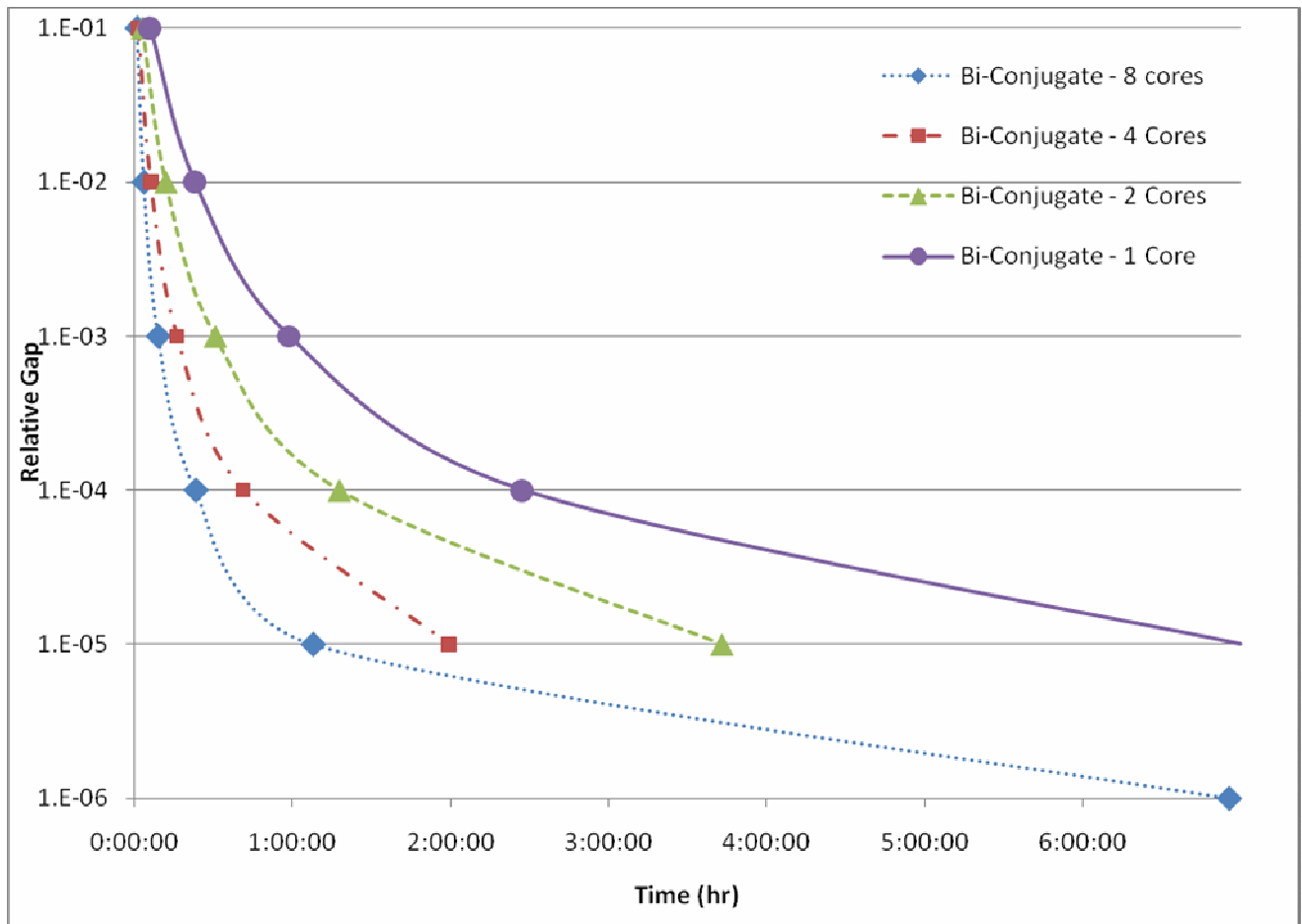
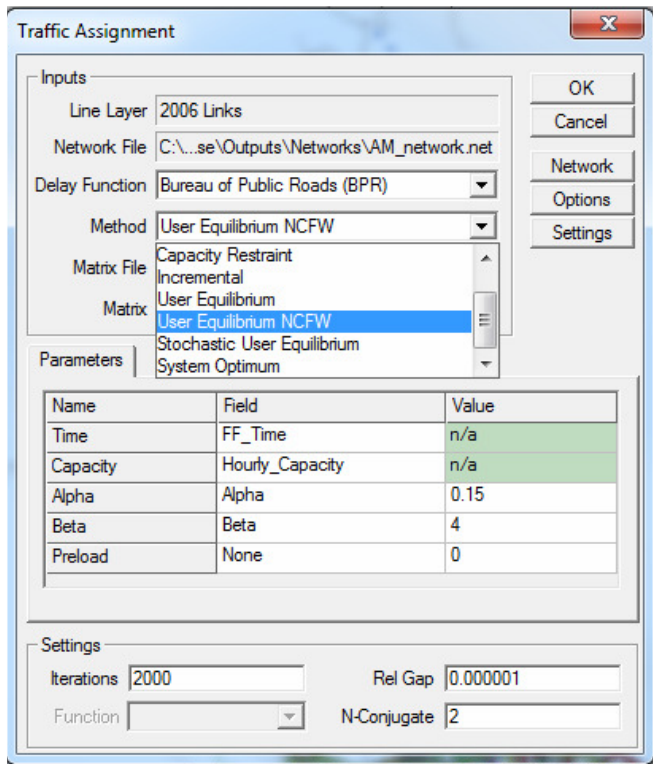


Figure 2: Convergence graph for the Bi-conjugate traffic assignment method with 1, 2, 4 and 8 computer cores

The results indicate a very significant improvement in efficiency using the bi-conjugate Frank-Wolfe method. Two things can be noted. First, the bi-conjugate method, as implemented in TransCAD, cuts the running time by a factor of 2 or more at low convergence and by a factor of 6 or more in terms of the time taken to reach a gap of 10^{-5} . It makes computing to that level realistic for this assignment problem. It can also be observed that multi-threading is very effective for this method, too, so further improvements in running time can be obtained by using more powerful computers.

If you are using TransCAD interactively, you can use the new bi-conjugate traffic assignment option by choosing it from the traffic assignment method pull-down list in the Traffic Assignment dialog box and setting the N-conjugate value to 2 as shown below.



If you want to run the new bi-conjugate assignment method in batch mode, change the traffic assignment macro code from

```
Opts.Global.[Load Method] = "UE"
```

to

```
Opts.Global.[Load Method] = "NCFW"
```

```
Opts.Global.[N Conjugate] = 2
```

In our experiments, we found the bi-conjugate option just as effective as the k-conjugate method for $k > 2$. However, because each network assignment problem can be different, some users may wish to experiment with these options.

A New Algorithm—OUE Assignment

Another way to solve an MMA UE problem is by using the OUE assignment option that was introduced in TransCAD 5 and has been evolved to handle all of the MMA procedure features and options. This method is based upon an algorithm developed by Robert Dial who worked with Caliper on its implementation. Dial's Algorithm B does not tail like a conventional assignment and, as a result, it can achieve unprecedented levels of convergence and do so quickly. The OUE assignment generates a solution for each origin's link flows and therefore requires more memory than FW and BFW methods. We chose to feature this method in TransCAD after evaluating and testing various other new UE assignment methods for which favorable claims were made in the literature. Specifically, we believe that it is significantly faster than route-based, projected gradient methods and other origin or bush-based methods and has other advantages as well. Importantly, it is not only faster, but unlike the FW or conjugate descent method, it drops inappropriate routes as it generates successive iterative solutions, yielding a cleaner assignment solution. In Figure 3, we show how OUE based upon algorithm B performs on the Washington Regional Network. You can observe that OUE benefits from multiple cores, but not nearly so much as the other assignment methods. The reason is that OUE is not fully multi-threaded at the present time.

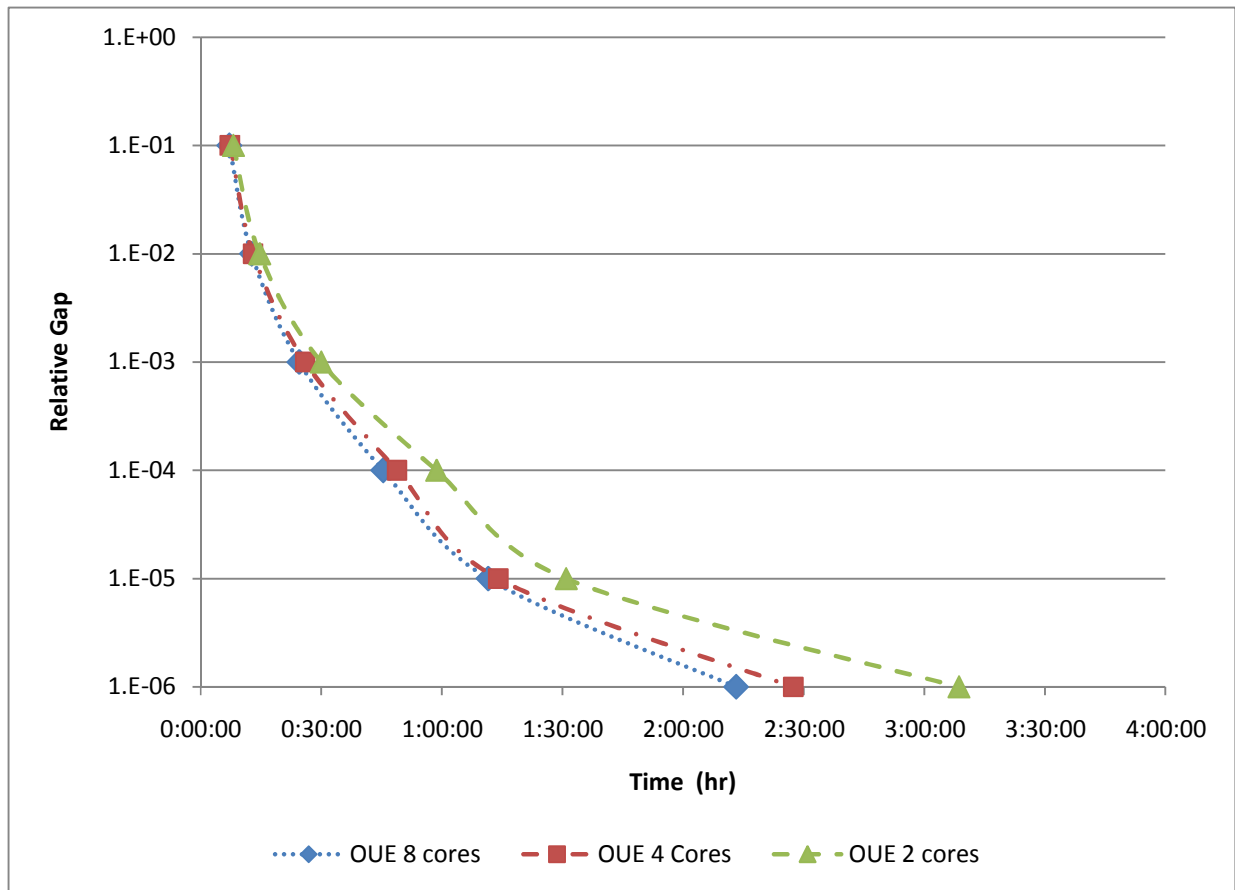


Figure 3: OUE Cold Start Convergence Rates PM Assignment

Figure 4 compares OUE with the FW and BFW methods. It illustrates the superior rate of convergence of OUE; however, it is not faster when started from scratch than BFW because of the multi-threading benefit afforded by the 8 cores until relative gaps are sought that are below 10^{-5} . For lower relative gaps, BFW tails but OUE does not.

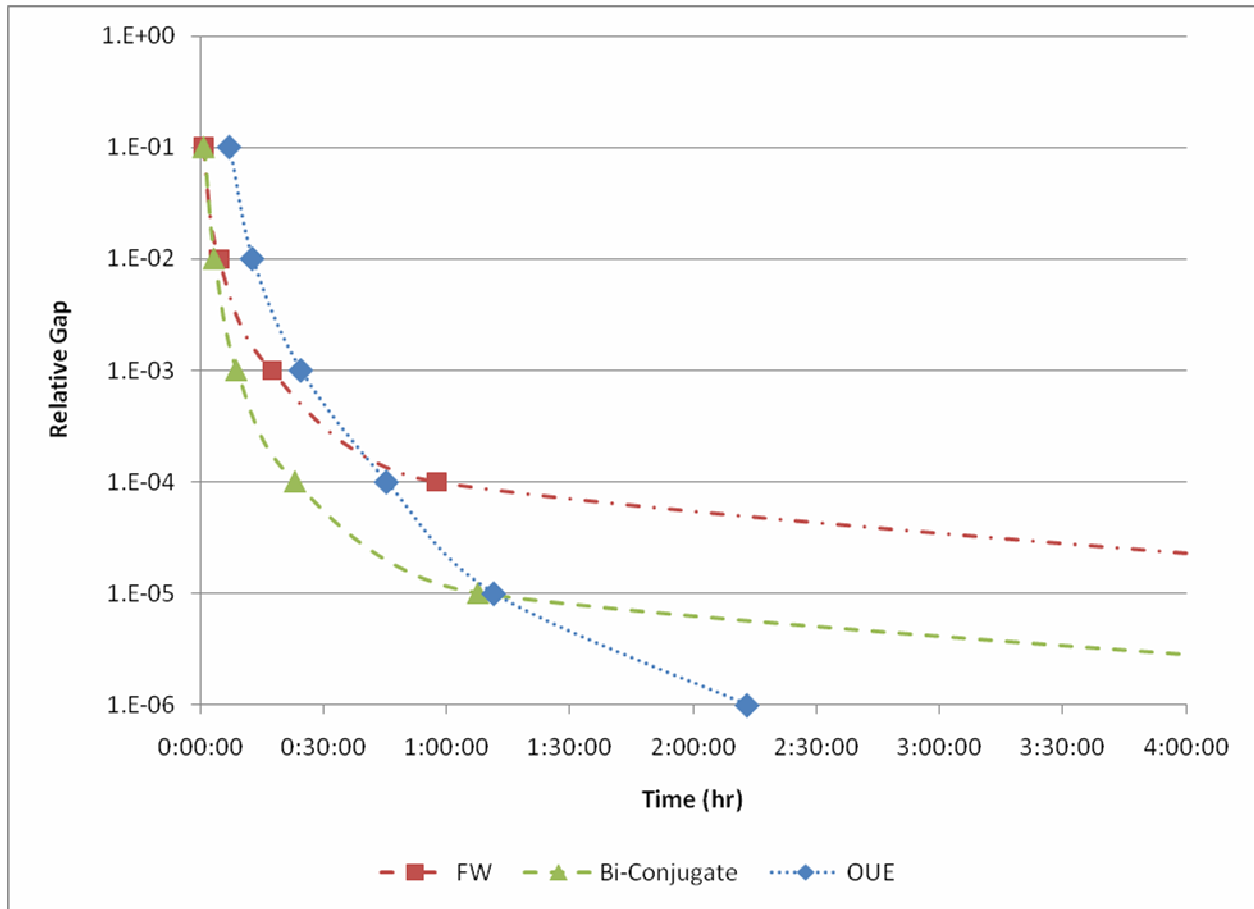


Figure 4: Cold Start Convergence Rates for the PM Assignment using 8 cores

When started without a prior solution or, in other words, from a cold start, OUE reaches a relative gap of .00001 in 1.5 hours on the machine with 4 cores and is only somewhat faster on the machine with 8 cores. This is acceptable for practitioners in comparison to the 12.5 hours that FW takes to get to the same level of convergence on the 4 core computer or the 6 hours for the 8 core computer. When compared to the BFW method, OUE is faster on this problem to 10^{-5} on a quad core computer, takes a similar amount of time on an 8 core computer, and can achieve orders of magnitude tighter convergence. In addition, OUE has further advantages to which we now turn.

Immediate Availability of Select Link Analysis as a Post Process

When the results are saved after the computation of an OUE assignment, select link analysis can be performed for any query. This means you do not have to specify the query before running the assignment, and you can save a great deal of time by not having to redo traffic assignments when another query is required.

Warm Start Speedups with OUE

A particularly important aspect of OUE (and Algorithm B) is that it can re-compute a new equilibrium from a prior solution, even if the trip table or network attributes have changed somewhat. Prior solutions are nearly always available because traffic assignments are computed over and over again in the course of model development. Of course, this requires that a previous solution has been saved, and there is some computational overhead associated with writing and reading a prior solution.

To illustrate this warm start advantage, we computed the test assignment to a relative gap of 10^{-5} with a quad-core computer and then investigated how long it takes to compute a new solution to the same level of convergence for a similar, but not identical trip table. As shown below in Table 1, when we randomly perturbed the trip tables for the Washington regional model by + or – 5%, we were able to calculate a new equilibrium to a gap of .00001 in less than 9 minutes compared to the 1.5 hours it takes from a cold start. With a 10% random perturbation, the time grows but only to about 11 minutes.

	Time to converge
Cold Start	01:28:02
+/-5% perturbation run 1	00:08:51
+/-5% perturbation run 2	00:08:53
+/-5% perturbation run 3	00:08:45
+/-10% perturbation run 1	00:11:10
+/-10% perturbation run 2	00:11:18
+/-10% perturbation run 3	00:10:00

Table 1: OUE Warm Start Convergence to 10^{-5} with Perturbed Trip Tables

From these results, it is clear that there is a massive efficiency gain from using the warm start which, for this model, yields roughly an order of magnitude improvement in computing time. With the warm start, the calculation of a tight equilibrium takes essentially no more time than calculating FW to a 1% relative gap.

In a model with feedback loops, trip tables change at each iteration while the network remains fixed. The warm start is particularly effective in this situation. In Table 2, we show the calculation times for a model run with 4 loops and traffic assignment convergence to a relative gap of .0001 and to a travel time skim RMSE of .1%.

Model Steps	Loop 1	Loop 2	Loop 3	Loop 4
All other Steps	25 min	16 min	16 min	16 min
AM Assignment	31 min	6 min	5 min	5 min
PM Assignment	1 hr 1 min	8 min	7 min	6 min
MD Assignment	35 min	9 min	6 min	6 min
Loop Time	2 hr 32 min	39 min	34 min	33 min

Table 2: Model run times (4 cores) with OUE assignment (0.0001 RG)

As is evident, the time needed to compute the traffic assignment steps declines significantly after the first loop and declines further although less so with successive loops. Loop 4 takes only one-fifth of the time taken for the first loop. Of course, one could use a warm start for the first loop using a prior solution, which would further reduce the overall running time necessary to achieve feedback convergence.

The warm start can also be used when the network changes somewhat. Generally speaking, if the prior solution is feasible for the modified network, the warm start will be effective. If the prior solution is not feasible, it may not work or, if it does, it may be much less efficient.

Inexplicable Link Flow Changes Need Not Be a Feature of your Model

Research that we have performed has confirmed other findings that poor convergence leads to large errors in traffic assignments and, thus, in all aspects of planning models and impact assessments. (Slavin et al, 2009, Boyce et al, 2004). In particular, poor convergence accounts for the implausible changes in link flows that are often observed in traffic assignments at locations that are remote from the specific projects being evaluated.

As an example, we edited the test network southeast of D.C. to reflect the opening of a new flyover ramp from the Capital Beltway to southbound Route 5 in Prince George’s County. Route 5 is heavily traveled and is used by more than 122,000 vehicles per day. The new ramp was intended to reduce congestion and ease a difficult merge at the interchange.

In Figure 5, we show the differences between the base case flows and the flows with the new interchange for different relative gaps. At a 1% RG, there is significant convergence error, and the convergence error dominates the project forecast. The green links represent those that had a gain in flow and red links are those that lost flow of more than 200 vehicles. While the flows on the new links are similar, there are spurious effects far away from the project.

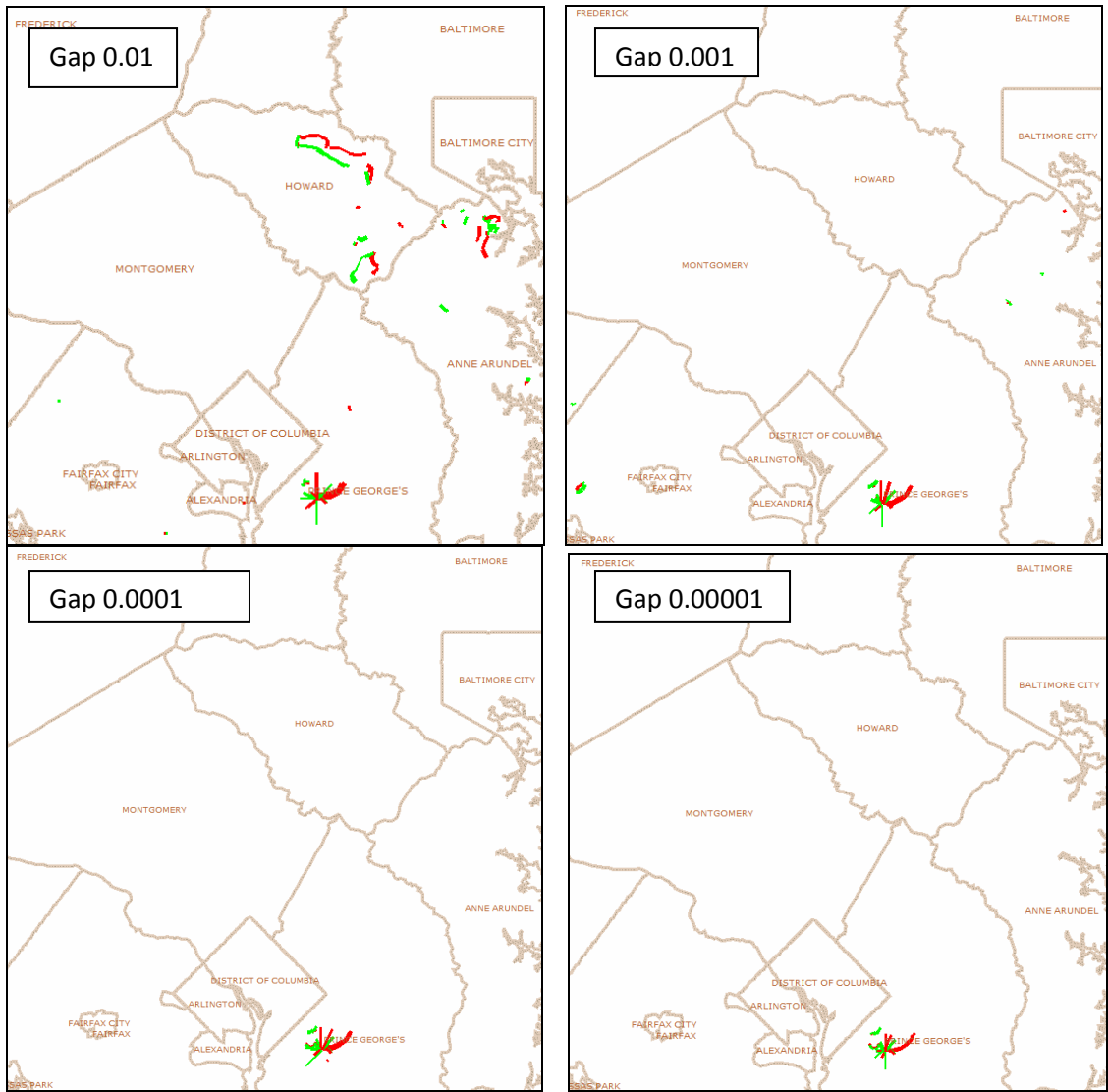


Figure 5: Links with flow changes greater than 200 vehicles – Interchange Project

Tighter convergence cleans up and localizes the projected impact. The spurious link changes are mostly gone at a RG of .0001, but you can also see that the solution is different and cleaner at a RG of .00001.

Importance of Convergence for Impact Analysis

The calculated VHT impacts show that poor convergence can lead to erroneous conclusions about project benefits. At a gap of 0.01, the model predicts that the VHT will actually increase by 408 vehicle hours after the modification of the interchange. At a gap of 0.001, the model still indicates a VHT increase of 94 vehicle hours. At a smaller gap of 0.0001, the model predicts a VHT saving of 20 vehicle hours. At an even tighter RG, the VHT savings are negligible, which may be due to replacing older ramps with longer ones.

Table 3 indicates the number of links in the scenario that have an absolute flow change of more than 200 vehicles from the base case, VHT for the base case, VHT for the scenario, and the VHT differences.

Gap	Number of Links with Abs_Flow_Diff > 200	VHT Base Case	VHT Scenario	VHT Saving (Veh-Hrs)
1e-2	162	1,105,726	1,106,134	-408
1e-3	56	1,091,153	1,091,247	-94
1e-4	44	1,090,136	1,090,116	+20
1e-5	45	1,090,090	1,090,088	+2
1e-6	47	1,090,084	1,090,079	+5
1e-7	45	1,090,087	1,090,085	+2
1e-8	45	1,090,089	1,090,086	+3

Table 3: Comparison of Base and Scenario – Interchange Project

Not only are estimates of project impacts affected by convergence levels, but so are overall VHT estimates. This means that it is important to run both base cases and forecasts to the same level of convergence, so as to avoid biased calculations.

While more research will be needed to establish general guidelines, evidence accumulated to date by us suggests that a relative gap of .0001 or better should be used for accurate results. This can be tested for specific models rather easily using TransCAD by computing OUE to high convergence and then assessing the link flow errors associated with lesser convergence.

The tests that we have conducted illustrate clearly that estimates of project benefits vary substantially with the convergence level utilized in the traffic assignment. The link flow errors are not randomly distributed, but will typically reflect a particular geographic bias associated with forecast use of some routes instead of others.

The bottom line is that achieving good convergence in the traffic assignment is essential for obtaining valid estimates of project impacts.

Why does Poor Convergence Plague Models?

One reason that poor convergence plagues most deployed models is that the disturbing and hard-to-explain spurious impacts are not very noticeable when evaluating an entire future plan scenario. It is much easier to identify problems when one can check for a logical connection between a specific project and its forecast impacts. But that is not the only reason why poor convergence has often been masked.

In particular, a variety of common mistakes have led to poorly converged models. Some of these problems, in other software, arose from use of a poor convergence measure, inaccurate calculations, and poorly conceived assignment options. For many years some older commercial packages offered equilibrium “closure” based upon VHT differences between successive iterations and did not have a valid gap measure. This caused modelers to think that their

assignments were converged when, in fact, they were far from it. Use of integer or limited precision link costs may also artificially inflate estimates of the convergence gap. Limiting the V/C ratio either explicitly or internally in the executable code has the same effect as does “dampening” of costs or flows.

Other practices are also problematic. Using a fixed number of iterations for an assignment is still common, but it is not a good practice. Often it results in arbitrary choice of the gap achieved, and it can easily lead to different levels of convergence for the base case and the forecast scenario thus introducing another source of error in any forecast.

Obviously execution time considerations have had a dominant impact. This is no longer necessary in most cases due to the various types of speedups available. Once again, testing the appropriate level of convergence is easy to do with TransCAD and can help you avoid convergence error in your model and its forecasts.

Most Likely Route Flows, Proportionality, and Select Link Analysis

The user equilibrium assignment computes a set of unique link flows if it is carried out properly and to a small enough relative gap. The route flows associated with any particular equilibrium assignment are not unique, despite the fact that estimates of these route flows are used when performing critical link analysis. Similarly, the class flows on each link derived from a multi-class assignment are not uniquely determined.

In a recent study, several university researchers examined the select link analysis results produced by various commercial packages with several different UE methods. (Boyce et al., 2010). This research, while not necessarily definitive due to limitations of several types, led to some interesting results. Perhaps the most important of these results was that unadjusted route-based and origin-based assignments lead to biased select link analysis, and conventional methods may also do so, but to a lesser extent. The reason is that iterative cost updates are made by origin for origin-based methods, and by origin or by origin-destination pair for route-based methods that enumerate routes. This leads to order dependence of the critical link assignment results because the specific new paths added are influenced by selection of prior ones.

Select link analysis based upon the FW or conjugate descent methods is more democratic in that no origin or route is given priority in terms of order of computation. Its defect is that in congested situations, the solutions reflect some unreasonable routes that are added in early iterations after the first set of shortest paths is computed and that are not dropped. This is mitigated, but only somewhat, by achieving good convergence. (With these methods, higher convergence also comes with a price in that there are more and more utilized paths with tiny fractional flows-- a result that is not realistic).

If there are multiple possible route flow solutions for a given set of equilibrium link flows, it becomes natural to try to identify the most likely route utilization. The maximum entropy or

most likely set of route flows provides an attractive solution to this problem. This methodology has not been commercially available until recently in **TransCAD 5, Release 3**.

In his Master's thesis, which still makes good reading today, Tom Rossi suggested the maximum entropy solution as a means of achieving a consistent method of allocating traffic to development projects (Rossi, 1987). Unfortunately his method of calculating the solution required route enumeration and use of a separate optimization package and might not be workable for the large networks used in current practice. Bar-Gera proposed the TAPAS assignment method (Bar-Gera, 2010) for achieving proportional route flows across paired alternative links in the solution for each origin, arguing that this very closely approached the maximum entropy solution for the route flows. This addresses a defect in origin-based and route-based traffic assignment with respect to select link analysis.

The idea of proportionality can be understood with a simple example, such as the one that follows. The network below shows the equilibrium flows resulting from a highly converged traffic assignment solution. This network has a demand of 1800 vehicles from origin node 1 to destination node 7 and of 1200 from origin node 3 to destination node 7. Such a case, in which different origins compete for limited infrastructure, is commonly found in real networks.

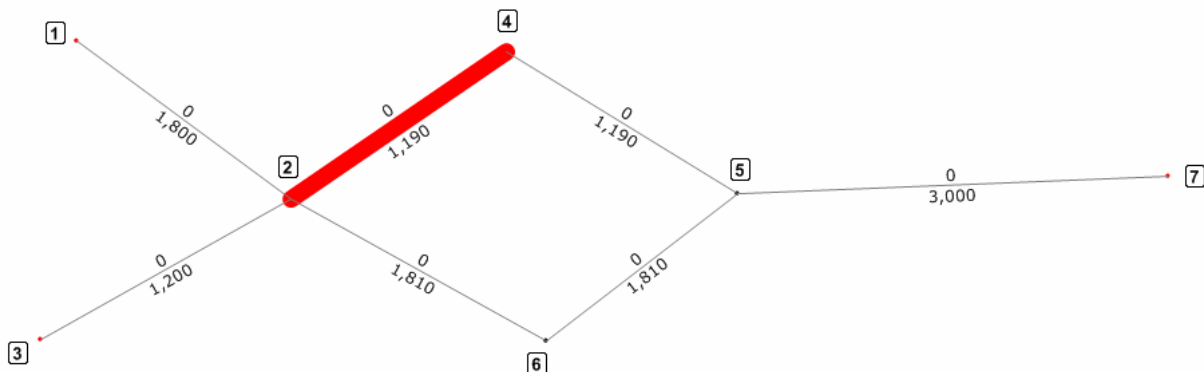


Figure 6: Proportionality Example

The answer to the question of how much flow is contributed to the total equilibrium flow on the highlighted link by each origin is that it should be in the same proportion as the flows on the two alternative routes, regardless of the point of origin. However, due to the order dependence of origin and path-based assignment methods, the select link results will be distorted. For OUE, we recognized that we could remedy this problem, and we immediately came up with a solution that achieves proportionality and does so with rather little additional computation. Table 4 below shows the select link analysis for highlighted red link with and without this correction.

Without proportionality, the flow on the critical link exclusively comes from the OD pair 1-7 and almost no flow comes from the OD pair 3-7. Note that this extreme solution still satisfies the link flows. If proportionality is used, the flows contributed by the OD pairs 1-7 and 3-7 are 714 and 476, which is proportional to the OD flows of 1800 and 1200 from node 1 and 3 respectively.

Origin	Destination	Flows without Proportionality Correction	Flows with Proportionality Correction
1	7	1189	714
3	7	1	476

Table 4: Select Link Flows With and Without Proportionality

In a forthcoming project report, Boyce et al. present findings demonstrating proportionality for select link analysis from TAPAS converged to 10^{-12} but compared the results with other methods converged only to 10^{-4} , presumably because not all methods in other software could go lower. The report suggests that extreme convergence is necessary to achieve route flow proportionality, but that is not so. Unlike Bar-Gera’s TAPAS method, OUE can achieve proportionality at the convergence levels that would normally be sought in large planning applications.

At present, TransCAD is the only commercial package that has this more robust select link analysis. Nevertheless, we should point out that the proportional route flow solution is a mathematical construct, so it is not a guarantee of behavioral realism.

A Note on Speed Comparisons

Research publications and sales literature are rife with claims from researchers and vendors stating that they have the fastest assignment method. We have made a practice of testing these claims to the extent possible and evaluating most of the new methods proposed, either from examination of published running times or through our own implementations. Based upon our investigations and published data, we believe that TransCAD is significantly faster in achieving convergence than other software. This statement applies to both our FW and to OUE in comparison to the new route-based methods from INRO and Citilabs and the LUCE algorithm in PTV software.

Even if performed with the same test problems, speed comparisons are fraught with difficulty due to variations in computer hardware, varying numerical precision of the software utilized, and use of differing convergence measures. Processor performance varies with the CPU architecture, on-chip memory caches and, of course, as a function of clock speed and the number of physical cores. As mentioned before, convergence behavior is strongly influenced by the precision of the calculations, with limited precision inflating convergence speed in some instances. There are many different metrics that are used to calculate the convergence gap in general and the relative gap in particular, so comparisons that use different measures are not

likely to be valid. We have also surmised that some results ignore the time it takes to actually test for convergence which, in of itself, takes significant computation time for large networks.

We have developed some comparative data using the well-traveled Chicago example popularized by David Boyce and Hillel Bar-Gera, and we have also had others make some runs for us with other commercial software. Based upon the evidence that we have gathered to date, the TransCAD OUE is significantly faster than the LUCE method in VISSUM and the projected gradient method in EMME/3. All of our assignments are many times faster than those in Cube Voyager. We plan on doing some further verification on different test problems before we publish these results. It remains to be seen how the performance of different algorithms and software implementations vary with problem size and other dimensions of difficulty. Also, there is always room for improvement, so don't be surprised if our assignment methods get faster in the future.

A Note on Multi-threading

You might wonder why the OUE and some other procedures in TransCAD are not yet fully multi-threaded. Multi-threading is a natural for computations that can be done in parallel, but brings another level of complexity to software development. If not done properly, multi-threading can actually slow down sequential calculations or result in a situation where the same computation yields different results on computers with differing numbers of cores. Slowdowns can occur when one execution thread has to wait for another to finish before proceeding. These so-called "race conditions" can also make the order of calculations somewhat dynamic. Changes in the order of calculation will often lead to small numerical differences in calculations, and, for iterative processes like UE traffic assignment, these small differences will be compounded into larger ones later in the calculation of results. Multi-threading can also greatly increase the amount of memory required leading to either insufficiencies or memory paging to disk, which has substantial performance penalties. So while multi-threading brings great benefits, it must be done with care.

We thought we should mention that a key component of our TransModeler traffic simulation software is multi-threaded, and that we recently did very well in a competition held by AMD to suggest the best use for their new 48 core offering. A You Tube video, for which we won third place, can be viewed at <http://blogs.amd.com/work/2010/04/15/winner-announced-what-would-you-do-with-48-cores/>. We are continuing to multi-thread more and more of both TransCAD and TransModeler, as large numbers of cores are clearly the wave of the future.

Deployment of OUE

OUE is deployed in several regional models for production use. Unlike some of the other competing new methods, it is feature complete and handles multiple user classes, entrance-to-exit tolls, and turn penalties, so there is no impediment to its applicability. However, for some very large problems with multiple user classes, OUE requires more than 2GB of memory. Thus a 64-bit solution is needed, and TransCAD 6, which will soon be in beta, will provide it.

References

- Bar-Gera, H. (2010) Traffic assignment by paired alternative segments, **Transportation Research B**, Article in press, doi.10.1016/j.trb.2009.11.004
- Boyce, D., Ralevic-Dekic, D., and Bar-Gera, H. (2004) Convergence of Traffic Assignments: How much is enough?, **Journal of Transportation Engineering**, ASCE, Jan./Feb. 2004
- Boyce, D. , Nie, Yu, Bar-Gera, H, Liu, Y, and Hu, Y. (2010) Field Test of a Method for Finding Consistent Route Flows and Multiple-Class Link Flows in Road Traffic Assignments, Prepared for the FHWA by the Transportation Center, Northwestern University, Evanston, Illinois. Publication forthcoming
- Caliper Corporation (2008) **Travel Demand Modeling with TransCAD, Version 5**, Newton, MA
- Daneva, M. and Lindberg, P.O. (2003) Improved Frank-Wolfe Directions through Conjugation with Applications to the Traffic Assignment Problem, Linkoping University, Dept. of Mathematics
- Dial, R. (1999) Algorithm B: Accurate Traffic Equilibrium (and How to Bobtail Frank-Wolfe), Volpe National Transportation Systems Center, Cambridge, MA July 25, 1999
- Dial, R. (2006) A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration, **Transportation Research B**, December 2006
- Florian, M., Constantin, I., and Florian, D. (2009) A new look at the projected gradient method for equilibrium assignment, TRB Annual Meeting CD-ROM, January 2009, Washington, D.C.
- Rose, G., Daskin, M., and Koppelman, F. (1988) An examination of convergence error in equilibrium traffic assignment models, **Transportation Research B**, Vol. 22B No. 4, 261-274
- Rossi, T. (1987) Traffic Allocation Methods for Use in Impact Fee Assessment, MS thesis, MIT Dept. of Civil Engineering, 1987
- Slavin, H., Brandon, J., and Rabinowicz, A. (2006) An Empirical Comparison of Alternative User Equilibrium Traffic Assignment Methods, **Proceedings of the European Transport Conference 2006**, Strasbourg, France
- Slavin, H., Brandon, J., Rabinowicz, A. and Sundaram, S. (2007) Achieving Planning Model Convergence, Presented at the 11th Transportation Planning Applications Conference, May Daytona Beach, FL, May 2007
- Slavin, H., Brandon, J., Rabinowicz, A., and Sundaram, S. (2009) Application of Accelerated User Equilibrium Traffic Assignments to Regional Planning Models, paper prepared for presentation at the 12th Transportation Planning Applications Conference, Houston, TX, May 2009

This page intentionally left blank.

B.2 Citilabs

Please answer the following questions about your travel demand forecasting software:

1. What is the latest version of your software and what assignment algorithms are available in it?
 - a. Cube 5.1.1 is the current officially released software version. We also provide updates to this release as Cube 5.1.2, which is available in pre-release form for customers under maintenance to download.
 - b. Static traffic assignment (convex combinations)
 - i. **Frank-Wolfe user equilibrium assignment** (COMBINE=EQUI), using either time-based (LAMBDASEARCH=EQUATION) or cost-based (LAMBDASEARCH=AREA) line search methods to solve for Lambda. Paths for up to 20 user classes may be built and loaded based upon any link array, such as TIME (using any user-specified TC, or congested time function), or COST (using any user-specified generalized cost function). Working matrix computations may be used to implement combined demand/supply models (such as diversion-based toll models). Gate-to-gate tolls may be incorporated into path costs by coding a toll matrix structure. Enhanced link-based Frank-Wolfe algorithms are also available:
 1. **Conjugate (ENHANCE=1) Frank-Wolfe user equilibrium assignment.** This option enables additional computations during the ADJUST phase to more accurately solve for LAMBDA, resulting in fewer path-building iterations required to reach higher-precision convergence.
 2. **Bi-Conjugate (ENHANCE=2) Frank-Wolfe user equilibrium assignment.** This option also enables additional computations during the ADJUST phase to more accurately solve for LAMBDA, resulting in fewer path-building iterations required to reach higher-precision convergence. Bi-Conjugate performs better than Conjugate Frank-Wolfe and is recommended by Citilabs as the best method for achieving high precision user equilibrium assignments without loss of proportionality in results.
 - ii. **Method of Successive Averages (MSA) or Volume Averaging equilibrium assignment (COMBINE=AVE).** COMBINE=WTD is used to perform a Method of Successive Weighted Averages or MSWA assignment. The Lambda series is fixed in either case. The Cube Voyager scripting language permits the user to build paths upon any working (LW.*) link arrays in addition to input (LI.*) and built-in (e.g.

TIME, DIST, COST) arrays. This permits the user to code different generalized cost functions for each user class, or implement any conceivable kind of stochastic assignment routine based upon Monte Carlo simulation of pseudo-randomly perturbed link costs. Built-in statistical distribution functions in the Cube Voyager scripting language include NORMDIST, GAMMADIST, POISSONDIST, LOGNORMDIST; others can be coded using mathematical computations in script. HCM-level analysis of intersection performance and delay is possible within assignment if a junction definition file is included as input.

- iii. Conveniently automated **stochastic user equilibrium (SUE) assignment** options are available as well for rectangular (COMBINE=BURRELL) and normal (COMBINE=PROBIT) link cost distributions. Both methods use Monte Carlo simulation of pseudo-randomly perturbed link costs; however all of the link array manipulations are automated for the user, requiring many fewer lines of script and reducing the potential for user error in implementation. Logit-based (i.e. so-called Dial or STOCH) assignment was implemented in TRANPLAN but is neither recommended nor implemented in Cube Voyager due to that algorithm's well-known tendency to violate Independence of Irrelevant Attributes (IIA) assumptions.
 - iv. **Incremental assignment (COMBINE=SUM)**, in which a fixed percentage of the total demand is loaded to the current minimum path set obtained in each iteration, and congested times and costs are updated after each iteration.
 - v. **Capacity-restrained assignment (COMBINE=ITER)**, in which each iteration is an all-or-nothing loading on the current minimum path set, and the results contain only the last iteration volume loadings.
- c. Static traffic assignment (non-link-based):
- i. **Path-based assignment** (gradient projection method, invoked via COMBINE=PATH). This algorithm is among the fastest user equilibrium assignment methods implemented in Cube Voyager. However, due to the zone order dependence inherent in this method as well as other non-link-based algorithms such as Origin-Based Assignment (OBA/OUE) and Algorithm B, intra-step distributed processing of path-based assignments using Cube Cluster is not possible at the current time. Furthermore, academic research has shown that non-link-based traffic assignment methods, including all current implementations of path-based assignment, as well as OBA, Algorithm B, and TAPAS, do not preserve proportionality in results, suggesting that although the same link flow solution is obtained by these algorithms, the path flow solution is not unique and may in fact be faulty. It is partly for this reason that select link/zone analysis and

turning movement as well as path file outputs are disabled in our implementation of path-based assignment. Cube Voyager users seeking a method that quickly converges to high levels of precision are encouraged to use the parameter setting COMBINE=EQUI, ENHANCE=2 (Bi-Conjugate Frank-Wolfe user equilibrium traffic assignment) instead of COMBINE=PATH until academic research proves that it is possible to implement non-link-based methods without loss of proportionality in results.

- d. **Cube Avenue** is an add-on to Cube Voyager that enables **Dynamic Traffic Assignment (DTA) with Mesoscopic Simulation**. The input networks to such models can be the same as those used in static equilibrium traffic assignment, however the input trip tables should be stratified by the user into multiple time segments within the assumed model period. These trips are disaggregated into real-valued “packets” of traffic, each of which is assigned to a randomly chosen continuous-time departure within the appropriate time segment. Packets are then simulated as they move through the network along time-dependent shortest paths and interact in queues formed at bottlenecks. This dynamic network loading process can be performed iteratively to construct a variety of DTA models defined by the user in script, including user equilibrium but also encompassing other, more general approaches.
 - i. **Dynamic User Equilibrium (DUE) assignment** is implemented using the Method of Successive Averages (COMBINE=AVE). By default in Cube 5.1.x, the MSA fractions are now interpreted probabilistically (PACKETS=PA, or “packet allocation”), meaning that at any given iteration a trip on a given origin-destination pair in a given time segment has a $1/N$ probability of switching to a new time-dependent minimum congested cost path (where N is the current iteration number). The resulting flows are virtually indistinguishable from the user equilibrium solution obtained using previous versions of Cube Avenue (PACKETS=PS, or “packet splitting”), wherein trips were assigned in a new set of packets with unique paths in each iteration, and all packet volumes were scaled by a factor of $1/N$ (where N is the iteration number). The main difference between the versions is a significant reduction in computational resources (mainly memory, but also processing power and time) required to perform a DUE assignment on a typical regional network.
 - ii. **Incremental DUE assignment** (ITERLOADINC>0) converges more quickly than conventional DUE by allowing the user to perform additional iterations of assignment for a given time segment before moving on to the next time segment in the period. This tends to isolate and fix problems in early time segments before they have the opportunity to spill over into and affect the entire model. As a result common problems such as network lockup and system-wide imbalance

of flows are generally averted and fewer iterations and hours of assignment are required to obtain a satisfactory DUE solution.

- iii. The user can build time-dependent shortest paths based upon any link input (LI.*), built-in (e.g. TIME, DIST, COST), or LinkWork (LW.*) array. The model period is divided into discrete time segments, and all path impedance link arrays are treated as two-dimensional, where the second dimension is the time segment index. Demand is loaded from input or working trip tables allocated by the user to these discrete time segments using prior or within-assignment matrix computations. Dynamic zone-to-zone path skims can be extracted for departures at any time point within the specified model period. These flexible functions allow the user to implement **stochastic DTA models** as well as **departure time choice or peak spreading models**. Tolls and other non-time costs can be incorporated into **generalized costs for multiple user classes** via techniques similar to those commonly employed within static Cube Voyager assignment; however, the tolls can either be input as time-varying charges or re-calculated dynamically within the assignment script. **Any facility characteristic, including speed, price, availability or capacity, can be set to vary dynamically** by time segment using script-based functions. Thus, Cube Avenue is an excellent tool for analyzing HOT lanes, managed lanes and similar dynamically priced road systems.
- iv. Many options are available to control the simulation of packets as they move through the network. The simulation is “mesoscopic”, meaning that a combination of macroscopic and microscopic techniques are used to evaluate congestion. Delays at capacity bottlenecks are evaluated microscopically, during simulation, by explicitly enforcing first-in-first-out (FIFO) queuing behavior at nodes. Cube Avenue can model truly **“horizontal” queues**, meaning that link storage in addition to flow capacity constraints are strictly enforced and queues spread backwards from link to link if storage constraints are exceeded. By contrast other simulation packages typically either model “vertical” queues (where storage constraints are not strictly enforced) and/or automatically remove demand from the network when queuing becomes excessive. Capacity flow constraints can be enforced at the link exit only in Cube Avenue by setting CAPLINKENTRY=N. If junction definitions are included as input, then queuing delays at intersections are based upon the application of a **junction model** (e.g. HCM 2000 equations). However, no cell transmission, gap acceptance or car-following models are included in Cube Avenue, meaning that lane-changing and weaving or merging behaviors are only implicitly considered through the evaluation of user-defined macroscopic speed-flow relationships (TC functions) based upon accumulated flows recorded during simulation. By using macroscopic relationships rather than explicitly simulating

traffic stream phenomena, Cube Avenue makes wide-area traffic simulation feasible with minimal alteration of input regional networks.

- v. Cube Avenue outputs a text-format (pseudo-XML) “packet log” file that contains the complete simulation results in a form that can easily be parsed or post-processed by Cube Voyager record processing scripts to identify exactly which packets (or trips) pass by a particular point within the network at a particular time. Furthermore, such processes can even extract the intended path and ultimate destination of trips that did not complete their journey before the end of the model period. Trips that meet such path-based criteria can then be isolated and assigned to a new path set using script-based controls. This functionality makes it possible to implement advanced **analyses of ITS strategies** such as variable message signs (VMS) and in-vehicle guidance or navigation systems, as well as a rich variety of user-controlled custom DTA algorithms.
- vi. Static assignment (i.e. PATHLOAD statements) can be included within Cube Avenue models. In such cases the statically loaded volumes are assumed to represent demand for the entire model period and interact with the dynamically loaded volumes using the total volume (V) and congested time (TC) functions. This technique may be used, for example, to perform **hybrid static/dynamic traffic assignment and simulation** within a detailed sub-area of a regional model, forgoing most simulation in the outskirts of the network while still considering the time required to traverse the network from a trip’s ultimate origin to the sub-area boundary.
- vii. Other assignment options available within Cube Avenue include fixed-Lambda COMBINE settings such as WTD (**weighted averaging**), SUM (**incremental**), and ITER (**capacity restrained**) assignment. EQUI is not currently a valid setting for Cube Avenue.

2. Do the algorithms converge to a relative gap of 10^{-5} ?

- a. Yes. The recommended parameter setting for high-precision convergence is COMBINE=EQUI, ENHANCE=2 (Bi-Conjugate Frank-Wolfe assignment). Path-based assignment (or any similar non-link-based quick-converge method) is not currently recommended for project evaluation purposes until academic research proves that it can be implemented in software without loss of proportionality in results.

3. What, if any, are the hardware requirements to run these algorithms?

- a. Cube Voyager will run on any modern Windows-based machine, including 64-bit as well as 32-bit XP, Vista, and 7 operating system versions.

- b. Cube Cluster is an add-on to Cube Voyager that enables intra-step and multi-step distributed processing of highway assignment on multiple CPU cores. Unlike multi-threaded applications, which rely upon the Windows operating system to balance resource load and can only access computing resources on the local machine, Cube Cluster allows the user to completely control the allocation of processing workload to available processors on a local area network. This actually reduces the cost of hardware in most cases because several quad or 8-core machines can typically be purchased for less than a single processing server having the same combined number of CPU's and amount of RAM available.
 - c. Cube Voyager is priced according to the number of processors available on the local machine where it is installed; the current "standard" license supports up to 4 cores on each machine, while upgrades are available to support machines having up to 8, 16, or 32 processors. The Cube Cluster add-on is included free of charge with "Cube Pro" upgrades from 4 to 8 processors.
4. When is your next release expected? Will it include any additions or enhancements to the assignment methodology?
- a. Cube Voyager 5.2 is targeted for release at the Futura International Citilabs User Conference in Fall 2010 (October-November) and will include the following enhancements:
 - i. a "warm-start" mode for static equilibrium traffic assignment (i.e. using paths generated by a previous run),
 - ii. performance improvements to path-based assignment, and
 - iii. some of the prioritized improvements for Cube Avenue, discussed further in the following section.
5. Are any additional enhancements or additions to the assignment methodologies currently planned for research and/or development?
- a. Static equilibrium assignment R&D
 - i. Route choice models have been identified as a desirable and fruitful area of research and development. A path-size logit route choice algorithm has been prototyped in Cube Voyager.
 - ii. We are continuing to improve our implementation of the path-based assignment algorithm and will continue to collaborate with the academic research community in working to identify methods of resolving the zone order dependence and proportionality issues currently limiting the application of non-link-based assignment models.

- iii. The ability to run assignment from a so-called “warm start” based upon saved paths has also been identified as a desirable improvement, for example permitting the use of path-based assignment during interim full model iterations, followed by a final detailed enhanced link based assignment with warm start from the path-based outputs.
 - iv. While not expected to significantly improve run times, we are researching other non-link-based approaches, such as OBA and TAPAS. These may be implemented to satisfy customer demand for testing and evaluating well-known alternative algorithms.
 - v. We are continually collecting feedback from our users and evaluating the potential for improvements to basic functionality such as multi user class generalized cost assignment.
- b. Dynamic traffic assignment R&D
- i. We have recognized the need to calculate convergence statistics in Cube Avenue differently than in static assignment; for example a path-based (rather than link-based) relative gap calculation has been prototyped in Voyager script and will probably be natively implemented within the software in the near term.
 - ii. Non-MSA equilibrium DTA algorithms are also a short-term goal. In this endeavor we hope to import to Cube Avenue some of the concepts and insights gained from our research into enhanced static equilibrium traffic assignment methods.
 - iii. The ability to programmatically manipulate saved paths and skims from multiple iterations may make it possible for users to implement a wide variety of custom DTA algorithms in script, possibly addressing the two points above.
 - iv. Currently, congested time (TC) functions are used as the primary means of representing macroscopic traffic stream flow phenomena in Cube Avenue. We would like to augment or replace these with speed-density (SD) functions evaluated continuously during simulation in the near future.
6. Are there plans for developing mesoscopic modeling capabilities?
- a. These capabilities currently exist; Cube Avenue, an add-on to Cube Voyager, performs dynamic traffic assignment with meso-scopic simulation. See responses to the previous questions for further details. In addition to its obvious applicability to corridor and sub-area studies as well as extensions to regional assignment, this tool is being used for evacuation planning purposes by the following organizations:

- i. The Virginia Modeling, Analysis, and Simulation Center
 - ii. The State of Florida (several separate districts funded a joint study)
 - iii. Houston-Galveston Area Council
 - iv. Santa Clara County, California
7. Is there other information surrounding this topic and your software that you believe we should be aware of?
 - a. On March 11, 2010 we conducted a webinar entitled “New Assignment Methods in Cube Voyager” which discusses the enhanced static equilibrium traffic assignment methods in Cube Voyager 5.1. The presentation slides as well as a recording of this webinar are archived and can be accessed at the following links:
 - i. <https://www1.gotomeeting.com/register/363881992>
 - ii. <http://www.citilabs.com/New%20Assignments%20in%205.1%20Webinar.pdf>
 - b. MWCOG and its consultants are also strongly encouraged to review the presentation generated by the academic research team of Hillel Bar-Gera, David Boyce, and Yu Nie entitled “Practical Implications of Finding Consistent Route Flows”, presented at the 2010 TRB Planning Applications Conference. Citilabs cooperated with these researchers, who conducted a broad survey and comparison of assignment approaches adopted by a variety of vendors. Their slides are available for download at the following location:
 - i. http://www.trb-appcon.org/TRB2009presentations/s14/01_bar_gera_nie_and_boyce_finding_consistent_route_flows.ppt
 - ii. As noted in the webinar, whereas other vendors have pursued quick-converge methods that sacrifice proportionality and route consistency in order to show fast convergence as measured by link-based statistics, Citilabs’ enhanced link-based assignment methods achieve the recommended levels of convergence (relative gap of 10^{-4} or less) within similar run times (or better, if Cluster is used), without sacrificing proportionality or consistency of route flows.
 - iii. It should furthermore be noted that the issue of proportionality and consistency of route flows is not a trivial matter, but in fact one of potentially serious policy implications. Without consistent route flows it is impossible to make credible statements about the likely origins and destinations of travelers using a particular node, link, or project. This undermines the validity of a fundamental component of project

environmental justice analysis, that is a comparison of the demographic composition of the project's users or beneficiaries with that of the populations immediately affected by said project's negative consequences or impacts. The need to perform this type of analysis when evaluating controversial transportation projects was underscored by the FTA's recent decision to withdraw funds from BART's proposed Oakland Airport Connector project, noting that "the fact that BART has not conducted the necessary service equity analysis for the OAC project or fare equity analysis raises concerns that [the] agency does not have procedures in place to monitor its subrecipients." (Quote from FTA reported in <http://www.planetizen.com/node/42991>.) The legal basis for this decision is not limited to transit project funding requirements, but includes Title VI and US DOT Environmental Justice regulations affecting highway projects as well.

Thank you very much for your participation.

You are most welcome! – *Citilabs*.

This page intentionally left blank.

Advanced Assignment Algorithms

B.3 Vendor Survey - INRO

Please answer the following questions about your travel demand forecasting software:

1. What is the latest version of your software and what assignment algorithms are available in it?

Emme 3.3.0 was released in March 2010. The following assignment algorithms are available:

a) Emme Traffic Assignment:

- **Standard Traffic Assignment.** User-optimal Frank and Wolfe equilibrium traffic assignment using linear approximation.
- **Parallel Standard Traffic Assignment.** Parallel computing version of the Standard Traffic Assignment, offering speedups on multiprocessor systems.
- **Path-based Traffic Assignment.** Fast-converging new user-optimal equilibrium traffic assignment offering dramatic performance improvements, better convergence, and faster analysis of path-based results, as well as warm starts for faster iterations in feedback.

All Emme traffic assignments offer single or multi-class functionality with generalized costs. The volume/delay (cost) functions are specified as open, user-configurable algebraic expressions. Whether it is a BPR, conic, or any other polynomial function, with or without an intersection delay term, the function is specified as an algebraic expression, interpreted at execution time without the need to compile special code. Emme traffic assignments can be adapted for a variety of assignment model variants that include system-optimal assignments, stochastic user-equilibrium assignments, and more sophisticated asymmetric cost assignment for turn-delay applications, or multiclass assignments where per-class link costs depend on factors such as average speed or vehicle class mix (for example, cars and trucks).

b) Emme Transit Assignment:

- **Standard Transit Assignment.** Sophisticated multimodal, multipath transit assignment that minimizes the expected travel cost.
- **Timetable Assignment.** Complete timetable assignment which uses detailed departure and arrival information to derive an optimal path.

-
- **Disaggregate Transit Assignment.** Permits the detailed analysis of individual transit trips, where the origins and destinations are specified as precise coordinates or node numbers rather than zone numbers.

The Emme Transit Assignment offers a sophisticated multipath assignment; a multipath or 'strategy' is a generalization of the concept of a path. In Emme, travellers may choose from elements that are more complex than a simple path toward a destination; for instance they may choose a set of paths, and let the vehicle that arrives first at a stop determine which of these paths to take.

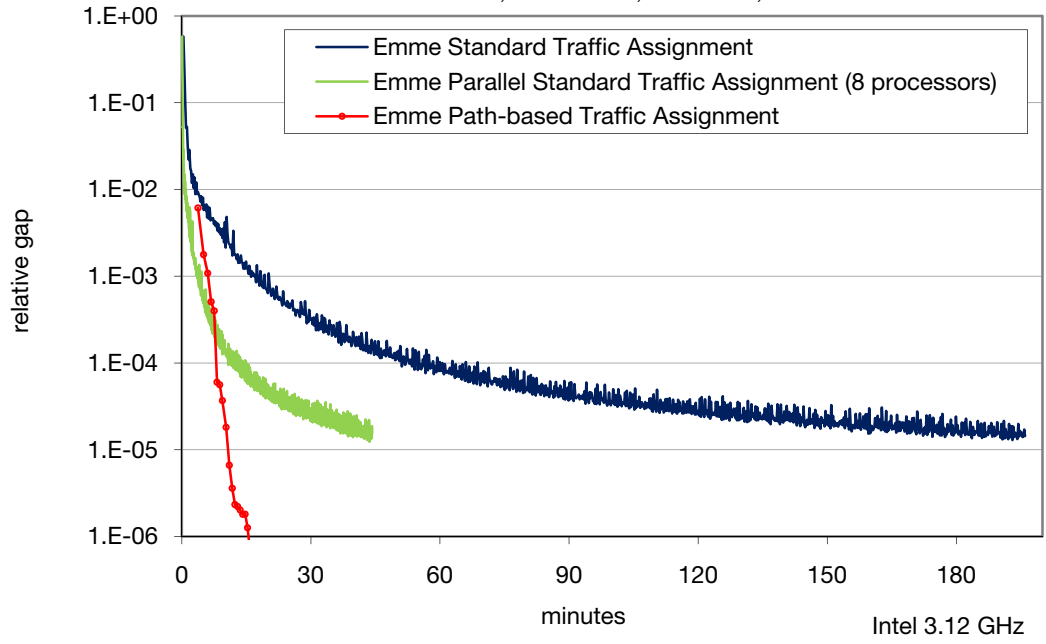
The Emme Transit Assignment supports multiclass assignments on multimodal networks, and transit time functions are open and fully user-specified. The Emme Transit Assignment also provides sensitivity to transit-segment specific attributes, allowing the possibility to implement sophisticated iterative solutions to effectively address congestion in transit networks, whereby crowding on certain transit segments can decrease the comfort of passengers in the vehicle, whereby passengers are prevented from boarding transit vehicles where there is no more room, or where transit vehicles on long transit-lines may "bunch" on congested corridors.

2. Do the algorithms converge to a relative gap of 10^{-5} ?

The Standard Traffic Assignment and the Parallel Standard Traffic Assignment algorithms converge to a relative gap of less than 10^{-4} and sometimes to just above 10^{-5} , depending on the given demand and network, before the convergence rate diminishes. The Path-based Traffic Assignment converges to a relative gap 10^{-5} and smaller, with reasonable computing times. See the convergence chart below.

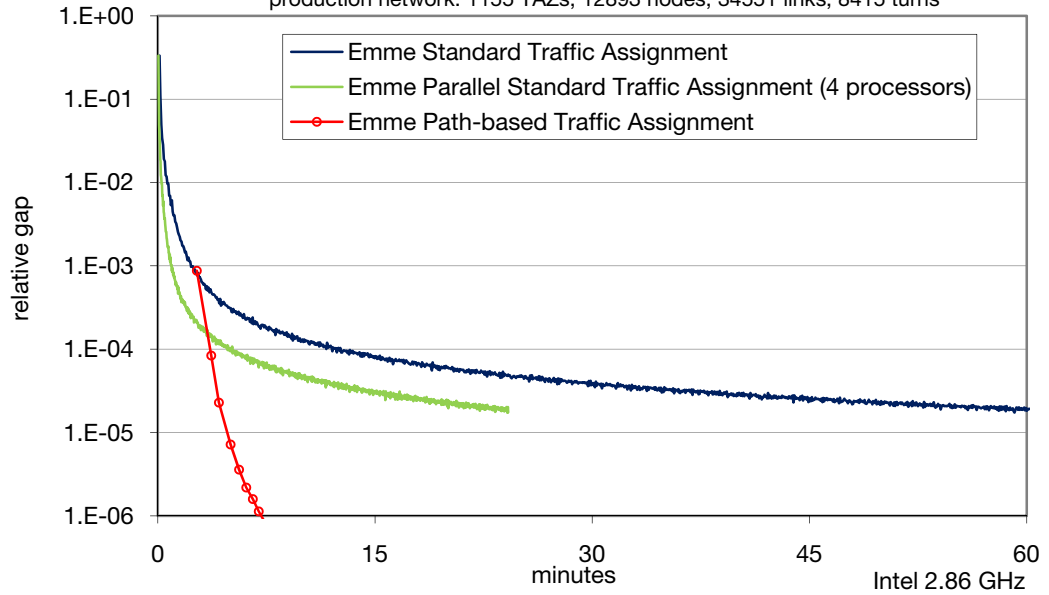
Maricopa Association of Governments

test network: 2041 TAZs, 12938 nodes, 39731 links, 1896 turns



Roads and Traffic Authority of New South Wales (Sydney)

production network: 1155 TAZs, 12893 nodes, 34551 links, 8415 turns



3. What, if any, are the hardware requirements to run these algorithms?

The Emme Standard Traffic Assignment performance will benefit from faster processor speed and I/O speed, though it can run on legacy hardware. Under 100 MB of physical memory are required, regardless of the problem size.

The Emme Parallel Standard Traffic Assignment requires a multiprocessor system and will consume physical memory depending on the network size, number of processors used, and analysis options specified. Larger networks with heavy multiprocessor use will typically consume a few hundred MB.

The Emme Path-based Traffic Assignment is more memory-intensive and will consume RAM depending on the network and demand characteristics of the model. The path-based traffic assignment is fastest when the paths generated can be kept in physical memory during the assignment process, but this is not a strict requirement and the procedure will respect the memory constraints available. Large applications will benefit from as much physical memory (RAM) as possible.

4. When is your next release expected? Will it include any additions or enhancements to the assignment methodology?

INRO is continuously working to enhance the efficiency and performance of the assignment algorithms. Upcoming releases will contain enhancements to both traffic assignment and transit assignment. Minor releases, i.e. Emme 3.x, occur at least once a year and patches, i.e. Emme 3.3.y, are more frequent. Emme 3.2 and 3.3 were released in September 2009 and March 2010, respectively. Two additional patches were released between these minor releases.

Emme users with a Software Support Agreement are eligible to receive Beta versions for advance access to new features and an opportunity to provide feedback to INRO.

5. Are any additional enhancements or additions to the assignment methodologies currently planned for research and/or development?

The research and development of new assignment algorithms is a continuous activity at INRO. At any one time we have several potential additions and enhancements which are the subject of investigation.

INRO's Beta program provides pre-release versions of assignment methods to organizations that have expressed a desire to work closely with INRO Support to deploy and test new software features.

6. Are there plans for developing mesoscopic modeling capabilities?

INRO offers Dynameq for Dynamic Traffic Assignment. With dozens of applications in North America, Europe, Asia and Australia, Dynameq is the leading professional software for mesoscopic modeling. A free Dynameq Trial

version is available from www.inro.ca/dynameq. References are available upon request.

7. Is there other information surrounding this topic and your software that you believe we should be aware of?

Any new assignment method that is released undergoes extensive testing, in house with data sets from clients and in the field with the collaboration of Emme users, to ensure that assignment results and performance are ready for practice. The assignment algorithm and path analysis capabilities (select link, generalized select link, cut off operator, etc.) are supported. The method is numerically stable; for instance, the Emme Standard Traffic Assignment and the Parallel Standard Traffic Assignment produce identical results using 1, 2, 4, 8, or more processors, which is non-trivial due to asynchronous flow updates that may affect numerical precision. The software enables the conscientious use of system resources: the user can choose to allocate a fraction of the available processors, reserving the remaining processors for additional model runs or other computing needs; similarly, the user can choose to cap the RAM consumed by an Emme Path-based Traffic Assignment.

Well-designed software features and development practices demonstrate INRO's dedication to delivering reliable software to professional modellers.

This page intentionally left blank.

B.4 PTV

Please answer the following questions about your travel demand forecasting software:

1. What is the latest version of your software and what assignment algorithms are available in it?

The latest version of our PTV Vision software is VISUM 11.0 and VISSIM 5.20.

VISUM offers a range of highway assignment methods, both static and dynamic.

The recommended static highway assignment method in VISUM is LUCE, an origin-based gradient method, which in recent comparisons achieved the highest convergence speed among all tested algorithms. As an example, LUCE solves the assignment for the 1800-zone Chicago Regional model, which serves as a common industry benchmark in about 7 minutes to a relative gap of 10^{-5} on a 2.66 GHz notebook. Convergence continues without slowing down significantly (10^{-6} after 9.5 min, 10^{-7} after 13min, 10^{-8} after 19 min). LUCE was devised by Prof. Guido Gentile (University La Sapienza Rome, Italy) as a further improvement of OBA (by Bar-Gera) and achieves much better convergence speed through the use of a better descent direction.

Largely for backward compatibility, VISUM also offers a path-based equilibrium assignment algorithm and incremental assignment.

For applications where relatively modest relative gaps must be reached very quickly, VISUM contains a variant of Frank & Wolfe (called Equilibrium Lohse after the original researcher). It shares the properties of F&W observed generally, i.e. it can be multithreaded easily, loads a rich path set, and achieves good convergence speed down to $\sim 10^{-3}$, but tails off after that.

Special-purpose variants of the methods above exist for assignment with road pricing (featuring continuously-distributed value of time), with detailed intersection delay calculation (instead of turn volume-delay functions, based on HCM 2000), and with blocking-back / downstream metering.

In addition to the various flavors of equilibrium assignment VISUM offers a stochastic user equilibrium assignment method.

Both stochastic and equilibrium highway assignments are also available in dynamic versions, the latter being based on the simplified theory of kinematic waves.

It is, and always has been a distinctive feature of VISUM that all path information is stored at the end of an assignment, so that all post-assignment functions (e.g. select-link, matrix estimation, skimming etc.) can be executed without having to re-run the assignment.

In addition to the built-in assignment methods listed above, VISUM provides a complete framework that supports interfaces with several external dynamic assignment methods. Within this framework, VISUM manages both the supply and demand natively and transfers input data necessary for assignment. Post assignment results are transferred back to VISUM where all of the standard analysis tools are available. Currently there are interfaces are to VISTA (www.vistatransport.com) and DynusT (www.dynust.net)

2. Do the algorithms converge to a relative gap of 10^{-5} ?

Yes, see above.

3. What, if any, are the hardware requirements to run these algorithms?

Standard off-the-shelf hardware is sufficient for most models. For large-scale models we recommend 64-bit Windows and RAM in the 8-16 GB range. Many of the computationally intensive calculations are multi-threaded, so having a multi-core machine will cut run times.

4. When is your next release expected? Will it include any additions or enhancements to the assignment methodology?

The next release will be 11.5 scheduled for summer 2010. The specialized assignment variants for toll assignment and for assignment with detailed intersection delays will be updated to improve run time and to support additional types of road pricing schemes.

5. Are any additional enhancements or additions to the assignment methodologies currently planned for research and/or development?

LUCE is derived from OBA and shares the property that turning proportions are consistent across all paths originating at one zone, but not across all origin zones. We plan to refine the method, so that the result comes closer to overall proportionality.

6. Are there plans for developing mesoscopic modeling capabilities?

Yes. For a while now we have engaged in discussions with PTV Vision users about additional features / requirements that would make dynamic traffic assignment (DTA) more practically useful. The majority of our users have expressed that DTA in the urban context needs to take into account intersection detail (geometry, control) in the impedance. Currently the preferred DTA methods in VISUM are based on theories that cannot easily incorporate junction detail (e.g. Dynamic User Equilibrium based on the simplified theory of kinematic waves).

This user input is confirmed by similar industry momentum and for that reason we believe that there is sufficient justification for adding a mesoscopic simulation to VISUM that incorporates junction detail in the near future. It is also our intent that the same assignment method could be run in hybrid mode with our VISSIM microscopic

simulations. This capability would speed up micro-simulation in larger networks where VISSIM's level of detail is not needed everywhere in the network.

In order to offer this mesoscopic simulation method/functionality we have been pursuing another academic partnership similar to that achieved with our origin-based LUCE assignment method. These discussions are well underway and currently the most promising candidate for integration into VISUM is Mezzo, developed by Wilco Burghout and Ingmar Andreasson at KTH Stockholm (www.ctr.kth.se/mezzo). Both researchers are keen to develop Mezzo further, but seek an arrangement where it would become a fully supported piece of software capable of supporting day to day project work. Based on other related academic work by Wilco Burghout will believe that Mezzo is also a good choice for hybrid simulations with VISSIM. A prototype for VISUM could be available later this year, with a full integration likely in VISUM 12, tentatively scheduled for mid-2011.

PTV has a long commitment to multi-resolution modeling. With this knowledge, experience and strong academic collaboration we are proud and confident in our ability to rapidly support users in the evolution of the mesoscopic space. This is perhaps best evidenced by the progress of our development efforts in the past 12-18 months offering users the industry's fastest OBA method and two-way interfaces with alternative DTA methods.

7. Is there other information surrounding this topic and your software that you believe we should be aware of?

As mentioned, VISUM is part of PTV's Vision suite and is integrated with the industry leading microsimulation sister product VISSIM. Paths from any VISUM assignment can be readily exported to VISSIM for microsimulation. Alternatively, a VISUM assignment can serve as a starting solution for dynamic assignment within VISSIM, speeding up the calculation.

VISUM is a fully multi-modal demand modeling package that contains an extensive transit data model with the industry's most advanced frequency-based and schedule-based transit assignments. We assumed the above questions were concerned only with highway assignment. If transit modeling capabilities are taken into account in the software evaluation process at hand we can provide more detail on these assignment methods.

new algorithm for user equilibrium traffic assignment:

LUCE in VISUM

For decades the state of the art in equilibrium assignment was dominated by methods like Frank & Wolfe (in most software packages) and by the path-based assignment in VISUM. But some years ago Hillel Bar-Gera's work on Origin-Based Assignment started a new era of research into alternative methods. Now it is time also for a quantum leap in VISUM: PTV introduces a completely new algorithm, called Linear User Cost Equilibrium (LUCE). LUCE was conceived by Professor Guido Gentile who during 2008 collaborated closely with PTV to produce a practical implementation of the method in VISUM.

LUCE represents a step change from the earlier path-based equilibrium assignment by improving runtime, convergence and path proportionality.

- runs much faster
- reaches higher levels of convergence
- loads a richer set of paths per OD pair

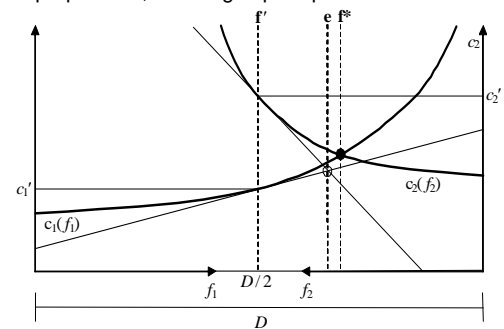
How does it work?

LUCE is an origin-based gradient method. Like all origin-based methods LUCE equilibrates volumes simultaneously on all paths from one origin to all destinations. Unlike VISUM's classic assignment method, those paths are not explicitly represented. Instead LUCE works on data structures called bushes which efficiently store all efficient paths for one origin. The way bushes encode path volumes ensures that at least all of the paths for one origin satisfy the fair proportionality condition recently proposed by Bar-Gera and Boyce. Like all gradient methods, LUCE does not directly equilibrate travel cost along alternative routes, but instead uses a linear approximation of travel cost around the current solution, by taking the gradient of the cost function with respect to volume.

Exploiting the inexpensive information provided by the derivatives, LUCE achieves a very high convergence speed, while it assigns the demand flow of each O-D pair on several paths at once. LUCE partitions the problem by origins. For a given origin the main idea is to seek at each node a user equilibrium for the local route choice of drivers coming from the origin among the arcs of its backward star. The travel alternatives that make up the local choice sets are the arcs that belong to the current bush – a bush is an acyclic sub-graph that connects each destination to the origin at hand. The cost functions associated to these alternatives express the average impedance from the origin to each intermediate node, linearized at the current flow pattern.

The unique solutions to such local linear

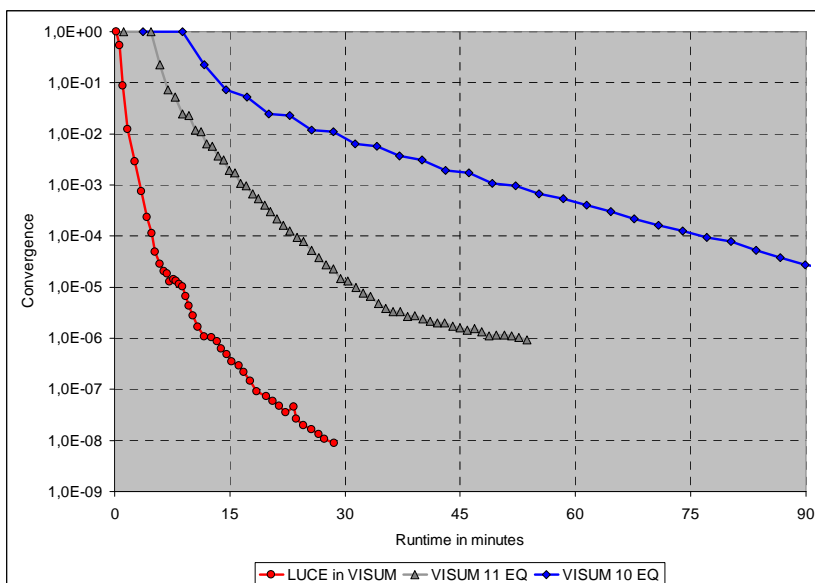
equilibria in terms of origin flows, recursively applied to each node of the bush in topological order, provide a descent direction with respect to the classical sum-integral objective function. The network loading is then performed through turn proportions, avoiding explicit path enumeration.



Schematic graph of linearisation of user cost function

How does it perform?

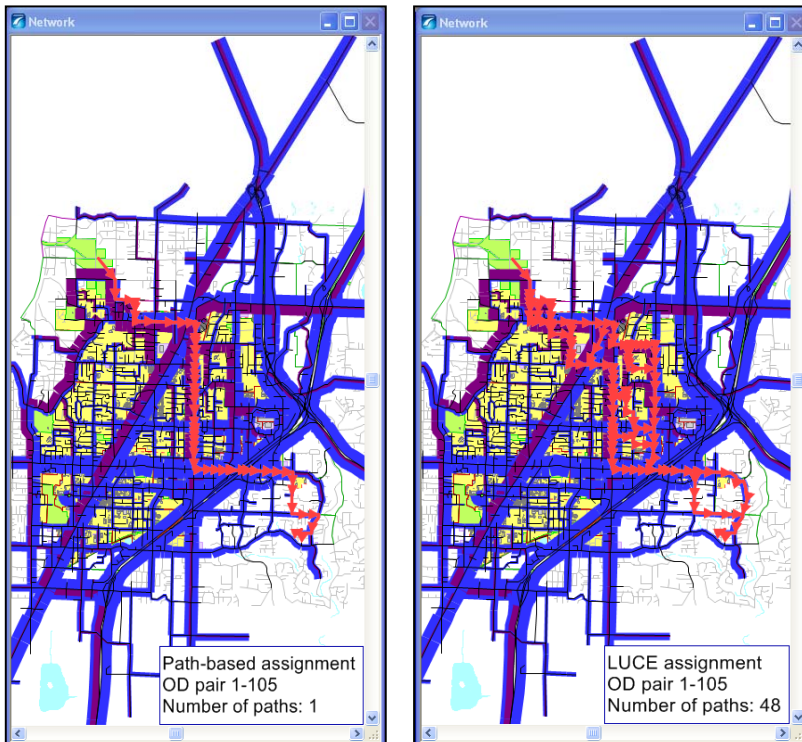
The effects on run-time of the improved algorithm are striking. The chart (left) compares run-times for the Chicago regional model used in many benchmarking exercises. This model contains ~1800 zones. The horizontal axis is run-time, the vertical axis is relative gap. Compared to equilibrium assignment in VISUM 10, the same path-based algorithm performs already ~3 times faster in VISUM 11, but LUCE is another improvement by a factor 3-5. To reach a gap of close to 10⁻⁵, VISUM 10 took over 1.5 hours. In VISUM 11 the path-based assignment takes only ~30 minutes to reach the same gap, but LUCE reaches it in ~8 minutes. Convergence continues down to around 10⁻⁸ at much the same speed. The same qualitative results were obtained for many other networks.



Comparison of performance of LUCE with existing equilibrium assignment methods

LUCE also loads a richer set of OD pairs. Note that while link volumes are uniquely determined in equilibrium, path volumes are not. In fact, a vast set of path volumes will fit a given link volume pattern. VISUM's path-based assignment is known to be parsimonious, i.e. it covers the equilibrium link volumes with a very small number of paths, often just 1 path per OD pair.

As long as path volumes are only a by-product on the way to good link volumes this can be neglected, but if the path volumes are analyzed in their own right, proportionality becomes important. The picture below illustrates how on a congested network VISUM's path-based assignment will still only load a single path for a given OD pair, whereas LUCE will load no less than 48 – all plausible.



Comparison of path proportionality

Current Status

At the time of writing (May 2009), the core method is stable enough to share it with our users, although some auxiliary functions are still missing and the post-assignment analysis methods still need to be optimized for LUCE. At this stage PTV provides LUCE as a prototype in VISUM 11, intended mainly for evaluation purposes. It does run on realistic networks, but it currently has some technical limitations: most importantly post-assignment analysis (skimming, path listing, matrix estimation, ...) is not possible yet.

Currently all these operations post-process paths, and are therefore not available in LUCE, which – through the improved proportionality – loads too many paths to be stored in memory. While this severely limits the use of LUCE in real projects, you can still compare link volumes, gaps and runtimes, to estimate how your model run times will decrease in the future. We strive to lift most of these limitations already in VISUM 11 bugfixes, i.e. well before the next major release.

REFERENCES

Gentile G. (2009), Linear User Cost Equilibrium: a new algorithm for traffic assignment, submitted to Transportation Research B.

Perspectives:

Some of the planned extensions are:

- ▶ **Warm start:** Like other assignment methods in VISUM, LUCE will be able to accept an existing assignment result as an initial solution. Because the warm start functionality requires a set of bushes, the prior assignment result must be of type LUCE
- ▶ **faster skim matrices, select-link analysis (flow bundle), etc.:** In the prototype, none of the post-assignment analysis functions are available for LUCE because of the memory footprint of explicit path storage (see above). Therefore, all of the analysis methods need to be re-implemented to work directly with the implicit bush representation, which will not only save memory, but also speed up the operation.
- ▶ **Bush storage:** LUCE will have an option to turn off bush storage and save memory, because bushes only need to be saved, if you plan to warm-start or use post-assignment analysis
- ▶ **General tuning:** Although we have tested LUCE on diverse networks we fully expect performance (in terms of memory and runtime) to vary with the characteristics of the networks, and it is quite likely that we will need to tune the implementation for the cases that have escaped us so far. You can help us by reporting to the VISUM hotline instances in which LUCE consumes unklky amounts of runtime or memory

With the next major release we expect LUCE to become the default equilibrium assignment method in VISUM .

Contact

PTV AG
Geschäftsfeld Traffic
Stumpfstraße 1,
76131 Karlsruhe
Germany

Tel.: +49 721 9651 - 300
Fax: +49 721 9651 - 562
E-Mail: info.vision@ptv.de
Internet: www.ptvag.com

This page intentionally left blank.

Short-Term Model Enhancements: Trip Purposes and Special Generators

■ 1.0 Introduction

Cambridge Systematics (CS) was tasked by staff of the National Capital Region Transportation Planning Board (TPB) to investigate two sets of issues on short-term enhancements of the regional travel demand model:

- Consideration of expanding the existing trip purposes, including potentially splitting the non-home-based trip purpose and/or splitting out the home-based school/university trip from the home-based other trip purpose.
- Consideration of developing additional or enhanced special trip generator models to handle special cases, including airport travel, special events, visitors, universities, regional shopping centers, group quarters, or military bases.

This report documents CS' reviews, findings, and recommendations on these issues. In performing this task, CS consulted documentation for adopted regional travel demand models for several of the largest metropolitan planning organizations, including those listed in Table 1. Based on these reviews and CS' experience, the pros and cons of splitting trip purposes are analyzed and a recommendation for splitting trip purposes and enhancing special generator models were made. Section 2 documents the work on trip purposes and Section 3 documents the work on special generators.

Table 1. Model Documentation Consulted

Agency Name	Metropolitan Region
Atlanta Regional Commission (ARC)	Atlanta, Georgia
Baltimore Metropolitan Council (BMC)	Baltimore, Maryland
Chicago Metropolitan Agency for Planning (CMAP)	Chicago, Illinois
Central Transportation Planning Staff (CTPS)	Boston, Massachusetts
Denver Regional Council of Governments (DRCOG)	Denver, Colorado
Delaware Valley Regional Planning Commission (DVRPC)	Philadelphia, Pennsylvania
East-West Gateway Council of Governments (EWGCC)	St. Louis, Missouri
Houston-Galveston Area Council (HGAC)	Houston, Texas

Table 1. Model Documentation Consulted (continued)

Agency Name	Metropolitan Region
Hampton Roads Planning District Commission (HRPDC)	Hampton Roads, Virginia
Metropolitan Council	Minneapolis-St. Paul, Minnesota
MetroPlan Orlando	Orlando, Florida
Metropolitan Transportation Commission (MTC)	San Francisco, California
North Central Texas Council of Governments (NCTCOG)	Dallas, Texas
North Jersey Transportation Planning Authority (NJTPA)	New Jersey
Northeast Ohio Areawide Coordinating Agency (NOACA)	Cleveland, Ohio
Ohio-Kentucky-Indiana Regional Council of Governments (OKI)/Miami Valley Regional Planning Commission (MVRPC)	Cincinnati; Dayton, Ohio
Portland METRO	Portland, Oregon
Puget Sound Regional Council (PSRPC)	Seattle, Washington
Regional Planning Commission	New Orleans, Louisiana
Sacramento Area Council of Governments (SACOG)	Sacramento, California
San Diego Association of Governments (SANDAG)	San Diego, California
Southern California Association of Governments (SCAG)	Los Angeles, California
Southeast Michigan Council of Governments (SEMCOG)	Detroit, Michigan
Southeastern Wisconsin Regional Planning Commission (SWRPC)	Milwaukee, Wisconsin
Southwestern Pennsylvania Commission (SPC)	Pittsburgh, Pennsylvania
Southeast Regional Planning Model (SERPM)	Ft. Lauderdale, Florida
Tampa Bay Regional Planning Model (TBRPM)	Tampa, Florida

■ 2.0 Trip Purposes

The TPB model currently has four trip purposes: home-based work (HBW), home-based shop (HBS), home-based other (HBO), and non-home-based (NHB). Based on a survey conducted for TRB Special Report 288, many regional models had seven or more person trip purposes (VHB, 2007), and less than one-third of large MPOs used four or fewer trip purposes. One of the key issues facing TPB staff, as part of developing the Version 2.3 travel model, is whether to split the NHB trip purpose into two trip purposes: non-home-based work-related (NHB WR) and non-home-based other (NHBO); alternatively, whether to split out school/university trips from HBO; or do both. Additional issues facing TPB staff include attraction estimation methods and the downstream effects of trip purpose splitting on the remaining model components. In the following subsections, we

discuss non-home-based trips, home-based college/university trips, and home-based school trips separately, in terms of our review findings, pros and cons, and recommendations.

2.1 Non-Home-Based Trips

Non-home-based trips are trips with neither end at home. The following summarizes the findings of our review:

- Nearly two-thirds of the regional travel demand models reviewed by CS have separate non-home-based-work-related (NHB WR) and non-home-based other (NHBO) trip purposes.
- Based on the survey conducted for TRB Special Report 288, the NHB WR trip purpose was used in 12 regional travel demand models out of a total of 36 large MPOs surveyed (VHB, 2007).
- Household travel surveys are commonly used to estimate trip production and attraction models for NHB WR and NHBO purposes.

While it is still considered state of the practice to not split the NHB trips in traditional four-step models, it is common for the large MPOs to split NHB trips into NHB WR and NHBO. However, there are both pros and cons for performing this split which should be considered.

Pros:

- NHB WR trips have very distinct travel behaviors that are very different from NHBO trips, including travel patterns, mode choice, etc. Modeling NHB WR separately allows consideration of those factors that affect work-related trips to be modeled explicitly.
- NHB WR trips have different values of time (VOT) than NHBO trips. Modeling the two separately will allow more accurate representation of VOT in the model, and help model and analyze pricing policies in the region.
- NHB WR and NHBO trips also have different time-of-day patterns, and thus splitting NHB trips into the two types will help with time-of-day modeling, which is important in the evaluation of pricing policies and peak spreading.
- Modeling NHB WR trips separately may also help the evaluation of transit alternatives because NHB WR may be an important transit market segment, particularly in the downtown area.
- Household survey data tend to be adequate to estimate trip productions and attractions.

Cons:

- Household survey data may not be adequate to develop mode choice models for separate NHB WR and NHBO trip purposes.

Modeling Methods

The methods used to model generation of NHB WR and NHBO trips are similar to those for the NHB trip purpose, including cross-classification, regression, logit-model, and trip rates. In the cross-classification method, trip production rates are developed using household survey data for a variety of cross-classifications, including household size and household income (e.g., BMC, PSRC), household size, number of workers and household income (e.g., SCAG), workers in households with and without children for NHB WR purpose, and persons in households with and without children for NHBO (e.g., SERPM), and household lifecycle (without children, with children, retired) and household income/number of workers (e.g., NJTPA).

Although NHB productions are estimated at the transportation analysis zone (TAZ)-level, they are often not used at household locations, but rather are used as a control total for the number of trips generated. The estimated attractions control where to allocate the control total among the TAZs. Portland METRO initially estimates NHB WR trips by TAZ of production (work), adjusts the estimated productions by a production factor (total employment scaled by 0.803 divided by total productions), and sums the result across all TAZs to arrive at a control total. A logit model is then used to allocate the control total among different TAZs, using the ratio of the TAZ's utility to the sum of all TAZs' utilities. Independent variables in the logit model include households and employment by categories (retail, service, government, and others) (Kim, 2008).

Regression models are very commonly used in estimating attractions for NHB WR and NHBO trips. Independent variables typically include employment by categories. NHB WR trip rates tend to vary with different categories of employment such as basic employment, retail, service, and government. NHBO trip rates may be related to the number of households, in addition to certain employment categories such as retail, service, and government employment. Regression models are used in production estimation in some models such as SCAG's regional model.

Few regional models develop separate mode choice models for NHB WR and NHBO trip purposes. Often, the two trip purposes are lumped together for the mode choice process, as done in models for CATS, CTPS, HGAC, Portland METRO, NJTPA, PSRC, SANDAG, and other MPOs.

2.2 Home-Based College/University

Home-based college/university trips refer to trips made between homes and colleges and universities. The following summarizes the findings of our review:

- Almost half of the large MPO regional travel demand models reviewed by CS for this task have a home-based college/university (HBU) trip purpose. An additional 20 percent of the large MPOs treated colleges/universities as part of special trip generators. Another 10 percent include the HBU trip purpose in the home-based school trip purpose, reflecting these trips through slightly higher home-based school (HBSch) trip rates.
- Based on the survey conducted for TRB Special Report 288 (VHB, 2007), the home-based college/university trip purpose was used in 48 MPOs out of a total of 219 MPOs surveyed.
- Household travel surveys are commonly used to estimate trip production models for home-based college/university trips. Trip attraction models for home-based college/university trips tend to use enrollment or population/households as independent variables.
- Of those MPOs having HBU trip purposes in trip generation, half include them only through the trip distribution process. Half of the MPOs do mode choice modeling for HBU. When not carried forward, HBU trips are generally folded back into the HBO trip purpose.

There are both pros and cons for adding HBU trips to a regional model.

Pros:

- College/universities tend to be major trip generators in a region; thus they can be regionally significant. In mid-sized regions, HBU trips can be a substantial element of local bus ridership. In larger regions, their contribution to transit and traffic congestion may be less noticeable.
- Home-based college/university trips have very distinct travel behaviors that are very different from other HBO trips, including trip frequencies, travel patterns, and mode choice, etc. They also are different from HBSch trips in many ways, particularly for mode choice and trip distance. Separating this trip purpose allows consideration of those factors that affect HBU trips to be modeled explicitly.
- Adding HBU in trip generation can potentially improve the accuracy of model validations downstream, even if HBU trips are not modeled separately in trip distribution and mode choice.
- College/university locations are quite stable over time and enrollment data are generally available for the base year and can be obtained or estimated for future years.

Cons:

- Household survey data may be inadequate in terms of sample size to estimate HBU trip generation models, and even more inadequate for mode choice modeling of HBU trips. Special surveys of college and university students, staff, and faculty may be needed. In many cases, universities may have their own surveys, but unless these are coordinated with the regional planning agency, it is likely the data will be unsuitable for travel forecasting use.
- Forecasting age cohorts, if needed, could be challenging and may require data or developing new assumptions that currently are not being made available by modeled jurisdictions.

Modeling Methods

Methods used to model trip generation of HBU trips are similar to the HBSch trip purpose, including cross-classification, regression, logit-model, and trip rates. Cross-classifications take different forms, including household size by age group (age of household head) as used by Portland METRO, household income and the number of household members with ages between 18 and 24 as used by PSRC (see Table 2) and SCAG, and household lifecycle (without children, with children, retired) and household income/number of workers as used by NJTPA. College age cohorts are estimated as part of the socioeconomic models in the SCAG and PSRC models. While it may seem counterintuitive for households without college-aged members to generate any HBU trips, these average rates include adults involved with continuing education as well as those in graduate studies, who are often older than 24 years old.

Table 2. PSRC College Trip Production by Presence of Household Members Aged 18-24

	Income				
	Less than \$15,000	\$15,000-\$24,999	\$25,000-\$44,999	\$45,000-\$74,999	\$75,000 and Above
No college age	0.14	0.06	0.04	0.03	0.03
1 college age	0.37	0.29	0.26	0.26	0.25
2 and more college age	0.62	0.54	0.52	0.51	0.51

Source: Cambridge Systematics, 2007, Page 6-5.

A potential bias with a household-based approach to estimating HBU trips is over-estimation in areas with a large proportion of senior households, or in areas with a very small proportion of young adult households. Alternatively, modelers can estimate HBU trip rates per student, but the challenge is to estimate the number of college students by zone-of-residence. Census data on school enrollment is the best existing source for

student data at the resident end, as there is a category for “enrolled in college, graduate school, or professional school.” At the attraction end, student and employee trip attraction rates vary significantly across colleges/universities. Some states, such as Florida, have conducted extensive college travel surveys to derive trip attraction rates.

Regression models are very common in estimating attractions for HBU trips. Independent variables typically include school enrollment and/or educational employment. College/university enrollment represents “students by place of attendance” and represents the total number of students enrolled in any public or private post-secondary school (college or university) that grant an associate degree or higher, located within a zone. Table 3 shows examples of HBU trip attraction models from selected MPOs.

Table 3. Summary of Selected MPO HBU Trip Attraction Models

MPO/Model	Variable	Attraction Equation
PSRC	Full-time college enrollment	$0.61 * \text{Full-time college enrollment}$
ARC	College enrollment	$0.8137 * \text{college enrollment (for motorized person trips)}$
Portland METRO	Students or staff	4-year college vehicle trips=students * 2.5 or staff * 9.8 2-year college vehicle trips=students * 1.5 or staff * 28.2
SCAG	College enrollment	$0.549 * \text{College enrollment}$
CTPS	Service employment (K-12) Service employment (college)	$9.25 * \text{K-12 employment} + 3.3 * \text{college employment}$
MTC	College full-time equivalent enrollment	$1.157 * \text{COLL_FTE}$

Notes: PSRC: These include both home-based and college dormitory trips. Portland METRO: modified ITE vehicle trips. MTC: total person trips.

Some regional models carry HBU trips through trip distribution and mode choice processes in a manner somewhat similar to other home-based trip purposes. PSRC distributes HBU trips, not stratified by income, using a gravity model with a logsum as impedance (Cambridge Systematics, 2007). The HBU mode choice model is a multinomial logit model to split total trips among transit walk access, drive alone, shared ride, bicycle, and walk.

The special characteristics of HBU trips often require their special treatment in trip distribution and mode choice. Those living in college dormitories may be very different from those living off-campus in their trip-making characteristics. The PSRC model recognizes this difference and uses a two-step process for the distribution of college trips (Cambridge Systematics, 2007):

1. Identify the number of students living in group quarters that are within walking distance of a college. These trips are then linked directly to the college attractions without a gravity model.

2. Distribute the remaining home-based college trips using a traditional gravity model formulation.

NJTPA uses a unique method to distribute HBU trips. Its process uses the university enrollment database as the controlling total, and temporarily inverts the typical “production-attraction” designations. Trip ends at the household level estimated from cross-classification are designated as attractions, while trip ends associated with off-campus residents at each university are treated as productions. A zero-iteration gravity model technique is used to allocate trip ends to the proportionate trip lengths observed from the survey data, but the model does not attempt to balance the unused “attractions” at the home end. After converting the trip end orientation back to the traditional home-based production-attraction format, the procedure reconciles any differences between the zonal-level HBU trip productions from the cross-classification process and the HBU trip productions allocated by the zero iteration gravity model. In all zones where unallocated trip ends occur, these trip ends are removed from the HBU trip purpose and added to the HBO trip purpose.

Trip tables and growth factors are used to model college trips in the Milwaukee regional travel demand model (SWRPC, 2006). The growth factor procedure is estimated based on projected enrollment and applied to total motorized person trips. A mode choice model is used to divide the total person trips into public transit and auto trips.

Some regional models carry HBU trips through trip distribution and lump them together with other home-based trips in the mode choice model. Examples are ARC, NJTPA, SERPM, and TBRPM. Part of the reason is lack of adequate observations of HBU trips in the travel survey to allow for mode choice model estimation.

2.3 Home-Based School Trips

Home-based school trips generally refer to trips made between homes and schools of kindergarten through 12th grade (K-12) education, namely high schools and below. They are sometime referred as home-based grade school trips to specifically distinguish them from home-based college/university trips. The following summarizes the findings of our review:

- Based on the survey conducted for TRB Special Report 288 (VHB, 2007), the HBSch trip purpose was used in 80 MPOs out of a total of 219 MPOs surveyed.
- CS’ review indicates that it is fairly common for large MPOs to include a distinct HBSch purpose. More than two-thirds of the regional travel demand models reviewed had a home-based school trip purpose.
- Household travel surveys are commonly used to estimate trip production models for home-based school trips. Trip attraction models for home-based school trips tend to use school enrollment or educational employment as independent variables.

Although not as common as home-based work, home-based other, and non-home-based trip purposes, the home-based school trip purpose is used in a number of travel demand models in the U.S., particularly those maintained by large MPOs. There are both pros and cons for establishing a separate HBSch trip purpose.

Pros:

- HBSch trips have very distinct travel behaviors that are very different from other home-based-other trips, including trip frequencies, travel patterns, mode choice, etc. Modeling HBSch separately allows consideration of those factors that affect school trips to be modeled explicitly.
- Developing a distribution model for HBSch trips, incorporating their unique distribution patterns and reflecting the use of school buses and other modes to travel to school, could result in improvement in the accuracy of model validation.
- Existing household survey data, which are used to develop models for traditional trip purposes, tend to be adequate to estimate school trip productions.

Cons:

- Although school locations and enrollment data are generally available for the base year, it may be time-consuming to assemble given the multiple jurisdictions covered by the TPB model area.
- Forecasting age cohorts, if needed, may require data or developing new assumptions that currently are not being made available by modeled jurisdictions.
- It may prove challenging to forecast future school locations and enrollment.

Modeling Methods

Methods used to model trip generation of HBSch trips are similar to other trip purposes, including cross-classification, regression, logit-model, and trip rates. Cross-classification is very common in modeling HBSch trip productions. Examples include models from BMC, Portland METRO, PSRC, SACOG, SCAG, and SEMCOG. Trip production rates are developed using household survey data, for a variety of cross-classifications, including household size and household income as by BMC, household size and number of children as by SEMCOG and Portland METRO, household size and number of workers as by SACOG, and the number of persons age 5 through 17 and household income as by PSRC (see Table 4).

The TPB Version 2.2 travel demand model uses a cross-classification approach to estimating trip productions. If the HBSch trip purpose is added, the cross-classification approach also can be used for HBSch, developing the trip rates from the household travel survey. The model's existing cross-classifications by household size and household income can still be used if the TPB decides not to add school age children as part of the socioeconomic forecasting submodel.

Table 4. PSRC Home-Based School Trip Productions per Household

Number of Children	Income				
	Less than \$15,000	\$15,000-\$24,999	\$25,000-\$44,999	\$45,000-\$74,999	\$75,000 and Above
0 school age	0.00	0.00	0.01	0.06	0.18
1 school age	0.71	0.79	1.01	1.14	1.26
2 school age	1.81	1.89	2.11	2.24	2.36
3 and more school age	3.20	3.28	3.50	3.62	3.75

Source: Cambridge Systematics, 2007, Page 6-7.

A logit-based trip generation model typically relates trip frequencies with independent variables such as travelers' socioeconomic characteristics. Different variables of policy significance such as land use density, design, diversity, and accessibility can be tested for their significance in these models. This form of model allows more independent variables to be incorporated than the cross-classification method and allows continuous independent variables rather than only classification variables to be used.

Regression models are very common in estimating attractions for HBSch. Independent variables typically include school enrollment and/or educational employment. School enrollment represents "students by place of attendance," including the number of K-12 students enrolled in all public and private schools located within a zone. Table 5 shows some examples of HBSch trip attraction models from selected MPOs.

Some regional models carry HBSch trips through trip distribution and mode choice processes in a manner somewhat similar to other home-based trip purposes. For example, BMC models HBSch as not stratified by income. Trip distribution of HBSch trips is modeled using a gravity model with a composite travel time similar to that used by TPB as the impedance value, while the mode choice model adds school bus as a separate choice at the top level, distinct from auto or transit modes. To better reflect trip distribution within and between school districts, K-factors are defined for 10 major jurisdictions used in the model, which also are school districts. The HBSch mode choice coefficients for system variables are essentially the same as those for HBO and HBS trip purposes, with the exception of the CBD dummy and long distance walk to transit variables (Allen, 2006).

Table 5. Summary of Selected MPO HBSch Trip Attraction Models

MPO/Model	Variable	Attraction Equation
PSRC	Educational employment	7.90 * Educational employment
SEMCOG	School enrollment	1.4 * School enrollment
BMC	School enrollment	Baltimore: 1.08 * ENROLL; Washington: 1.18 * ENROLL
Portland METRO	Household size and number of kids	Set=HBSch Productions (cross-classification trip rates)
SCAG	K-12 school enrollment	1.326 * K-12 enrollment
CTPS	Service employment (K-12) Service employment (college)	9.25 * K-12 employment + 3.3 * college employment
MTC	Population age 5-13 High school enrollment	Grade School: HBGSA=HBGSP=POP0513 * 0.923 * 1.314 High School: HBHSA=HSENROLL * 1.314
TBRPM	K-12 school enrollment	1.5 * K-12 enrollment
SACOG	K-12 school enrollment	1.55 * K-12 enrollment
SERPM	School enrollment (for private schools, colleges/universities)	1.85 * enrollment for Palm Beach and Broward Counties 1.75 * enrollment for Miami-Dade County

PSRC distributes HBSch trips using a gravity model with a logsum as impedance (Cambridge Systematics, 2007). These HBSch trips are not stratified by income. The HBSch mode choice model is a simple binomial logit model that splits trips into motorized and non-motorized categories. Motorized modes are further split into transit, drive alone, shared ride two and shared ride three-plus, using mode split factors derived from their household travel survey. Transit trips are assumed to be all bus trips. Non-motorized trips are split into bicycle and walk trips.

Some regional models carry HBSch trips through trip distribution and combine them with other HBO trips in the mode choice model. Examples include ARC, HGAC, NOACA, SEMCOG, SERPM, and TBRPM.

Some regional models do not carry all HBSch trips through the trip distribution and mode choice processes. After estimating HBSch trips in trip generation, one option is to split out school bus trips and carry forward all trips (apart from school-bus trips) into trip distribution. School districts often have data about the number of students taking school buses in their districts. Additionally, school districts may have data about the specific busing routes within their boundaries. These data can be used to estimate school bus trips and, potentially, their distribution; although data assembly could be time consuming and complicated by the number of jurisdictions.

Public schools are different from private schools in terms of school trip making. Southeast Regional Planning Model VI, which covers the Miami area, treats public and private

schools differently (Corradino Group, 2008). For public schools, school districts have their own district boundaries to enroll students. All TAZs within a school district are defined as the production zones, and the TAZ in which the school is located is defined as an attraction zone. HBSch trip tables are directly built for the public school students using actual school board student enrollment information for each school and the related school boundary. However, private schools, which are not bound by districts, draw students from all TAZs. As a result, the gravity model is used to make the trip distribution for private schools among all TAZs. Universities and colleges are typically treated as private schools without school boundaries in models with a home-based university (HBU) component. Similarly, the private school approach can be adapted to address situations such as public magnet schools.

With well-defined school district boundaries, HBSch trip tables can be derived from the observed data and school district policies. School trip tables and growth factors are used in the regional travel demand model for the Southeastern Wisconsin Regional Planning Commission (SWRPC, 2006). Growth factors were applied to the observed 2001 trip tables of school trips, by mode (i.e., automobile, school bus, and public transit). The growth factors were estimated based upon forecast changes in population, and were adjusted to account for potential changes in school service boundaries and the construction of new schools.

2.4 Conclusions

It is a best practice to represent distinct trip purposes in the regional model, so purpose-specific travel behavior such as trip frequencies, trip patterns, and mode choices can be explicitly modeled. Specific variables relevant for specific trip purposes can be explicitly incorporated in the model. More refined market segmentation not only helps improve the model accuracy but also increase the model's capability in policy and planning analysis.

Based on our review and experience, it is recommended that TPB consider the following options for short-term enhancements of the regional travel demand model:

- Split NHB into NHB WR and NHBO trip purposes and model them at least through trip distribution.
- Establish a HBU trip category and model in trip generation and distribution, assuming data are available to support it.
- Establish a HBSch trip category for trip generation and distribution, assuming data are available to support it.

■ 3.0 Special Generator Modeling

The TPB model currently handles the three commercial airports as special generators, but not universities, group quarters, regional shopping centers, or military bases. TPB staff is considering whether the TPB model should enhance and/or explicitly include special generators in the upcoming Version 2.3 model update.

Special generators can be subdivided into three groups for travel forecasting purposes, based on the terminology of Kurth et al. (1997a): “regular” special generators, “periodic” special generators, and “special” special generators. “Regular” special generators produce trips on a regular, weekday basis, for examples, airports, regional shopping centers, hospitals, and schools. These are and should be accounted for in regional travel demand forecasting, because travel forecasting for transportation and air quality planning is usually done for an average weekday. “Periodic” special generators do not produce trips on a regular weekday basis, for example, some convention centers, infrequently used stadia and arenas, and fairs and festivals. They also are called special events. Because of their relative infrequency, these special generators are usually not considered in the regional travel demand model for transportation planning. “Special” special generators include those sites or activities that cannot be easily classified as regular or periodic special generators, for example, ski areas and historic mountain community gaming districts.

3.1 Modeling Special Generators or Not

In the Fiscal Year 2009 Task Reports, Cambridge Systematics recommended enhancements to or addition of three special generators (airports, visitors, and special events) and discussed data needs, model estimation, validation, and implementation of three such models in the TPB model framework.

Modeling airport trips is a common practice in regional travel demand models of large MPOs. In a survey conducted in 2007, about two-thirds of the 23 MPOs responding reported modeling airport passenger trips in their regional models (Gosling 2008). Most of the 23 MPOs are among the country’s 35 largest MPOs. CS’ review of regional models as part of this task indicated that over 80 percent model air passenger trips in some way.

TPB’s ongoing air passenger survey program can provide information in support of airport trip modeling. TPB, in conjunction with Maryland Aviation Administration (MAA) and Metropolitan Washington Airport Authority (MWAA), has conducted Air Passenger Surveys in the region for nearly 30 years. Surveys currently are conducted at the three major commercial airports over two weeks (weekdays and week-ends) every two years. The surveys collect information on origin airport, destination airport, residency, socioeconomic data, and arrival trip data (origin location, purpose, and mode). It is worth noting that the surveys include household income, household size, and vehicle availability variables, which are critical to the estimation of mode choice models.

The latest air passenger survey shows that transit modes (including regular transit and airport transit) account for approximately 20 percent of total local weekday originations (59,900 in the modeling domain) to the airports (Milone, 2010), a significant transit share. In particular, Washington National Airport (DCA) has almost 6,000 transit trips to the airport, accounting for 28 percent of the total. With the extension of Metrorail to Dulles International Airport, transit trips to Dulles will potentially increase. These data suggest the importance of developing airport ground access mode choice models.

Visitor trips are often overlooked in regional travel demand models. Of the regional models reviewed by CS for this task, only a quarter explicitly account for trips made by visitors. Almost all of the models that include visitor trips treat them as special generators, without formal visitor models used. However, areas with many tourists, such as Orlando and Miami, Florida, and San Diego, California do tend to explicitly account for visitor/tourist trips.

Washington, D.C. is among the top tourist destination cities in America. The “National Mall and Memorial Parks” are ranked third on the list of top 25 most visited tourist destinations in America according to Forbes (Murray, 2010). According to the Washington, D.C. Convention and Tourism Center, D.C. welcomed 16.2 million visitors, including 1.2 million international visitors in 2007, which translates to 44,000 tourists on an average day (Destination DC, 2008).

Another important component of visitor or non-resident trips is those people who come to work or shop in the region. These travels are represented by external-to-internal trips. But those people making external-to-internal trips also generally make internal trips within the modeling domain. Many non-residents enter the Washington region by autos. Based on the TPB Version 2.2 model, external to internal trips by auto drivers are projected to grow from 486,000 in 2000 to 842,000 every day in 2030, an increase of approximately 70 percent. Accurately representing visitor trips is, therefore, potentially important for travel demand forecasting and transportation planning in the region, particularly for downtown Washington D.C.

Very few MPOs model special events as part of the regional travel demand model. Of the regional models reviewed by CS, less than 10 percent model special events as special generators. However, special events can be very important for major investment studies such as New Starts applications, because of their potential to significantly benefit from the proposed transportation investment.

Very few MPOs model shopping centers specifically as a special generator in their regional travel demand model. Of the regional models reviewed by CS, less than 10 percent model shopping centers as special generators. One reason may be that it is uncertain whether the travel related to the existing shopping centers will continue the same way 20 or 30 years in the future.

It is not a common practice to model group quarters, military bases, and colleges/universities as special generators. Based on the Census definition, a group quarter is “a place where people live or stay, in a group living arrangement, that is owned or managed by an entity or organization providing housing and/or services for the residents.” Group

quarter facilities may include correctional facilities, juvenile facilities, nursing homes, hospitals with long-term care facilities, college or university dormitories, fraternities and sororities, dormitories for workers, religious group quarters, shelters, and group homes. Residents of group quarters represent a diverse group of people with distinctly different travel behaviors (e.g., students in dormitories versus senior citizens in nursing homes).

Of the regional models reviewed by CS, only a quarter explicitly treat these trip generators as special generators. As indicated in the previous section on HBU trips, it is a common practice to treat HBU trips as a distinct trip purpose. Group quarters and military bases are often missing from household travel surveys, and are thus not often (or easily) represented in regional models. For the Washington region, the group quarters population totaled approximately 120,000 (about two percent of total population) in 2000, and is projected to increase to 150,000 in 2030. The current TPB model does not use group quarters population in the household-based trip generation estimation process. However, considering adding representation of travel for group quarters population, even in a crude way, is probably better than ignoring it completely as a short-term enhancement.

3.2 Modeling Methodology

Regular special generators are often modeled in regional travel models in different ways:

- As a separate trip purpose (the approach often used for schools and colleges);
- As part of a trip purpose (the approach often used for shopping centers, i.e., being part of home-based shopping trip purpose); and
- As a stand-alone model with its own modeling components (the approach often used for airport trips).

Based on a recent Airport Cooperative Research Program study, the current best practice in modeling airport trips is to develop a separate model based outside the modeling stream for traditional trip purposes and to include ground access mode choice models with a nested logit structure for at least four market segments (resident business trips, resident non-business trips, non-resident business trips, non-resident non-business trips) (Gosling, 2008). A suggested nesting structure has three nests: private vehicle modes (with different parking options as a second-level nest), exclusive ride on-demand modes (taxi and limousine), and group shared-ride scheduled modes (public transit with different transit access options as a second-level nest and scheduled airport bus). In terms of trip generation and distribution, airport trips, taken as a special case, are usually modeled as a combined direct demand model with a single destination. District-level socio-economic variables are often used to estimate regression models forecasting the number of trip ends based on a set of socioeconomic variables.

Periodic special generators or special events are not usually included in most regional travel demand models, but are sometimes considered in major investment studies, such as New Starts applications. Periodic special generators can be modeled:

- As part of the existing four-step model framework with some special treatments; and
- As a stand-alone model with its own modeling components.

Kurth et al. (1997a) developed a methodology of modeling periodic special generators or special events by using the existing four-step model framework, without collecting special survey data. This methodology treats periodic special generators or special events as home-based non-work and NHB trips and uses attendance projections for trip generations. A gravity model distribution is made on a distance basis, or non-distance basis, depending on whether the event under consideration has any substitution. Professional sports events do not have any substitution and are considered to be independent of travel time or distance. For these types of events, distribution is made on a non-distance basis with all friction factors set to one. Mode choice models for home-based non-work and NHB trips in the regional model can be used for modeling mode choices for periodic special generators or special events, with special consideration of event parking costs and/or transit fare.

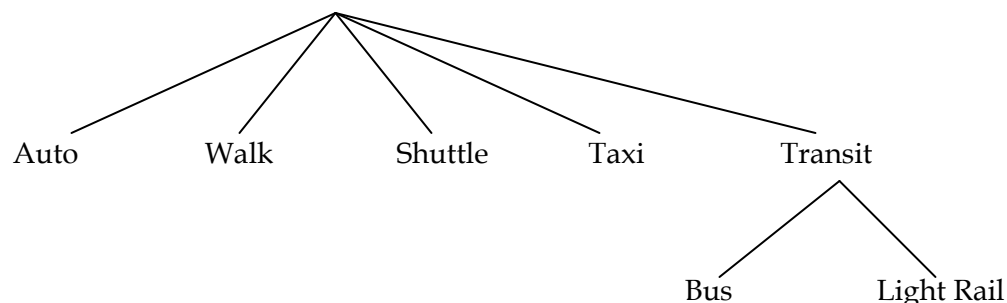
Cambridge Systematics developed a visitor model and a special events model for the Dallas Area Rapid Transit (DART) (Kuppam, Johnson, and Rossi, 2009). The visitor modeling process consists of trip generation, destination choice, and mode choice model components. These visitor models were developed using hotel survey data collected by DART in the Dallas CBD area. The special events models were estimated using special event survey data collected by DART at the American Airlines Center and the Meyerson Symphony Center.

In the DART visitor models, a trip production is defined as the hotel end of a hotel-based trip or as the origin of a non-hotel-based trip. The visitor trip production model estimates person trip rates (per occupied hotel room) on an average weekday for each of the seven hotels surveyed in the CBD area. Trip rates were estimated for business visitors and leisure visitors for four trip purposes (meal, business, other, and non-hotel-based).

A trip attraction is defined as the non-hotel end of a hotel-based trip or the destination of a non-hotel-based trip. A regression model was developed to establish a linear relationship between trips attracted by zone and socioeconomic characteristics in each zone (e.g., service employment). Separate regression models were developed for business and leisure travelers.

The visitor destination choice models represent the choice of trip destinations made by visitors while they are staying in hotels in downtown Dallas. Four separate multinomial logit models were developed based on purpose of visit in Dallas – business and non-business (leisure); and land use type of destination zones – CBD and Non-CBD. Variables include logsum from the mode choice model, distance, and total number of trip attractions to a destination zone.

The visitor mode choice models are nested logit models developed for business visitors and leisure visitors, separately, with the structure illustrated in Figure 1. Independent variables include transportation level-of-service, distance, auto availability, and a CBD dummy.

Figure 1. DART Visitor Mode Choice Model Structure

The DART special events models consist of two components: origin location and mode choice. The origin location model estimates the locations from which event attendees travel to the event and return to after the event. The special events origin location models are estimated as destination choice models with a structure of multinomial logit models (MNL). Variables include logsum from the mode choice model, distance, and total number of trip attractions to a destination zone. The mode choice model estimates the probability of using DART light rail to travel to and from the event. The special events mode choice models have a binary logit structure with two modal alternatives: auto and light rail. Independent variables include transportation level-of-service and auto availability.

A special interest has been expressed by TPB in the area of modeling group quarters population travel. Modeling this as a special generator often relies on special travel surveys or Institute of Transportation Engineers (ITE) trip rates. In the PSRC trip-based model (CS, 2007), trip generation rates for non-institutional group quarters were derived from a variety of sources, including a university trip model developed for the University of Michigan, a special generator model developed for the MacDill Air Force Base in Tampa Bay (Florida), a retired person's model developed in Tucson (Arizona), and ITE trip generation rates. Table 6 summarizes the PSRC trip generation rates per person by type and trip purpose, compared to the regional model averages.

BMC differentiates group quarters trip rate assumptions between institutional (IGQ) and non-institutional (NIGQ) populations (institutionalized population includes people under formally authorized, supervised care or custody). In the BMC model it is assumed that IGQ people generate trips at the same rate (per person) as low-income, one-person households, except that they probably make no HBW, journey to work (JTW), journey at work (JAW), or HBSch trips. NIGQ people are assumed to generate trips at the same rate as one-person, one-worker, income level three households, except for HBSch trips (Allen, 2006).

Table 6. PSRC Non-Institutional Group Quarters Trip Rates (Per Person)

Trip Purpose	College Dormitories	Military Quarters	Retirement Homes	PSRC Regional Average for Household
Home-Based Work	-	0.37	0.10	0.60
Home-Based College	1.18	-	-	0.08
Home-Based School	-	-	0.03	0.29
Home-Based Shop	0.40	0.74	0.70	0.40
Home-Based Other	1.24	1.09	1.49	1.09
Non-Home-Based	1.00	0.76	1.04	1.03
Total	3.82	2.97	3.37	3.48

Source: Cambridge Systematics, 2007.

3.2 Conclusions

In summary, it is recommended that TPB consider the following options for short-term enhancements of the regional travel demand model:

- Develop an airport trip submodel set incorporating the current best practice, taking full advantage of the ongoing air passenger travel survey data. This airport model would include ground access mode choice models with a nested logit structure for at least four market segments (resident business trips, resident non-business trips, non-resident business trips, and non-resident non-business trips).
- Plan a visitor travel survey and a special events survey in support of model development for a visitor model and a special events model. At the same time, an evaluation of the exogenously generated visitor/tourist auto driver trips data can be made to see if the data fully account for trips made by visitors within the region. Interim enhancements can be made using some simple assumptions and adjustments.
- Model HBU trips as an independent trip purpose as recommended in Section 2.2. Other college-related trips from college dormitories and other group quarters trips should be estimated using simplified assumptions or using trip rates from other similar regions, and checked against the ITE trip rates.
- It is not recommended to treat shopping centers as special generators.
- Explore use of ITE trip generation rates for treatment of group quarters trip generation.

■ 4.0 References

Allen, William G. (2006). Baltimore Metropolitan Council Travel Forecasting Model Calibration Report.

Atlanta Regional Commission (ARC) 2008. The Travel Forecasting Model Set for the Atlanta Region: 2008 Documentation.

Cambridge Systematics (1997). Enhancement of DVRPC's Travel Simulation Models - Task 4: Transit Assignments. Prepared for Delaware Valley Regional Planning Commission, April 1997.

Cambridge Systematics (2002). Southeast Michigan Council of Governments (SEMCOG) Travel Model Documentation.

Cambridge Systematics (2004). North Central Texas Council of Governments (NCTCOG) Mode Choice Model Estimation.

Cambridge Systematics (2007). PSRC Travel Model Documentation (for Version 1.0): Updated for Congestion Relief Analysis. Final Report.

Cambridge Systematics (2008). FSUTMS-Cube Framework Phase II Model Calibration and Validation Standards Final Report. Prepared for Florida Department of Transportation Systems Planning Office.

Cambridge Systematics (2009). Fiscal Year 2009 Task Reports. Final Report. National Capital Region Transportation Planning Board.

Central Transportation Planning Staff (2009). Methodology and Assumptions of Central Transportation Planning Staff Regional Travel Demand Modeling: North Shore DEIR Support.

Corradino Group (2008). Southeast Regional Planning Model VI. 2000 and 2030 Models, TECHNICAL REPORT 2, Model Calibration and Validation. Submitted to Florida Department of Transportation - District IV.

Delaware Valley Regional Planning Commission (DVRPC) (2008). 2000 and 2005 Validation of the DVPRC Regional Simulation Models.

Destination DC (2008). "Washington, DC Posts 7% Increase in Visitation in 2007 as Visitor Spending Tops \$5.5 Billion." Press Release. August, 27.

DKS Associates (2002). Sacramento Regional Travel Demand Model: Version 2001 (SACMET 01) Model Update Report. Prepared for Sacramento Area Council of Governments (SACOG).

Gannett Fleming (2007). Tampa Bay Regional Planning Model, Cube Voyager Conversion, Technical Report No. 1, Validation Report. 2007.

Gosling, G.D. (2008). Airport Ground Access Mode Choice Models. Airport Cooperative Research Program Synthesis 5. Washington, D.C.: Transportation Research Board.

Houston-Galveston Area Council (HGAC) (2009). 2005 Cube Voyager Model Validation and Documentation Report.

Kim, K. (2008). Metro Travel Forecasting 2008 Trip-Based Demand Model Methodology Report. Portland METRO.

Koppelman, F.S. and C. Bhat (2006). A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models. Prepared For Federal Transit Administration.

Kuppam, A.R., P. Johnson, T. Rossi, A. Banerjee, S. Kumar (2009). Visitor Travel Model, Special Events Model, and Parking Location Choice Model for Studying Transit Improvements: Downtown Dallas Transit Study. Presented at 12th Transportation Research Board National Transportation Planning Applications Conference, Houston, Texas. May 17-21.

Kurth, D., B.V. Meter, S. Myung, M. Schaefer (1997a). Incorporating Special Generator Ridership in Major Investment Analyses and Transit Corridor Studies. Presented at 76th Annual Meeting of the Transportation Research Board, Washington, D.C.

Kurth, D., B.V. Meter, S. Myung, M. Schaefer (1997b). Quantifying Special Generator Ridership in Transit Analyses. Sixth TRB Conference on the Application of Transportation Planning Methods, Dearborn, Michigan.

Martin, W. A. and N. McGuckin (1998). Travel Estimation Techniques for Urban Planning. NCHRP Report 365.

Milone, R. (2010). TPB's Development of Airport Passenger Trip Forecasts, Presented to the Travel Forecasting Subcommittee on March 19.

Milone, R., H. Humeida, M. Moran, and M. Seifu (2009). FY 2009 Development Program for TPB Travel Forecasting Models. Draft report. Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board (COG/TPB). June 30.

Milone, R., H. Humeida, M. Moran, and M. Seifu (2008a). TPB Travel Forecasting Model, Version 2.2: Specification, Validation, and User's Guide. Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board (COG/TPB). March 1.

Milone, R., H. Humeida, M. Moran, and M. Seifu (2008b). TPB Travel Forecasting Model, Version 2.3: Specification, Validation, and User's Guide. Draft report. Washington, D.C.:

Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board (COG/TPB). June 30.

Murray, V. (2010). "America's Top Tourist Attractions," Forbes.com <http://www.forbes.com/2010/05/20/top-tourist-attractions-lifestyle-travel-magic-kingdom-disneyland-times-square.html>. Accessed May 28.

North Central Texas Council of Governments (NCTCOG) (2000). Dallas-Fort Worth Regional Travel Demand Model (DFWRTM): Description of the Multimodal Forecasting Process.

North Central Texas Council of Governments (NCTCOG) (2009). Dallas-Fort Worth Regional Travel Demand Model (DFWRTM): Model Description.

Northeast Ohio Areawide Coordinating Agency (NOACA) (2005). Travel Demand Model Validation: NOACA Model Update with 1,236 Zones and 82 External Stations.

Purvis, C. L. (1997). Travel Models for the San Francisco Bay Area (BAYCAST-90), Technical Summary, Metropolitan Transportation Commission.

San Diego Association of Governments (SANDAG) (2008). 2030 Regional Growth Forecast Update: Process and Model Documentation.

Southeastern Wisconsin Regional Planning Commission (SWRPC) (2006). A Regional Transportation System Plan for Southeastern Wisconsin: 2035. Planning Report No. 49.

Southern California Association of Governments (SCAG) (2003). SCAG Regional Transportation Model: Model Validation and Summary.

URS Corporation (2008). North Jersey Regional Transportation Model Enhanced. Prepared for North Jersey Transportation Authority (NJTPA).

Vanasse Hangen Brustlin, Inc. (VHB) (2007). Determination of the State of the Practice in Metropolitan Area Travel Forecasting: Findings of the Surveys of Metropolitan Planning Organizations: Prepared for Committee B0090 TRB/National Research Council.

This page intentionally left blank.

Short-Term Model Enhancements: Time of Day Model, Queue Delay Function, and Multi-Run Assignment

■ 1.0 Introduction

The National Capital Region Transportation Planning Board (TPB) tasked Cambridge Systematics (CS) to investigate four topics pertaining to short-term enhancements of the regional travel demand model:

1. Further consideration of time of day models, including providing concrete examples of how developing such a model would be beneficial for pricing studies; providing feedback on how a time of day model might be developed using INRIX data; providing additional information on data requirements for validation and calibration; and providing discussion of proper usage of the time period traffic assignments generated during application of the existing model;
2. Consideration of a queuing delay function, including review of other metropolitan planning organizations (MPOs) using these and recommendations for the use of such a function by TPB;
3. Expanded discussion of procedures to eliminate the multi-run assignment process – skimming high-occupancy/toll (HOT) facilities once for both HOT and high occupancy vehicle (HOV) paths – to reduce model run time; and
4. Discussion of experience and recommendations regarding Citilabs' Application/Scenario Manager as an alternative to the batch file application approach that TPB currently employs.

This report documents CS' reviews, findings, and recommendations on these four issues. CS reviewed the state of practice on time-of-day representation in regional models from a sample of the country's largest metropolitan planning organizations (MPOs). Based on these reviews and CS' experience, this memo discusses the time of day choice model, its benefits, particularly associated with pricing studies, its development issues such as data requirements, model validation and calibration, and finally we provide some direction and guidance on the issue.

The staff at the TPB has incorporated a queue delay function (QDF) into the Version 2.2 model as a method for minimizing overloading the highway and ramp links in the model highway network. CS reviewed the state of practice for consideration of a queue delay in regional models, and discusses two approaches for incorporating queue delay in regional models. CS also contacted several MPOs for their experiences and current practice. As part of this task and in order to better understand the impact of the QDF, CS conducted a

test model run to examine the effects of applying the existing queue delay function to all facility types in the network.

The current Version 2.2 model requires two separate model runs to be performed, a “base run” and a “conformity run.” The “base run” produces HOV skims for the HOT facilities which treat the HOT facilities as HOV facilities. This process was developed to accommodate the stated policy of Virginia Department of Transportation (VDOT) that HOT facilities will not degrade the operations of HOV users. CS proposed combining the two runs into a one-run process to save model run time. This concept was tested and the results are documented in this memo.

Finally, this memo discusses CS’ experience with Citilabs’ Application/Scenario Manager and makes a recommendation regarding its use in the next model update.

■ 2.0 Time of Day Models

While not one of the main components of the four-step modeling process, many regional travel demand models use one of two primary methods to estimate the spread of trips across different times of the day: 1) fixed factors or 2) choice models. The more common method is to divide the day into several time periods (typically between three and five) and develop a set of fixed factors to split trips into time periods based on trip purpose, direction, and sometimes mode. In trip-based models, this typically occurs either before or after the mode choice step while productions and attractions are being converted into origins and destinations. Factors are most often developed on a regionwide basis from household travel surveys and traffic count data, although other sources can be used. TPB’s use of three time periods (morning peak, evening peak, and off-peak) and fixed factors by mode, purpose, and direction is consistent with the current state of the practice.

Time of day choice models are less commonly used in trip-based models. This method apportions trips to defined time periods based on characteristics of the trip and the trip maker such as travel times, trip purpose, socioeconomic characteristics of the traveler, and density variables. The effects of congestion levels on the choice between time periods can be evaluated, as well as the sensitivity of travelers by trip purposes and income. Although not necessarily true, choice models typically use far more time periods than the fixed factor methodology, allowing for a more fine grained analysis of peak period travel.

2.1 Benefits of Time of Day Choice Models

There are two main advantages for implementing a time of day choice model. First, choice models are able to forecast changes in time of day splits in the future, which is not possible using the fixed factor methodology. Secondly, time of day choice models are able to capture the effects of “peak spreading” and time of day changes caused by congestion and variable pricing policies. In a recent FHWA research project, a time of day choice

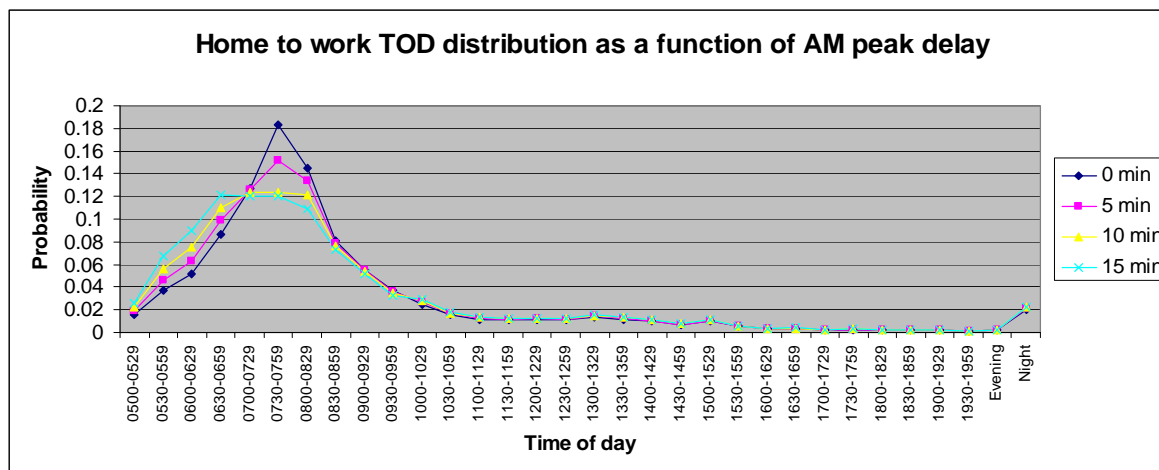
methodology for trip-based models was developed, and tested and validated the methodology in Denver and San Francisco. Later, the methodology was refined for the Puget Sound Regional Council (PSRC) model. These models are sensitive to travel behavior related to time of day, such as demographic and land use factors, as well as changes in time and cost across detailed time periods (e.g., 30-minute periods). As a result, they can predict peak spreading, as well as the shifts in time of day that may occur in the future under various dynamic or congestion pricing scenarios. These experiences demonstrate a promising, applicable approach to model detailed time-of-day choices using available data sources. The PSRC model incorporates a range of variables that describe individual travelers and individual trips including household income, household size, carpooling, and level of congestion (Kuppam, Outwater, and Hranac, 2008).

By incorporating elements that change over the course of the day (travel time and cost) and over the time of years (demographics like household income or size) into the decision of when to travel, the time of day choice model is able to capture the reasons behind the decision, not just the results that are incorporated by the fixed factor method. Whereas the fixed factor method assumes the same time of day splits in future years, a time of day choice model can predict future time of day splits indicative of future conditions.

Because congestion and travel cost are not explicitly included in the fixed factor method, the model cannot capture time of day patterns caused by these types of changes. The choice model approach can address this by incorporating these variables into the decision process. In order to increase the sensitivity of the time of day choice model and to capture the effects of peak spreading, the number of time periods used in the model should be increased. For example, in the PSRC time of day choice model, the number of time periods was increased from five to 32, 30 of which are 30-minute periods that previously comprised the morning peak, evening peak, and midday periods. This allows for analysis of more fine-grained changes in time of day behavior during the peak periods, when heavy congestion or high tolls may cause travelers to shift their travel by 30 minutes or so. Figure 1 shows an example of the effects of peak spreading caused by travel delay in the PSRC model. As shown, most of the shifting occurs within the peak, with a higher probability of choosing the shoulder periods – particularly the earlier shoulder. There is also some probability of moving out of the period altogether into one of the other larger periods (e.g., some of the p.m. peak trips moving into the midday and evening periods).

The model's sensitivity to changes in congestion levels and travel costs allows planners to analyze the effects of peak spreading on a transportation network. This is a very important component when considering pricing policies, including congestion pricing and variable toll facilities. Using this methodology, prices can be varied for each time period (e.g., a different toll rate for each 30-minute period), and the effects of these changes can be observed. Different toll rates can be analyzed to determine the optimal configuration for a specific facility, and accurate revenue projections can be developed.

Figure 1. Home to Work Distribution as a Function of Morning Peak Delay



Source: WSDOT, 2009.

Different travel market segments (e.g., households by income, various trip purposes) react to the congestion level and pricing differently and make/change their choices on time of day travel based on their willingness to pay or experience travel time delays. Choice models take these factors into account, with the ability to represent travelers' shift in choice of the time of day based on transportation system performance and the traveler's sensitivity to travel costs. Value of time, which is already developed for each market segment as part of the mode choice model, would also therefore be included in the time of day choice model.

The updated PSRC time of day choice model has been used to analyze the effects of different pricing scenarios in the Seattle region including fully tolled facilities and HOT lanes (WSDOT, 2009). The time of day choice model reveals appropriate sensitivity to the characteristics of the travelers and their trips such that the success of different toll policies can now be measured using the PSRC model.

2.2 Development of a Time of Day Choice Model

The development of the type of time of day model described in the previous section requires the estimation of logit choice models for different market segments. In the case of PSRC's model, six multinomial logit models were estimated by trip purpose and direction. Estimation of similar models by TPB would require time of day information about trips and trip makers from a household survey. The time of day information provided in the survey must be at the same level of detail as the time period structure. Data requirements for estimation would include socioeconomic characteristics of trip makers (i.e., household income and size), land use characteristics (i.e., total employment levels accessible within certain time or distance radius), and level of service characteristics

of the trip (i.e., generalized in-vehicle cost in minutes). It is likely that not all of these variables will be carried through into the final model specifications.

Time of day choice models can be difficult to estimate due to the nature of the choice being modeled. Time-of-day choice is dependent on the congested travel times, while the congested travel times are also dependent on the time of day choices made by travelers. This two-way relationship can make estimation of a regression type model difficult, even given large data sets.

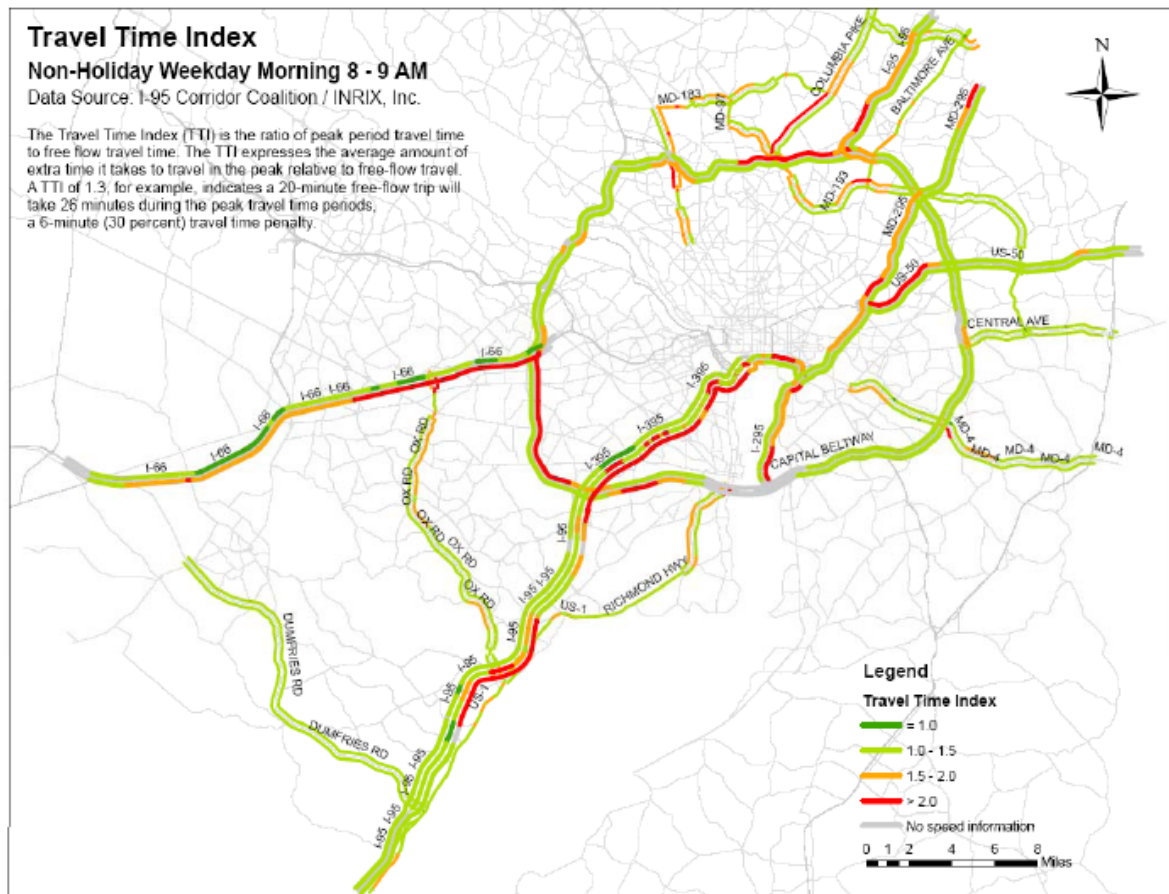
As part of the development of a time of day choice model for TPB, it is recommended that the daily trip tables be divided into at least four or preferably five large time periods. This would involve breaking the off-peak period into midday and night periods; an evening time period could be included as well. This is recommended in order to effectively capture the full effects of peak shifting, which may cause trips to shift out of the peak periods all together. The off-peak period is currently too large with different, distinct sub-periods, and thus its breakdown into a set of small, more specific off-peak periods would help capture travel behaviors in different off-peak sub periods and help model peak spreading.

Data Requirements

Most of this required data would be available from a household travel survey and existing land use forecasts. What is missing includes service characteristics of the trips. Travel times and travel costs – combined into a measure of generalized cost – for each time period are necessary for the development of the choice model. Cost elements, including vehicle operations and tolls, are generally a matter of policy and should be easily defined for all trips in the network. Travel times for different time periods are somewhat more difficult to develop. In the PSRC model, initial travel times are calculated based on the difference between the congested travel time during the period in question and the uncongested travel time. Travel times are necessary for each origin-destination pair.

As part of the Congestion Management Process (CMP), TPB has obtained data detailing traffic speeds on major highway facilities across the region (Pu and Meese, 2009). The data provided by INRIX, Inc. were used to calculate the ratio between congested and free flow speeds as a “Travel Time Index” for each hour of the day. As shown in Figure 2, the INRIX data do not cover the majority of roadways in the region, and therefore cannot provide travel time data for all origin-destination pairs. Origin-destination specific travel times are necessary in order to accurately capture the conditions of an individual time of day choice. This data cannot be used as the primary data source for developing travel time data for a time of day choice model. Instead, a procedure similar to that used for PSRC could be implemented, where congested travel times are developed for each large period (morning peak, evening peak, and off peak) based on the existing fixed time of day factors and compared with uncongested travel times. These times could be used as the starting point for developing congested speeds for each of the smaller time periods.

Figure 2. INRIX Travel Time Index Coverage



Source: Pu and Meese, 2009.

Validation and Calibration

A time of day choice model will need to be calibrated and validated prior to application; as with the PSRC model, a two-step process is recommended. PSRC felt that the most important check was the traffic levels on the roadway network; as such the first tier of validation and calibration involved matching the estimated model results to observed traffic levels. This was mostly measured as vehicle miles of travel (VMT) by time of day, although other aggregate regionwide measures by time period could be used as well. Another useful validation measure could be a comparison of congested versus free flow speeds by functional class and/or area type, although congested speeds are not always estimated well by link-based highway assignment methods. In this measure, the INRIX Travel Time Index data might be useful to provide an outside source of congested speeds at a fine-grained level of time detail for freeways and some other major facilities.

In the second tier, the estimated model results should be calibrated and validated so that they match the results of the survey data. This is chosen for the second tier since reported time of day travel is a less reliable data source than observed traffic data.

2.3 Usage of Existing Time of Day Model Results

The current TPB time of day model divides the daily trip table into three time periods (morning peak, evening peak, and off-peak) using static factors for each mode (auto driver, drive alone, and carpool), trip purpose, and direction. These factors are developed at the regional level to match the time of day spread across the metropolitan area. Because of the regional nature of the calibration, the time of day factors are likely not accurate for individual facilities or intersections. It is therefore important when analyzing traffic flows in these small areas that model users do not use the time of day outputs from the regional model directly. Instead, it is recommended that jurisdictions develop time of day factors specific to each facility based on existing traffic data. Using existing traffic data, appropriate factors can be applied to the daily travel demand forecasts for any period of interest. These results will better represent the local daily travel patterns at these smaller areas of interest.

■ 3.0 Queue Delay Function

The queue delay function (QDF) was introduced as a method for minimizing overloading on highway and ramp links in the TPB network. The function attempts to mimic the delay caused by queuing that occurs in a congested network by adding minutes of delay to a link's congested travel time. The QDF is used only on highway and ramp links that are associated with interchanges or locations where capacity drops significantly between adjacent links. Based entirely on the volume-to-capacity (V/C) ratio for the link, the QDF is added to the volume delay function (VDF), ranging in value from zero (at a V/C ratio of 0.8 or less) to 14 minutes (at a V/C ratio greater than 1.4). The queuing delay is not related to the length of the link, so it is possible for a very short link to have a very high level of queuing delay.

In order to ensure that queue delay is only incorporated on the appropriate links in the network, three link attributes are coded into the network for each of the three time periods. This excludes all links with certain characteristics (those one-way links with a single one-way link in and a single one-way link out with the same number of lanes) from having queuing delay added to the travel time.

The VDFs used in the Version 2.2 TPB model incorporate minimum speeds that occur when the V/C ratio has a value of 3.0 or higher. These “speed floors” vary by facility type, as follows:

- Varies from 0.90 to 1.10 mph for freeways;
- Varies from 0.86 to 1.74 mph for arterials; and
- Varies from 1.16 to 2.33 mph for collectors.

By adding up to 14 minutes in additional queuing delay, the QDF serves to further decrease speeds below these minimum levels on appropriate links. Because the QDF is not a function of distance, speeds can be affected by the QDF regardless of the link length. For example, the speed on a 1.0 mile long link operating at 2.0 mph with a V/C ratio of 1.5 will decrease to 1.4 mph, while a 0.25 mile long link operating under the same conditions will decrease its congested speed to 0.7 mph.

3.1 Use of Queue Delay Function in Regional Models

The relationship between traffic speed and flow has long been a subject of interest in the transportation modeling community. While the theories form the basis of traffic operation analysis and much of microscopic traffic simulation, the major application in travel demand forecasting has typically been in calculating delays caused by congestion. Most regional travel models address delay during the assignment phase through the use of a VDF which calculates congested speed based on the V/C ratio. These VDFs are designed to capture congestion delay along a route and were traditionally based on link volumes. On non-freeway links the source of delay in a network is often from the movements at intersections and junctions.

Two main methods have been utilized in practice to account for increased delay in congested networks. First, new functional forms of link-based VDFs have been developed that more accurately represent the breakdown in traffic flow at very high volume levels. These functions may incorporate delay at intersections and junctions implicitly by increasing the level of delay experienced on congested links. A second and less commonly used method is to explicitly incorporate the delay experienced at intersections and junctions by developing a VDF that is both link-based and node-based. This is the method that has been adopted by TPB in the form of a QDF described previously and has been implemented in various forms by several other MPOs around the country.

Volume Delay Functions

The volume delay functional forms traditionally used in travel demand models were those of the Bureau of Public Roads (BPR) and its derivatives. New functional forms of VDFs have been developed that provide additional benefits. Conical VDFs were primarily developed to improve model run times, but have the additional effect of lowering delays on links with very high volumes when compared to the BPR functions. Version 2.2 of the TPB model uses a conical VDF to calculate delay on links in the network.

Akçelik VDFs, a newer functional form, have shown certain advantages in calculating network delay when compared to both the BPR and conical forms. This functional form is consistent with queuing theory (Dowling, Singh, and Cheng, 1998) in that:

- For V/C ratios greater than 1.0, the speeds predicted by the Akçelik curve drop fairly rapidly at the rate predicted by the queuing theory; and

- The Akçelik curve has the property of maintaining a linear increase in link travel times for V/C ratios greater than 1.0.

Studies indicate that the Akçelik functional form produces significantly more accurate speed estimates than other VDF forms. In addition, the properties of Akçelik functions result in higher delays on facilities with very high volumes than the conical functions and are computationally more efficient than the BPR functions.

Dowling and Skabardonis (2006) evaluated several different speed-flow relationships such as exponential, BPR, and Akçelik functions, against observed data for under-capacity arterials in the Los Angeles area. The fitted speed-flow equations along with the standard BPR equation were then evaluated for their ability to predict delays for congested conditions where one or more intersections on the arterial are over-capacity. The theoretical delay due to vehicles waiting their turn to clear the bottleneck intersection on the arterial was computed using classical deterministic queuing theory. Of the VDFs tested, the Akçelik equation performed the best for over-capacity situations and performed as well as other forms for under-capacity conditions. VDOT also recently evaluated speed-flow relationships (Raw et al., 2004) using empirical data to test and calibrate the BPR, conical, and Akçelik VDFs. The initial findings of this analysis confirm that the Akçelik function works very well.

Several MPOs have addressed their facility overloading issues through the implementation of Akçelik VDFs, including the Metropolitan Transportation Commission of the San Francisco Bay Area (MTC) and the Southern California Association of Governments (SCAG) in Los Angeles. A San Francisco Bay Area study (Singh and Dowling, 1999) shows that the Akçelik function results in more realistic assigned traffic volumes that tend to cluster more closely around a V/C ratio of 1.0, with far fewer links with V/C ratios in excess of 1.5 than obtained with traditional BPR curves and their variants. Furthermore, the Akçelik curve did not adversely affect equilibrium assignment model run times. SCAG has also adopted Akçelik VDFs for arterials and freeways in the regional travel demand model.

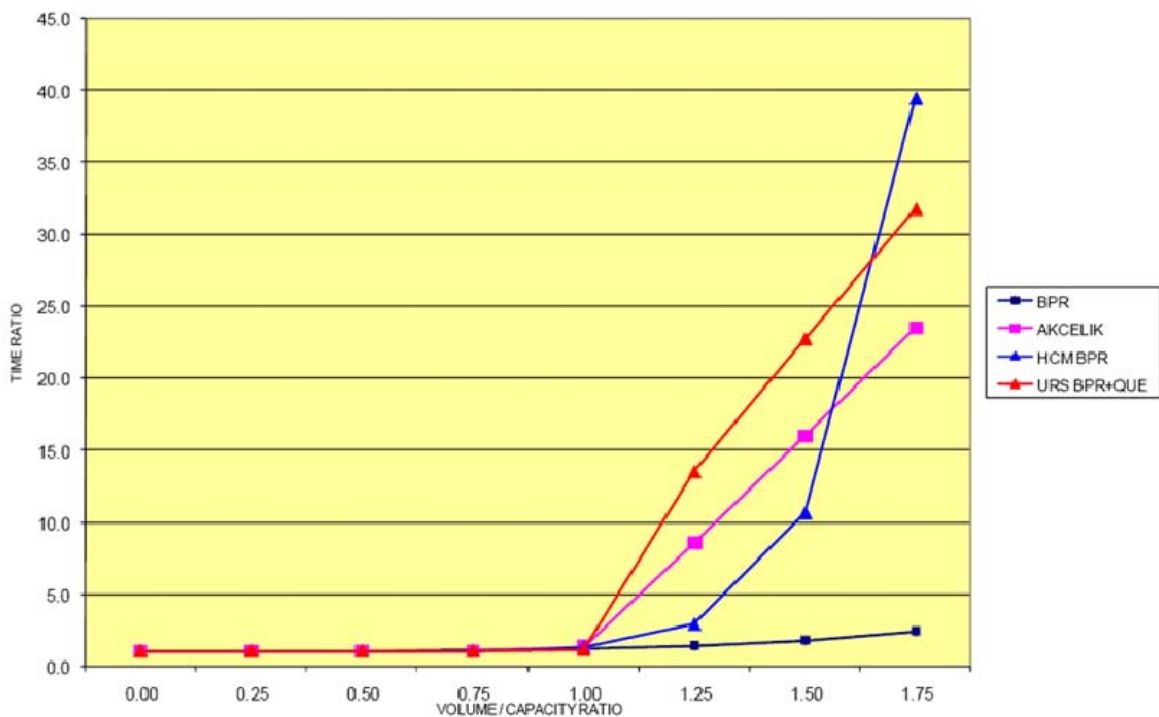
Incorporation of Delay at Nodes/Intersections

While the Akçelik VDFs appear to implicitly account for delay that occurs at intersections and other junctions, other MPOs have elected to explicitly incorporate node delays into their volume delay calculations. Agencies have used multiple methods for accomplishing this depending on the software platform used, the available data, and issues in each specific region. Agencies that explicitly incorporate node delay include the North Jersey Transportation Planning Authority (NJTPA), San Diego Association of Governments (SANDAG), Greater Buffalo-Niagara Regional Transportation Council (GBNRTC), and Portland METRO. Pima Association of Governments (PAG) did a pilot study on incorporating node delay in their VDF.

NJTPA

Primarily in response to the issue of bottlenecks in the network not accounted for by the BPR functions, NJTPA developed a new VDF that was designed specifically so that the delay was *not* proportional to the length of a link (NJTPA, 2008). NJTPA evaluated several potential VDF forms as shown for freeways in Figure 3. As can be seen, all functions produce similar estimates of travel time when V/C ratios are smaller than 1.0. But when V/C ratios are higher than 1.0, the NJTPA function (labeled as URS BPR+QUE in Figure 3) increases travel delay very rapidly. At a V/C ratio of 1.5, the ratio of congested time to free-flow time is about 22.5 for the NJTPA function in comparison with 16 from the conical function used in the TPB Version 2.2 model.

Figure 3. NJTPA Volume-Delay Function Comparison



Source: NJTPA, 2008.

The VDF adopted by NJTPA is a hybrid formula of the 2000 Highway Capacity Manual (HCM) volume-delay functions and a simplified queuing formula from the Institute of Transportation Engineers (ITE) Transportation and Traffic Engineering Handbook. The second term in the equation only applies when the V/C ratio is greater than one.

$$T_F = T_0 \left(1.0 + a \left(\frac{v}{c} \right)^b \right) + \left(\frac{120}{2} \right) \left(1 - \left(\frac{c}{v} \right) \right)$$

where:

T_F = final travel time

T_0 = initial travel time

v = volume
 c = capacity
 a, b = parameters

The coefficients in this equation vary by facility type, allowing for different delay properties on different types of roadways. The first term in this VDF is the standard BPR function. The second term in the VDF incorporates delay caused by queues at intersections, although the actual capacity at the node is not specifically calculated. Both terms are based on the V/C ratio of the link and are used in all links in the network. To facilitate the equilibrium process to converge, the model also incorporates a “speed floor” which sets a minimum initial operating speed for each link.

PAG

The Pima Association of Governments (PAG) recently conducted a pilot travel-time study on two major corridors in the Tucson, Arizona region to investigate the effects of intersection delay on vehicle travel time (Sun and Chae, 2009). The resulting VDF is a combined function that incorporates both the link delay in the first term (estimated as a standard BPR function) and the intersection delay in the second term. The intersection delay term only applies to signalized intersections and is dependent on the V/C ratio at both midblock and intersection locations, the percentage of green time allocated to the approach (g/c ratio), and the percentage of through traffic. The second term incorporates the signal delay as calculated in NCHRP 387 and a congestion adjustment based on the BPR function.

$$T_F = T_0 \left[1 + \alpha_l \left(\frac{v_l}{C_l} \right)^{\beta_l} \right] + P \left(\frac{c}{2} \right) \left(1 - \frac{g}{c} \right)^2 \left[1 + \alpha_t \left(\frac{f v_t}{C_t} \right)^{\beta_t} \right]$$

where:

T_F = final travel time
 T_0 = initial travel time
 v_l = link volume
 c_l = link capacity
 c = cycle time
 g = green time
 f = through traffic
 v_t = intersection volume
 c_t = intersection capacity

The α , β , and P are coefficients or adjustment factors determined by facility and area types. Because of its increased complexity, this model requires additional detailed information about intersection operations. Testing of this link-and-intersection based VDF produces more realistic estimates of travel time than the standard BPR function.

SANDAG and GBNRTC

The MPOs in the San Diego and Buffalo regions both use the TransCAD software platform, which includes several built-in VDFs. SANDAG and GBNRTC both use the logit-based VDF, developed by the Israel Institute of Transportation Planning and Research, which incorporates both link delay and node delay (Caliper, 2007). Again, the first term quantifies link delay based on the V/C ratio of the link. The second term follows the same functional form and calculates intersection delay based on the V/C ratio of the intersection.

$$d = t_0 * c_1 \left[\frac{1}{1 - \frac{c_2}{1 + \exp\left(c_3 - c_4 \left(\frac{v}{c_i}\right)\right)}} \right] + d_0 * p_1 \left[1 + \left(\frac{p_2}{1 + \exp\left(p_3 - p_4 \left(\frac{v}{c_i}\right)\right)} \right) \right]$$

where:

d = link delay

t_0 = free-flow link travel time

c = link capacity

v = traffic flow

d_0 = free-flow travel time of intersection

c_i = intersection capacity

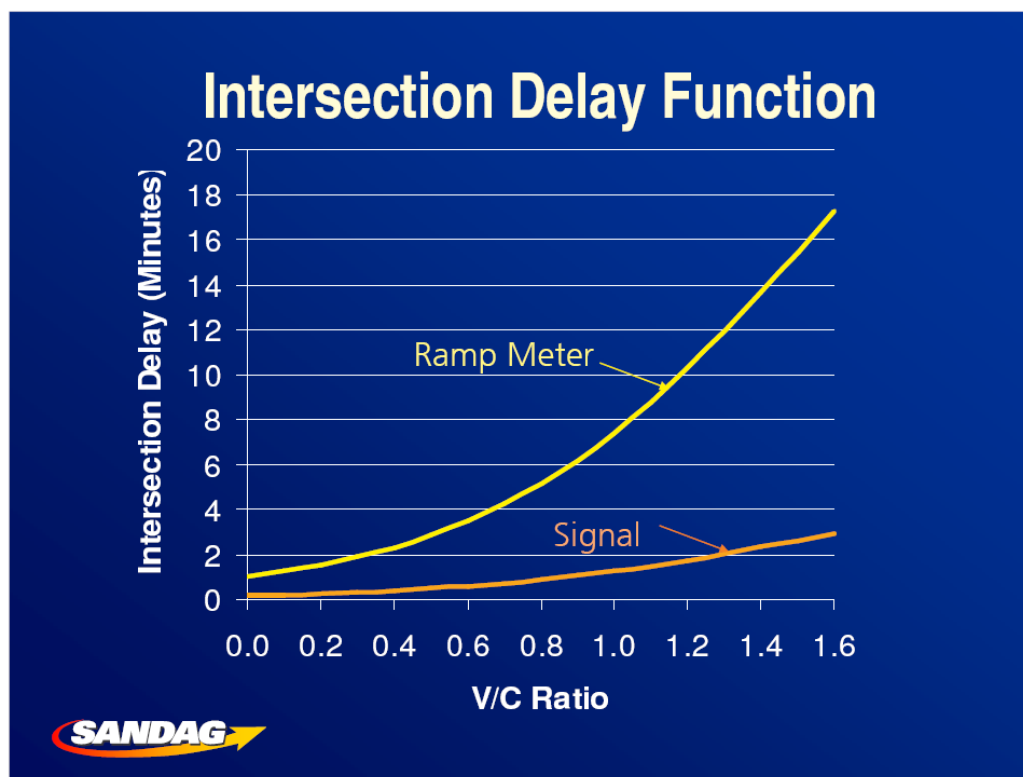
$c_1, c_2, c_3, c_4, p_1, p_2, p_3, p_4$ = parameters

While both agencies use the same VDF, implementation varies somewhat between the two agencies. The implementation of this VDF requires that a capacity be developed for all applicable intersections in the network. Both GBNRTC (Cambridge Systematics, 2007) and SANDAG (2008) include node delay for signalized and unsignalized intersections; SANDAG also calculates node delay at ramp meter locations while GBNRTC includes toll booths and roundabouts. This is in contrast to the TPB methodology which only applies node delay to freeway and ramp links.

Figure 4 illustrates the intersection delay component of the VDF used by SANDAG. Under these functions, intersection delays range from 10 seconds to more than two minutes while delays at ramp meters can range from one to more than 15 minutes.

Calculating intersection approach capacity can be dependent on a large number of variables including the number of through lanes, the signal timing, functional classifications of all streets at the intersection, and the number of turn lanes on the approach. Obtaining and maintaining this level of detail for all intersections in a metropolitan region would be very complicated and time consuming. To avoid this problem, both agencies have developed lookup tables that determine approach capacity for intersections based on simpler characteristics such as the number of approach lanes and the number of crossing lanes. In the GBNRTC model, tollbooth and roundabout capacities are calculated based on separate formulas.

Figure 4. SANDAG Intersection Delay Component



Source: SANDAG, 2004.

Portland METRO

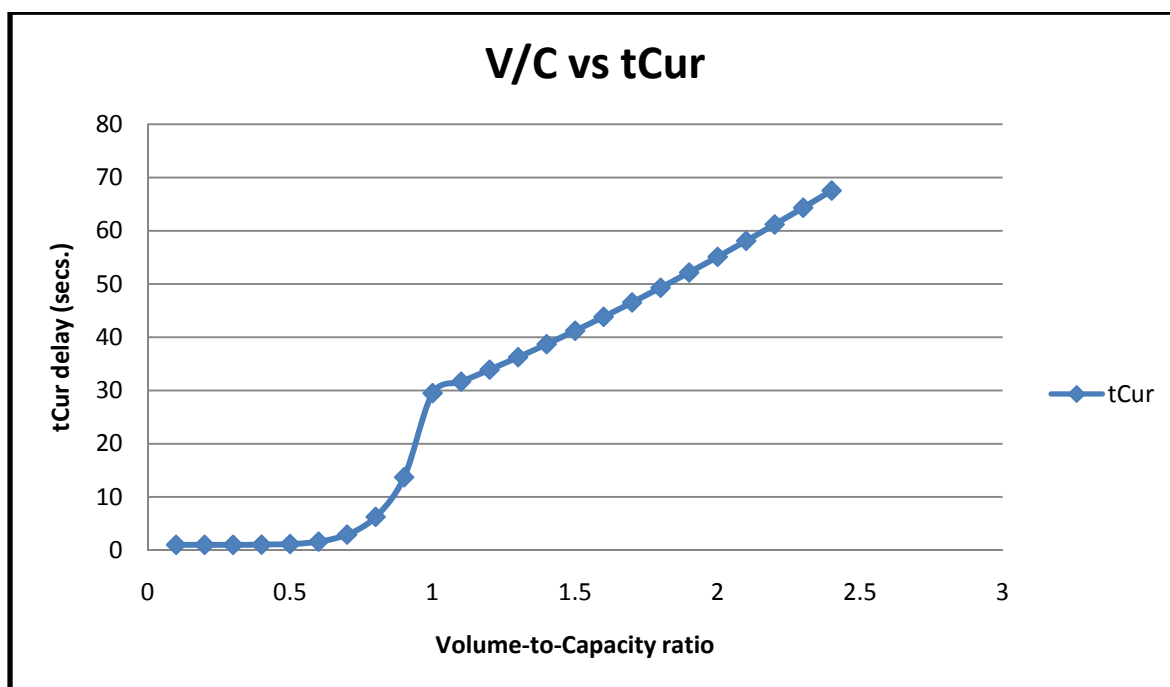
After significant experimentation and evolution with VDFs, Portland METRO's model includes what may be the most advanced and most detailed representation of queuing delay. Experimentation with conical functions proved to be insufficient for modeling speed variations along facilities, which are necessary for accurate air quality emissions forecasting. In 1998, Portland METRO developed a VDF which incorporated intersection delay based on data collected locally on different facility and area types. However, Portland METRO migrated to the VISUM platform in 2008, which incorporates a more advanced version of intersection delay based on individual movements through an intersection.

The intersection delay functionality in VISUM used by Portland METRO attempts to more accurately represent the delays caused by particular movements on the network; for example, this methodology recognizes the increased delays often associated with left turns. As with other node delay functions discussed in this memo, Portland METRO's function assigns a capacity to each turning movement at each intersection based on the speed, the number of lanes, and the facility types involved. Figure 5 shows the intersection delay function used by Portland METRO. As shown, the function includes a break at the saturation point (where volume equals capacity) of the movement. Delays on

freeways and ramps are calculated using a conical VDF as queuing does not appear to be a major concern in the Portland METRO model.

Portland METRO staff reported being very happy with their VDF, and cited the wider range of analytical options available with this tool as a major benefit to calculating delay based on turning movements. Specifically, this type of function allows the agency to easily and quickly calculate the benefits derived from implementing BRT specific improvements, such as queue jump lanes (Hauger, 2003). This would not have been possible without this type of model structure. Staff also feel that this type of model makes intuitive sense to both planners and the public.

Figure 5. Portland METRO Turn-Based Volume Delay Function



Source: Portland METRO.

3.2 TPB Queue Delay Testing

Full testing of alternative queue delay methods and model specifications is beyond the scope of this task. However, based on the procedures reviewed at other agencies, CS was able to determine several possible directions for changes in the TPB queuing delay function. For all surveyed agencies except for one, a QDF is usually applied only to non-freeway and non-ramp links. As a first step in the testing process, CS tested different applications of the current volume and queue delay functions, including applying the QDF to all facilities in the network and applying the QDF only to surface streets. The second set of tests involved removing the QDF and making changes to the VDF used.

The TPB model applied in this exercise is TPB's Version 2.2 regional travel demand model used in the air quality conformity determination for the October 21, 2009 amendment to the 2009 CLRP/FY 2010-2015 TIP (approved by the TPB on June 17, 2009). The 2005 land use inputs to this travel model are based on the Round 7.2A Cooperative Forecasts. A range of standard TPB model output tables and reports were used to analyze the results. The test scenarios include the following:

- **Base** scenario using existing conical VDF and QDF only on freeway and ramp links;
- **QDF-All** scenario using existing conical VDF and QDF on all facility types in the network;
- **QDF-Surface** scenario using existing conical VDF and QDF only on surface streets (not freeways or ramps);
- **No-QDF** scenario removing the QDF from all facility types and extending the VDF so that links with V/C ratios higher than 3.0 are penalized according to the conical function; and
- **Akçelik** scenario removing the QDF from all facility types while using a sample Akçelik curve for the VDF on all facility types. The sample Akçelik function was developed based on recommended coefficients from the Highway Capacity Manual 2000 and a functional form used by the Metropolitan Transportation Commission (MTC, 2004) in the San Francisco region.

These scenarios were not designed as standalone options, but to provide insight into the direction for further research and development of TPB's VDF. None of these scenarios, especially the Akçelik scenario, have been calibrated or validated for the TPB model, and further work on the subject would be expected. The purpose of these tests was to determine how the QDF affects travel on different facility types in the network and the comparative levels of congestion on those different facility types. CS also looked at how closely each of the scenarios matched observed volumes at the 35 TPB screenline locations.

Table 1 shows how each of the five scenarios performs at the 35 TPB screenlines. As shown, the Base scenario overestimates total traffic on the screenlines by approximately 2.4 percent. Of the remaining test scenarios, only the QDF-All scenario matches the screenline volumes closer, with an overestimation of 0.1 percent. The QDF-All scenario also shows that approximately three quarters of the screenline locations improve or stay the same in terms of match between estimated and observed volumes. The results for the QDF-Surface and the No-QDF scenarios are very similar, with a 5.5 percent overestimation of screenline volumes and improvement over the Base scenario at almost half of the screenline locations.

Of all the tested scenarios, the Akçelik scenario provides the worst match to the observed values, with an overestimation of 7.8 percent. Because the VDF used in this scenario has not been properly calibrated or validated for the TPB model region, this poor match should not be viewed as a failure of the Akçelik scenario. A properly calibrated and

validated Akçelik function would most likely have a closer match to observed values than the one used for testing purposes.

Table 1. QDF Test Scenario Screenline Performance (2005)

	Base	QDF-All	QDF-Surface	No-QDF	Akçelik
Estimated/Observed	1.024	1.001	1.055	1.055	1.078
Screenlines Better than Base		23	16	16	8
Screenlines Worse than Base		12	19	19	27

Table 2 shows the daily vehicle miles traveled (VMT) for each scenario, and the percent of daily VMT on each facility type. As shown, the addition of the QDF to all facility types slightly decreases the daily VMT in the region while each of the other scenarios increases total VMT. Despite the changes in total VMT, only small changes are observed in the composition by facility type across all of the tested scenarios. As expected, adding the QDF to the surface streets in the QDF-All scenario shifts some traffic between different categories of surface street, while including the QDF only on surface streets (QDF-Surface) shifts traffic from the larger (and likely more congested) Major Arterials back onto the freeways. The No-QDF and Akçelik scenarios also result in some shifting of traffic onto the freeways when compared to the base scenario.

Table 2. QDF Test Scenarios - VMT by Facility Type (2005)

	Base	QDF-All	QDF-Surface	No-QDF	Akçelik
Freeways (1)	39%	39%	40%	40%	40%
Major Arterial (2)	38%	37%	37%	37%	37%
Minor Arterial (3)	13%	13%	13%	13%	13%
Collector (4)	5%	6%	5%	5%	6%
Expressway (5)	4%	4%	4%	4%	4%
Ramp (6)	1%	1%	1%	1%	1%
Total	151,739,310	150,658,834	153,631,038	153,626,506	158,926,361

Another important metric used to analyze the test scenarios is the level of congestion experienced throughout the network. Table 3 and Table 4 detail the percentage of network links that experience different levels of congestion in each of the test scenarios.

The Base scenario has 4.1 percent of links operating in hyper-congested conditions (V/C ratio greater than 1.5) and 15.5 percent in over-congested conditions (V/C ratio above 1.2) during the evening peak. Moderate levels of congestion (V/C ratio between 0.6 and 1.1) are found in 41 percent of the links in the Base scenario during the evening peak.

All of the test scenarios reduce the percent of links experiencing hyper-congested conditions during the evening peak period, although two scenarios merely maintain the level of hyper-congestion in the morning peak. The QDF-All scenario, which applies an additional time penalty to congested links of all facility types, predictably shifts traffic from over- and hyper-congested links to links with moderate levels of congestion. This is also reflected in the shift between facility types, as the larger facilities such as major arterials tend to be more congested than collectors and local roads. This scenario shows the most drastic decrease in the number of hyper-congested links, a decrease of over 40 percent in the evening peak and 67 percent in the morning peak.

Table 3. QDF Test Scenarios - Percent Links by Level of Morning Peak Congestion (2005)

V/C ratio	Base	QDF-All	QDF-Surface	No-QDF	Akçelik
0.6 - 1.1	36.6%	40.9%	35.5%	35.6%	38.7%
> 1.2	4.8%	1.6%	4.9%	4.9%	4.5%
> 1.5	1.2%	0.3%	1.2%	1.2%	0.7%

Table 4. QDF Test Scenarios - Percent Links by Level of Evening Peak Congestion (2005)

V/C ratio	Base	QDF-All	QDF-Surface	No-QDF	Akçelik
0.6 - 1.0	41.0%	48.2%	39.5%	39.6%	40.0%
> 1.2	15.5%	9.1%	15.6%	15.6%	15.7%
> 1.5	4.1%	1.2%	3.8%	3.8%	2.7%

The QDF-Surface scenario shifts the QDF congestion penalty from freeways to the other facility types, thus making travel more attractive on freeways than on surface streets. This test was implemented based on CS interviews and research which indicate that the

majority of agencies that model node delay, do so only on surface streets. This practice is in keeping with the theory that capacity and therefore delay at highway merge points is determined only by the upstream and downstream lane capacities. As predicted, this scenario does increase the portion of traffic using freeways, although the number of hyper-congested links in the network decreases by seven percent in the evening peak.

The No-QDF scenario attempts to increase the delay experienced on congested links without the use of the QDF, which was not found to be common practice in the industry. The conical VDF currently used by TPB has been validated and therefore is a valuable starting point for this exercise. TPB currently uses a lookup table to represent the volume delay relationship at varying levels of congestion (V/C ratios between zero and three, in intervals of 0.1). V/C ratios higher than three do not experience any additional delay under this VDF structure. In order to increase the delay in extremely congested links simply, this test extended the lookup table by calculating delay values for V/C ratios up to five and removed the QDF entirely. This effectively allowed traffic speeds to approach much closer to zero as increasing levels of congestion occur.

The No-QDF scenario has limited effects on congested links, decreasing only the portion of hyper-congested links in the evening peak by 7.5 percent, as shown in Table 3 and Table 4. In both time periods, this scenario also shifts traffic away from links with moderate congestion levels and onto uncongested facilities.

The final scenario, Akçelik, was designed only to determine if the use of this functional structure, which places heavier delay penalties on heavily congested links, would maintain the traffic splits by facility achieved through the use of the QDF. Changes in these splits by facility type are minimal, as highlighted in Table 2, as some traffic shifts from congested major arterials to freeways and collectors. In addition, the Akçelik function shifts a substantial amount of traffic away from hyper-congested facilities (decreases of 34 percent and 42 percent in the evening and morning peaks, respectively). It is possible that a calibrated and validated Akçelik VDF could have the desired affect without the need for the queue delay function.

The results of these tests indicate that there are a number of methods that could result in the desired effect of shifting traffic away from severely congested facilities, especially freeways and ramps. As expected, the addition of the QDF to any subset of links serves to shift traffic away from those facilities. This is true for the current application of the QDF and the QDF-Surface scenario, which does not substantially ease the use of heavily congested links. The QDF-All scenario evens the field across all facility types, but adds a severe delay penalty onto links with a V/C ratio higher than 0.8. This essentially creates an additional penalty for congestion, causing traffic to shift to facilities that are less congested (and often of a lower functional class). While this scenario achieves some of TPB's goals, it does so using a rather complex and somewhat arbitrary method. The No-QDF scenario achieves approximately the same results without the need for a QDF while using a VDF that has been validated for the Washington region. The Akçelik function also shows some promise in achieving TPB's goals.

3.3 Future Considerations

Research into queuing and VDFs indicates that several options exist for further enhancement of the methodology currently used by TPB. First, the TPB model is the only one that CS encountered which applies queuing delay only to freeway links. As detailed in the GBNRTC model documentation, the capacity of freeway nodes – including merging locations – is “controlled by the capacity of the freeway segments upstream and downstream of the merge/diverge area, or the ramp itself” and that the merge point has no inherent capacity. Therefore, the QDF may not be the most accurate way to capture the desired network constraints.

Re-calibration and validation of the link-based VDF, to discourage overloading of links by inherently including additional delay on the links with the highest levels of congestion is indicated by the state of the practice and the results of the testing scenarios. Calibration would require the use of new or existing data detailing volumes and speeds for different facility types, while validation could be performed using existing traffic count data. Several options, including continued use of an expanded and/or re-calibrated conical function or switching to an Akçelik form, may have some of the desired effects. Neither of these changes would result in additional network coding requirements. It is also possible to employ different functional forms of VDFs on different facility types (i.e., freeway versus surface streets).

Should it still be advantageous to model node-related delay in the TPB model, further investigations should be made into the process, starting with the procedures detailed in this report. As already outlined, node-based delay is typically incorporated onto non-freeway links at signalized and unsignalized intersections, although TPB will have to determine how best to treat each different kind of node. As part of this type of function, the capacity of the intersection – not the capacity of the approaching link – must be determined. Developing this type of function will require significant data collection, both to estimate the amount of delay encountered at different types of intersections and to categorize each node in the TPB highway network. Maintaining this type of data as changes occur in the region will also require some additional work. The addition of queue delay only to surface streets has the potential effect of shifting traffic onto freeways, although re-estimated delay functions could help to counteract an over-compensation.

■ 4.0 Multi-Run Assignment Process

The current Version 2.2 model framework requires two model runs to be performed in order to address HOV policy and capture the impacts of HOT lanes. This process was developed to accommodate the stated policy of Virginia Department of Transportation (VDOT) that HOT facilities will not degrade the operations of HOV users. The “base run” captures the travel time for unimpeded flow of HOV traffic on HOT lanes consistent with the stated operational policy. The “conformity run” substitutes the HOV skims thus obtained for the HOV skims that would otherwise be obtained by simply skimming the

networks with HOT lanes in operation. Only the HOV skims are taken from the “base run;” skims for all other modes are taken from the “conformity run.” Under this framework, the “base run” serves solely as a means for measuring times for HOV traffic on HOT facilities. CS has proposed combining the two runs into a one-run process to save model run time and to provide more consistency in mode choice modeling.

4.1 Effects of the Multi-Run Assignment Process

The multi-run modeling process extends the run time required to execute the model. Two full assignment runs are necessary to implement this approach, and elimination of one of these runs could cut the time required to run the assignment portion of the model in half.

There is an additional benefit of enhancing the consistency in mode choice modeling, by eliminating the use of skims from different runs for different modes. As previously described, the HOV skims are obtained on a network that does not allow non-HOV vehicles to use HOT facilities. This ensures free flow traffic conditions as guaranteed by the regional HOT operational policy. The accuracy of the HOV time skims on the non-highway links (especially arterials that load onto the highways) are likely affected by the fact that fewer vehicles will be using the tolled highway paths, since the free flow HOT option is not available in the “base run.” Low occupancy vehicles (LOVs) that would have selected the path using the HOT facility are then forced to select another route; the traffic flows on arterials in the “base run” therefore may be different from those in the “conformity run” resulting in different time skims for non-freeway links being used in the final assignment for the HOV and HOT modes.

4.2 Eliminating the Multi-Run Assignment Process

In order to be able to execute the model with only one HOT facility skim, the tolls must be adjusted to achieve free flow traffic conditions that are consistent with the regional operational policy. In implementing these toll adjustments, we further recommend that the toll rate continue to be set based on link capacity rather than speed, as is done in the Version 2.2 model. Using link capacity helps with arriving at a toll which achieves the stated operational policy for the specific facility in the most straightforward and successful fashion¹. When the single-run travel time skim is performed, the path costs for the HOV users would be calculated so that tolls are not included. This change provides HOV skims that more accurately reflect traffic conditions on the arterial approaches to the HOT facilities, which could impact the HOV users’ path choice. It also serves to provide

¹ Speed is typically not represented as well in the travel demand forecast model assignment algorithm. Speed and time function as “impedance” in the assignment process and are used in an algorithm that is based on a simple decay function that lacks the sensitivity of more complex vehicle flow models.

consistent link travel times for HOV and HOT paths. This leads to a forecast which is more appropriately balanced against competing modes.

4.3 Testing Results

A test of this procedure was conducted to quantify the effects that this structural change would have on the model results. The test combined the “base” and “conformity” runs into a single process by consistently applying the HOV3 skims generated from a single-run model run. The TPB model applied in this exercise is TPB’s Version 2.2 regional travel demand model used in the air quality conformity determination for the October 21, 2009 amendment to the 2009 CLRP/FY 2010-2015 TIP (approved by the TPB on June 17, 2009). Year 2030 was selected as the model horizon year for testing because baseline conditions do not include HOT facilities.

Overall, the test performs as expected, with a shift away from HOV travel as more accurate travel time skims are used. Effects on other aggregate measures in the network are virtually unchanged, which indicates that this procedure has the desired benefit of accurately measuring HOV travel times without negatively affecting the results of the assignment procedure. Some of the findings of the test include:

- As shown in Table 5, the number of HOV auto person trips decrease by 8.5 percent due to the more realistic HOV-3 impedances generated from the single-run model in which HOT operations are included. Both LOV auto person and transit trips increase insignificantly as a result.
- The test assignment method shows that total auto person trips decrease only slightly (1,857 person trips out of 30 million or approximately 0.006 percent) due to the lower HOV market share.
- The total number of vehicles before traffic assignment is only 4,437 vehicles higher out of 27 million (or approximately 0.016 percent).
- The total daily VMT is approximately 0.019 percent lower (only 37,016 fewer) due to higher transit usage.
- The total number of vehicles crossing all screenlines is only 6,000 fewer (approximately 0.022 percent).
- The distribution of volume to capacity ratios by facility type for both the morning and evening peak periods are nearly identical between the one-run and multiple-run assignments.

Table 5. Percent Change in Number of Trips by Mode

Mode	Percentage Change in Trips
LOV	0.04%
HOV	-8.52%
Transit	0.18%

■ 5.0 Application and Scenario Manager

According to the Cube Help File, the Application and Scenario Manager “provides a convenient interface for defining scenarios, editing and running them, and reviewing input data and output data results by scenario.” In essence, the program allows modelers to run a complex, multistep modeling process with varying inputs through a simple graphical user interface (GUI). The TPB model currently uses a batch file methodology to run numerous executables and scripts repeatedly.

CS staff has used Citilabs’ Application and Scenario Manager routinely and generally feel that it serves its purpose well. In our collective experience, the graphical interface makes model application easy for users, although it can be a little tricky to set up at the beginning. Also, it can make error checking a little more time consuming since locating error messages in different files is more challenging than the current structure.

CS has used the Application Manager of Cube extensively to organize, manage, and implement models. Given that many current models, including the TPB model, are very large and complex, CS has found it is not always desirable for running complex models. The hierarchical flow chart design of the Application Manager may make it seem that it is very easy to start the program at any point in the system, but having outside executable programs can add complexity in the Application Manger environment. When using the Application Manager on a project, some CS staff have reported that the full model does not run as a group and that some steps have to be run directly from the script instead.

From CS experience, Cube’s Scenario Manager is not always a convenient interface for managing and implementing various scenarios when the model is complex and takes advantage of outside programs. The Scenario Manager’s file organization system is a convenient interface for viewing and editing data and helps to clearly visualize how the model works and contains all file editing within one program. On the other hand, it does not allow users to copy or paste, which means that some file editing is easier done outside Cube. In addition, although relationships between similar networks exist in Scenario Manager, the program does not cascade changes such that an edit in the parent network is reflected in the child network. This may be an unexpected outcome for an untrained user.

■ 6.0 Acknowledgments

We would especially like to thank all the agencies who took the time to speak with us, including:

- North Jersey Transportation Planning Authority (NJTPA);
- San Diego Association of Governments (SANDAG);
- Greater Buffalo-Niagara Regional Transportation Council (GBNRTC);
- Puget Sound Regional Council (PSRC); and
- Portland METRO.

■ 7.0 References

Akçelik, R. (1991). Travel Time Functions for Transport Planning Purposes: Davidson's Function, Its Time Dependent Form and an Alternative Travel Time Function. *Australian Road Research*, 21(3), September 1991, pg. 49-59.

Caliper (2007). *Travel Demand Modeling with TransCAD 5.0: User's Guide*.

Cambridge Systematics (2007). *GBNRTC Travel Model Refinement*.

Dowling, R. and V. Alexiadis (1997). A Blueprint for Applying EMME/2 to Ramp Metering Analyses. Presentation at 12th Annual International User's Group Conference San Francisco October 22-24, 1997

Dowling, R.G., R. Singh, and W.W.K. Cheng (1998). The Accuracy and Performance of Improved Speed-Flow Curves. Paper Presented at the Transportation Research Board Annual Meeting in Washington, D.C., January 1998.

Dowling, R. and A. Skabardonis (2006). Urban Arterial Speed-Flow Equations For Travel Demand Models. Presentation and Publication at Innovations in Travel Modeling 2006, A Transportation Research Board Conference, May 21 to 23, 2006 in Austin, Texas. Revised April 19, 2006.

Hauger, K (2003). Using Intersection-based Delay Algorithms to Determine BRT Operation Effectiveness. Presented at 9th TRB Planning Methods Applications Conference, Baton Rouge, Louisiana.

Kuppam, A.R., M. L. Outwater, M. Bradley, L. Blain, and Robert S. Tung (2005). Application of Time-of-Day Choice Models Using EMME/2, Washington State DOT (WSDOT) Congestion Relief Analysis. Presented at the 19th Annual EMME/2 Users

Conference, October 2005. <http://www.inro.ca/en/pres_pap/international/ieug05/CRA-TODModel-1018-F.pdf>

Kuppam, A. R., M. L. Outwater, and R. C. Hranac (2008). Innovative Methods for Pricing Studies. Innovations in Travel Demand Modeling: Summary of a Conference. Volume 2: Papers.

Metropolitan Transportation Commission (2004). Baycast-90 Users Guide, San Francisco Bay Area, Travel Demand Model System (Cube/Voyager version), Appendix D. August 2004.

North Jersey Transportation Planning Authority (2008). North Jersey Regional Transportation Model Enhanced: Model Development Report. Developed by URS Corporation.

Pu, W. and A. Meese (2009). Update on the Congestion Management Process. Presentation at the Travel Management Subcommittee. November 24, 2009. <<http://www.mwcog.org/uploads/committee-documents/a15aW1hd20091123142201.pdf>>

Raw J., J. Lee, P. T. Agnello, D. Munn, and K. D. Kaltenbach (2009). Investigation of Speed-Flow Relations and Estimation of Akçelik Volume Delay Functions for Travel Demand Models in Virginia. TRB Planning Applications Conference – May 19, 2009.

SANDAG (2008). 2030 Regional Growth Forecast Update: Process and Model Documentation.

SANDAG (2004). SANDAG Model Calibration with GPS Data. September 17, 2004. <http://www.srtmug.org/SANDAG_Model_Calibration_with_GPSdata.pdf>

Singh, R. and R. Dowling (1999). Improved Speed-Flow Relationships: Application to Transportation Planning Models. Proceedings of the Seventh TRB Conference on the Application of Transportation Planning Methods.

Sun, A. and K. Chae (2009). Estimating Link Travel Time with Explicitly Considering Vehicle Delay at Intersections. Presented at 12th TRB National Transportation Planning Applications Conference, Houston, Texas, May 17-21, 2009.

WSDOT (2009). Congestion Relief Analysis – Phase 2: Exploring Value Pricing in the Central Puget Sound Region. Prepared by Fehr & Peers/Mirai. <http://www.wsdot.wa.gov/NR/rdonlyres/536AC165-A8E8-4279-ADD6-B246DE4BA97A/0/CRA2ExecutiveSummary_Jan09.pdf>

Short-Term Model Enhancements: Transit-Related Enhancements

■ 1.0 Introduction

The National Capital Region Transportation Planning Board (TPB) tasked Cambridge Systematics (CS) to investigate potential transit-related enhancements of the regional travel demand model in four areas:

- Representation of the fare system in the regional travel demand model;
- Methodology for capturing fare subsidy programs in the model;
- Treatment of bus speeds in the regional travel demand model; and
- Estimation of mode choice models in the context of Federal Transit Administration (FTA) guidelines and requirements.

This technical memorandum documents the review, findings, and recommendations in the above four areas in the sections that follow. In performing this task, CS consulted documentation for adopted regional travel demand models for several of the largest metropolitan planning organizations (MPOs), including those listed in Table 1.

Table 1. Model Documentation Consulted

Agency Name	Metropolitan Region
Atlanta Regional Commission (ARC)	Atlanta, Georgia
Baltimore Metropolitan Council (BMC)	Baltimore, Maryland
Chicago Metropolitan Agency for Planning (CMAP)	Chicago, Illinois
Central Transportation Planning Staff (CTPS)	Boston, Massachusetts
Denver Regional Council of Governments (DRCOG)	Denver, Colorado
Delaware Valley Regional Planning Commission (DVRPC)	Philadelphia, Pennsylvania
East-West Gateway Council of Governments (EWGCC)	St. Louis, Missouri
Houston-Galveston Area Council (HGAC)	Houston, Texas
Hampton Roads Planning District Commission (HRPDC)	Hampton Roads, Virginia
Metropolitan Council	Minneapolis-St. Paul, Minnesota

Table 1. Model Documentation Consulted (continued)

Agency Name	Metropolitan Region
MetroPlan Orlando	Orlando, Florida
Metropolitan Transportation Commission (MTC)	San Francisco, California
North Central Texas Council of Governments (NCTCOG)	Dallas, Texas
North Jersey Transportation Planning Authority (NJTPA)	New Jersey
Northeast Ohio Areawide Coordinating Agency (NOACA)	Cleveland, Ohio
Ohio-Kentucky-Indiana Regional Council of Governments (OKI)/Miami Valley Regional Planning Commission (MVRPC)	Cincinnati; Dayton, Ohio
Portland METRO	Portland, Oregon
Puget Sound Regional Council (PSRPC)	Seattle, Washington
Regional Planning Commission	New Orleans, Louisiana
Sacramento Area Council of Governments (SACOG)	Sacramento, California
San Diego Association of Governments (SANDAG)	San Diego, California
Southern California Association of Governments (SCAG)	Los Angeles, California
Southeast Michigan Council of Governments (SEMCOG)	Detroit, Michigan
Southeastern Wisconsin Regional Planning Commission (SWRPC)	Milwaukee, Wisconsin
Southwestern Pennsylvania Commission (SPC)	Pittsburgh, Pennsylvania
Southeast Regional Planning Model (SERPM)	Ft. Lauderdale, Florida
Tampa Bay Regional Planning Model (TBRPM)	Tampa, Florida

■ 2.0 Representation of Fare Systems

CS reviewed the representation of fare systems in the existing TPB regional travel demand model with a particular focus on the fare calculation involving transit fare zones (discussed in Section 2.1). CS also reviewed the state of the practice on fare representation in regional models in a sample of the country's largest MPOs (discussed in Section 2.2). Based on these reviews and experiences elsewhere, an explicit representation of transit fares by provider and mode appears to be a preferred method for use in the regional travel demand model and should be considered as a short-term model enhancement (discussed in Section 2.3).

2.1 Fare Representation in the TPB Model

The Washington metropolitan area has a complex transit system, with a complex transit fare structure. All three major transit fare calculation methodologies exist in the Washington region:

- A distance-based fare system based on the distance traveled between boarding and alighting stations, currently used in the Metrorail system;
- A flat fare system where a boarding fare is collected for passengers on a given route, currently used by bus systems in the region; and
- A zonal fare system based on the boarding and the alighting stations, currently used by the Maryland Transit Administration/MARC system and the Virginia Rail Express (VRE) system.

Various types of discounts also are available in the region's transit system, including passes such as One Day Metrorail Pass, 7-day Rail Fast Pass, and Metrobus Weekly Pass, as well as discounts such as student fares and reduced fares for senior citizens and people with disabilities. In addition, paid parking is a feature of the heavy rail transit stations (handled outside of the fare representation).

The TPB travel demand model uses two processes, MFARE1 and MFARE2, to compute transit fares for use in the mode choice model. The MFARE1 program calculates distance-based, peak and off-peak Metrorail fares between station pairs and produces a station-to-station distance and fare matrix. The MFARE2 program is then used to calculate zone-to-zone transit fares, incorporating the Metrorail fares and bus/rail fares as appropriate. The outputs coming out of the process are transit fare matrices by time period (morning peak and off-peak), by access modes. In the draft Version 2.3 TPB model, transit fare matrices are further stratified by submode (i.e., commuter rail, Metrorail only, Metrorail/bus combination, and bus only).

The model does not include a station choice program, but there is a procedure for defining drive access to stations considering driving distances. In the Version 2.2 TPB model, drive-access to park-and-ride (PNR) lots is determined based on the airline distance between each origin zone and PNR lot, with four miles used as the cut-off threshold for core jurisdictions, five miles for the inner suburban jurisdictions, and eight miles for the outer jurisdictions. An equivalency file establishes the relationship between a PNR and a transportation analysis zone (TAZ) that represents the PNR. The PNR-TAZ equivalency file is used to then permit the method for calculating transit fares for drive access to follow the same methodology as transit fare calculation for walk access to transit. Parking facility capacity constraints are not considered in developing station demand.

In the Version 2.2 TPB model, transit fare between one TAZ and another is represented as an approximate average of transit fares for transit users taking all transit modes between the two locations. The transit fare calculation takes into account transit fares by all modes

available for the TAZ pair, walk access to transit, rail-to-bus discounts, and a deflation factor.

A system of primary and secondary transit service areas/zones is maintained, as a particular TAZ may be accessible to several modes of transit. The “bus fare matrix” represents transit fares between fare service areas/zones in the model. This fare matrix, in fact, includes fares for commuter rail service as well as for buses. When calculating the fare matrix, “the least expensive fares available are used to reflect what the majority of regular work trip commuters would pay and are averaged for areas with multiple services and fare structures” (Milone, Snead, et al., 2009). This straight-average approach may or may not lead to weighted fares that are representative of the average fares experienced.

Transit fare for each zone-to-zone pair is calculated based on whether a Metrorail trip is involved in the transit trip chain.

Non-Metrorail-related transit fare (bus-only). For non-Metrorail-related transit trips, the transit fare is the average of all fares between the origin bus fare zone (primary and secondary zones, if any) and destination fare zone (primary and secondary zones, if any). If both origin and destination TAZ are in only primary zones, the fare between them is simply the value from the fare matrix.

Metrorail-related transit fare. To calculate the transit fare for the Metrorail service portion of the trip, walk access to transit and bus transfers are taken into account as follows:

$$\begin{aligned} \text{Transit Fare} = & (\text{Bus Access Fare} \times (1.0 - \text{Origin Walk Percent})) \\ & + \text{Metrorail Fare} \\ & + \text{Bus Egress Fare} \times (1.0 - \text{Destination Walk Percent}) \end{aligned}$$

Bus access (egress) fares are the average bus-only fares between the origin (destination) fare zone and the boarding (alighting) station fare zone, diminished by one-half of the rail-to-bus discount, which varies by jurisdiction.

A deflation factor is used to convert all transit fares to be expressed in terms of 1994 dollars. The deflation factor is based on the consumer price index (CPI) from 1994 to 2008. Model years before 2009 use the actual CPI-based deflation factor, and any model years after 2008 use the 2008 deflation factor. Because of the use of the same fare structure for a future year, future transit fares must also be deflated to 1994 dollars. Likewise, auto operating cost is assumed to remain constant in the Version 2.2 TPB model with a value of 8.3 cents per mile (in 1994 dollars) in 2005 and 8.2 cents per mile (in 1994 dollars) in 2010 and in future years (Milone et al., 2010). In the draft Version 2.3 TPB model, the auto operating cost is assumed to be 10 cents per mile in 1994 dollars, constant over time.

The Washington Metropolitan Area Transit Authority (WMATA) increased Metro fares effective January 6, 2008. Version 2.2 and draft Version 2.3 of the TPB model have adopted this latest change and the revised WMATA Tariff #23 (effective June 2004) fare

structure. Historically, WMATA transit fares have not kept pace with inflation. Over the past 12 years, the WMATA base fare has risen 12 percent for bus and 23 percent for rail, while inflation has lifted prices 37 percent. Due to budget concerns, though, WMATA fares are likely on an upwards trajectory that will meet or exceed inflation in the future (witness the approximately 18 percent fare increase approved by the WMATA Board of Directors on June 24, 2010).

2.2 Fare Representation in Other Regional Travel Demand Models

CS reviewed fare representation in regional models from a selection of large MPOs in the country. This review indicates that fares are typically represented explicitly in terms of transit provider/operator and mode combinations. For example, Baltimore Metropolitan Council (BMC) represents the single-trip cash fare by transit operator and mode in the model (Allen, 2006). Similarly, Atlanta Regional Commission (ARC) assigns a separate mode for each transit provider and each mode they operate as shown in Table 2 (ARC, 2008).

Usually fares are represented in terms of cash fare, but some MPOs account for the discount available to certain groups (e.g., students, seniors) or due to pass usage (e.g., monthly or weekly passes). In these cases, the fares are usually weighted averages based on revenue composition of different types of users. For example, for the 2003 model, Southern California Association of Governments (SCAG) calculated all boarding fares as a weighted average of Year 2003 fare rates in Year 1999 dollars, considering the revenue composition of different fare types, such as monthly passes, weekly passes, senior and disabled citizen discounts, student fares, etc. These weighted fares vary by a combination of transit providers and transit modes as shown in Table 3 (SCAG, 2003).

San Diego Association of Governments (SANDAG) codes cash fare in the model, and its transit skimming procedures develop the cash fare skims based on the existing cash fare structure. In the mode choice model, these cash skims are discounted to account for pass usage, using cash fare discount factors which are computed and applied by trip mode, trip purpose, and income level. Note that calculating these discounted rates requires a survey of transit riders containing information both on pass usage and income level. SANDAG's transit survey showed that commuter rail users are more likely to use passes than local bus riders, riders making incidental nonwork trips are less likely to use passes than those making work trips, and pass usage declines with higher incomes (SANDAG, 2008).

Delaware Valley Regional Planning Commission (DVRPC) represents discounted transit fares (base fare and transfer fares) directly, using weighted average fares which are calculated using the relative usage of pass, tokens, and cash. These coded fares vary by operator and transit mode (see Table 4), reflecting the fact that different operators charge transit fares through a variety of discount passes, tokens, and cash fares. The per ride fare varies significantly depending on the type of fare instrument used and usage (DVRPC, 2008).

Table 2. ARC Transit Fare and Mode

Operator	Mode	Mode Number	Fare Coding Approach	Fare
All	Transfer	1	N/A	N/A
All	Drive to Transit	2	N/A	N/A
All	Walk to Transit	3	N/A	N/A
All	All Park and Ride Lots	4	Link Fare	Parking Fee If Applicable
N/A	Shuttle Bus	10	N/A	Free
MARTA				
	Local Bus	14	Mode Fare	\$1.75
	Heavy Rail	15	Mode Fare	\$1.75
	Express Bus	16	Mode Fare	\$1.75
	Light Rail/Bus Rapid Transit (BRT)	18	Mode Fare	\$1.75
CCT				
	Local Bus	24	Mode Fare	\$1.75
	Express Bus	26	Mode Fare	\$1.75
	Light Rail/BRT	28	Mode Fare	\$1.75
Clayton County				
	Local Bus	34	Mode Fare	\$1.75
	Express Bus	36	Mode Fare	\$1.75
	Light Rail/BRT	38	Mode Fare	\$1.75
Gwinnett County				
	Local Bus	44	Mode Fare	\$1.75
	Express Bus	46	Mode Fare	\$1.75
	Light Rail/BRT	48	Mode Fare	\$1.75
State Owned				
	Local Bus	54	Mode Fare	\$1.75
	Express Bus	56	Mode Fare	\$1.75
	Commuter Rail	57	Link Fare	TBD
	Light Rail/BRT	58	Mode Fare	\$1.75
	Intercity Rail	59	Link Fare	TBD
Greyhound	Express Bus	66	TBD	TBD

Table 2. ARC Transit Fare and Mode (continued)

Operator	Mode	Mode Number	Fare Coding Approach	Fare
Hall County				
	Local Bus	74	Mode Fare	\$1.00
	Express Bus	76	Mode Fare	N/A
	Commuter Rail	77	Link Fare	TBD
	Light Rail/BRT	78	Mode Fare	N/A
	Intercity Rail	79	Link Fare	TBD

Source: ARC, 2008, Pages 257-258.

Table 3. SCAG Transit Fare by Mode

Year 2003 Transit Boarding Fare by Mode		
Transit Mode	Description	Boarding Fare
11	MTA Local Bus	\$0.75
12	MTA Express Bus	\$0.75
13	Urban Rail (MTA Metrorail)	\$0.75
14	Los Angeles County Express Bus	\$1.03
15	Los Angeles County Local Bus (Group 1)	\$0.69
16	Los Angeles County Local Bus (Group 2)	\$0.40
17	Los Angeles County Local Bus (Group 3)	\$0.19
18	Los Angeles County Local Bus (Group 4)	\$0.00
19	All Other Local Bus	\$0.75
20	All Other Express Bus	\$0.75
22	MTA Rapid Bus	\$0.75

Source: SCAG, 2003, Page 46.

Note: Transit boarding fares are in 1999 constant dollars.

Table 4. DVRPC Transit Fare by Mode

Operating Company	2000 Base Fare (€)	2000 Transfer Charge (€)	2000 Zone Increment (€)^a
SEPTA			
City Division	132	26	-
Victory Division	143	38	20
Frontier Division	162	32	20
Regional High Speed Rail	205	205	Varies with distance
NJ TRANSIT			
Mercer Division	110	45	20
Southern Division	110	45	20
NJT Rail	110	-	Varies with distance
PATCO	100	**	Varies with distance
Pottstown Urban Transit	150	35	Varies with distance
<hr/>			
Operating Company	2005 Base Fare (€)	2005 Transfer Charge (€)	2005 Zone Increment (€)^a
SEPTA			
City Division	148	29	-
Victory Division	160	31	23
Frontier Division	181	36	23
Regional High Speed Rail	230	230	Varies with distance
NJ TRANSIT			
Mercer Division	121	54	22
Southern Division	121	54	22
NJT Rail	121	-	Varies with distance
PATCO	153	**	Varies with distance
Pottstown Urban Transit	150	35	Varies with distance

Source: DVRPC, 2008.

^a All coded fares are averaged over pass, token, and cash fare rates using relative usage.

** SEPTA City Division has special PATCO transfers.

2.3 Recommendations

Several factors should be taken into consideration when developing a fare structure for the regional model:

- Accuracy of representing costs incurred by transit users;
- Intended uses of the model for policy analysis, e.g., whether the model will be used to analyze the effects of fare hike or fare-free zones, or changes in the fare payment system;
- Whether existing transit surveys can provide the required level of detail;
- Flexibility for future fare structure changes associated with new modes and/or transit providers;
- Ease of coding and summarization of model results by mode and by transit provider; and
- Software capability and flexibility in accommodating the transit fare system.

Based on these considerations and experiences elsewhere, an explicit representation of transit fares by provider and mode appears to be a preferred method for use in the TPB regional travel demand model. Coding transit fares explicitly by a combination of transit providers and modes, enables this approach to accurately represent the complex transit fare structures in the region. Because of this explicit coding, the model can accommodate providers with different fares for different transit services. This method also allows for potential expansions in transit modes and providers in the future. Importantly, it also offers an easy way to test the effects of fare policies on transit usage and ridership, particularly for different providers and/or modes.

The Cube modeling software can accommodate a wide range of fare coding approaches, although clearly the Washington region has a very complex transit fare structure. Cube TP+ can accommodate up to 255 “modes” in the model stream. Using a two-part mode code can thus permit up to 25 different operators (first two digits of the code) using up to 10 modes (last digit of the code). While the TRNBUILD transit module relies on the mode numbers themselves, Cube Voyager can further use the OPERATOR feature to differentiate transit fares in its Public Transport (PT) transit module. It is recommended that TPB consider incorporating the revised treatment of transit fares among its short-term model enhancements.

■ 3.0 Representation of Fare Subsidies

TPB staff has undertaken research on how to incorporate employer subsidies into the transit fare inputs for the model. This effort was prompted in part by findings from the 2007 Metrorail passenger survey that showed the proportion of subsidized commuting trips on the Metrorail system is 60 percent (Milone, Humeida, et al., 2009). Although participation varies by employer and the exact benefits may vary widely, the fact that the SmartBenefits program can, by law, potentially represent a subsidy of up to \$230 per month as of March 1, 2009, indicates its potential importance in mode choice.

TPB's research derived the percentage of Metrorail work trip attractions that are subsidized at each Metrorail station based on the 2007 Metrorail passenger survey. Then "the monthly subsidy [was translated] into a reduced or discounted average daily transit fare between station pairs" (Milone, Humeida, et al., 2009). The subsidy was assessed for the Metrorail trip only due to limitations in the available data from the 2007 survey. The final station-to-station fares are simply the weighted average of regular fares and discounted fares, with the weight defined as the percent of Metrorail work trip attractions that are subsidized at each Metrorail station. The maximum allowable monthly benefits are assumed to be available for commuters to use. This amount can vary each year based on authorizing legislation¹.

The scan of regional travel demand models indicates that no large MPO has consideration of transit fare subsidies built into their travel demand model, though a few incorporate fare-free zones. Also, some agencies include treatment of discounted fare media as part of the fare system representation.

Several issues appear to complicate incorporating the researched transit subsidy approach directly into the production version of the TPB travel demand model, including:

- Application of fare subsidies solely to the Metrorail trips may potentially bias the forecast of trips in other transit modes such as bus and commuter rail, as illustrated in the model tests performed by TPB. Applying fare subsidies only for Metrorail trips increased Metrorail and Metrorail-related trips and decreased bus and commuter rail trips.
- Employers provide varying subsidy levels for transit. The currently proposed method assumes the maximum benefit is provided, but the prevalence of this policy outside of the Federal government needs to be verified.
- It is not easy to implement a similar method for bus-related trips. Although the regional bus survey may provide some insights on transit subsidies provided to bus

¹ The maximum potential monthly benefit has increased in a series of uneven steps from \$65 in 2000 to \$230 as of March 1, 2009.

riders, it would be difficult to derive a stop or TAZ-based percent of work trip attractions that are subsidized.

- Station-based subsidies of work trip attractions are tied to the employers close to individual stations, which may change locations in the future. While many employers and government agencies are relatively stable in their office locations, office locations may be changed over a typical planning horizon. Recent Base Closure and Realignment (BRAC)-related relocations are an example. It is unclear whether an office would move to increase their relative transit accessibility or, conversely, whether they would drop the SmartBenefits program once established, even if they move further from a station. It is uncertain that 20-year forecasts of the percentages of subsidized work trip attractions by station will be valid, yet this is required for forecasting purposes. While the SmartBenefits program is still growing in popularity, it is not clear whether a linear or exponential growth in program take-up should be considered for future years, whereas fixing these percentages to base-year conditions also is problematic.

Recognizing that, given their prevalence, fare subsidies do seem to have a significant influence on travel behavior in the Washington, D.C. region, consideration of other methods of incorporating the fare subsidy which could be applied to all transit modes is encouraged. In doing so, it may be useful to consider the application of transit subsidies in the context of other important employer transportation subsidies, such as for parking costs. Taking transit and parking subsidies into account in a similar manner in the regional travel demand model may reflect the out-of-pocket travel cost of the different travel options. The TPB model uses a parking cost model to estimate average zonal parking cost based on zonal employment density. The estimated parking costs in the model currently reflect subsidized parking costs.

A few possible options exist for continuing the exploration of representing the transit fare subsidies in the regional model. One possible avenue to explore would be to use recent travel surveys to look at fare subsidy availability/use for non-Metrorail riders and pursue an expanded treatment of transit subsidies along the lines already considered for Metrorail, but applied to all modes. Although this approach would not address all of the limitations identified in the research approach, it would potentially eliminate some of the modal bias.

Another avenue is to explore the relationship between fare subsidy availability/use and the type of employment in the attraction TAZ. If this relationship can be established, it can be used to forecast the potential effects of employment relocation and redistribution in the future. The current land use forecasts for the TPB model do include employment by different categories, such as retail, office, industrial, and other. As an example, it can be hypothesized that office employment is more likely to offer transit subsidies than retail employment. If this relationship can be tested and confirmed with the survey data, it could be applied on a regionwide basis.

Similarly, an exploration using the same dimensions used to reflect parking cost subsidies might yield a set of more parallel approaches to addressing subsidies of both parking and

transit. For example, the parking subsidy could potentially be represented explicitly as a separate calculation from the parking cost and the same underlying look-up methods used to apply parking subsidies and fare subsidies.

A third avenue for exploration is more general distributions of fare subsidies on a geographic basis (e.g., district-to-district or county-to-county) which might be more stable over the planning horizon. This type of aggregate approach would mirror the approach used by other MPOs to represent the use of discounted fare media.

■ 4.0 Representation of Bus Speeds

CS reviewed the treatment of bus speed in the TPB regional travel demand model, which is based on bus run time and degradation factors for future years to account for increasing traffic congestion. CS also reviewed the state of the practice on bus speed treatment in regional travel demand models in the country's largest MPOs. These methods were synthesized into three categories, and strengths and weaknesses of each method are discussed in the subsections that follow.

4.1 Current Bus Speed Treatment in the TPB Model

The current TPB model uses fixed end-to-end run times for buses for the base-year transit time, based on bus schedule times from transit agencies. Local bus travel times for future years are reduced by fixed degradation factors, which vary by time of day (peak versus off-peak), by types of buses (WMATA and primary local buses versus secondary local buses), and by year. Express bus travel times are the same as the schedule-based coded run time.

4.2 State of the Practice for Bus Speed Treatment in Regional Models

The review of regional travel demand models used by large MPOs shows that it is a state of the practice technique to estimate the travel time of transit modes operating in mixed traffic as a function of congested highway time. All the MPO models reviewed make the linkage between transit travel time and highway travel time in one way or another. It also is a state of the practice technique to directly code transit travel time or speed for transit modes operating on exclusive right-of-way transit lines (such as fixed guideway and bus lanes) independently.

There are some variations on how the relationships between highway travel time and mixed flow transit travel time are represented in the MPO models. Overall, they can be categorized into three groups: bus speed curves, a regression model, and highway time/speed with bus delay, depending on how various elements of transit time are explicitly or implicitly represented, including:

- Auto travel speed/time on roadway network;
- Acceleration/deceleration of transit vehicles;
- Dwell time at stops/stations; and
- Recovery time at the end of each trip.

Bus Speed Curves

In this method, bus speed curves are developed to relate bus speed with auto speed on highways, generally by facility types, area types, and perhaps submodes. Bus speed curves are piecewise linear and are defined by three points. The first point is zero miles per hour for autos and zero miles per hour for buses. A second point defines an intermediate break in the auto/bus speed relation; a third point defines the maximum travel speed for auto and for bus. Between each pair of points the auto/bus speed relationship is linear.

In 1997, CS implemented transit speed curves based on facility types and area types for the DVRPC model in the Philadelphia region. These curves established relationships between bus speeds and auto speeds. This procedure was later adopted for the Southeast Michigan Council of Governments (SEMCOG) travel demand model for the Detroit area. Table 5 identifies which curve to use for a particular area type and facility type combination. Table 6 shows the auto and bus speeds for the three points, which define each curve.

Table 5. SEMCOG Auto/Bus Speed Relationship Curves

	Area Type 1 (CBD)	Area Type 2 (Fringe)	Area Type 3 (Urban)	Area Type 4 (Suburban)	Area Type 5 (Rural)
Facility Type 1 (Freeway)	Curve 2	Curve 2	Curve 2	Curve 2	Curve 2
Facility Type 2 (Expressway)	Curve 2	Curve 2	Curve 2	Curve 2	Curve 2
Facility Type 3 (Principal Arterial)	Curve 3	Curve 3	Curve 5	Curve 8	Curve 10
Facility Type 4 (Minor Arterial)	Curve 3	Curve 3	Curve 5	Curve 8	Curve 10
Facility Type 5 (Collector)	Curve 4	Curve 4	Curve 7	Curve 8	Curve 10

Source: Cambridge Systematics, 2002, Page 3-22.

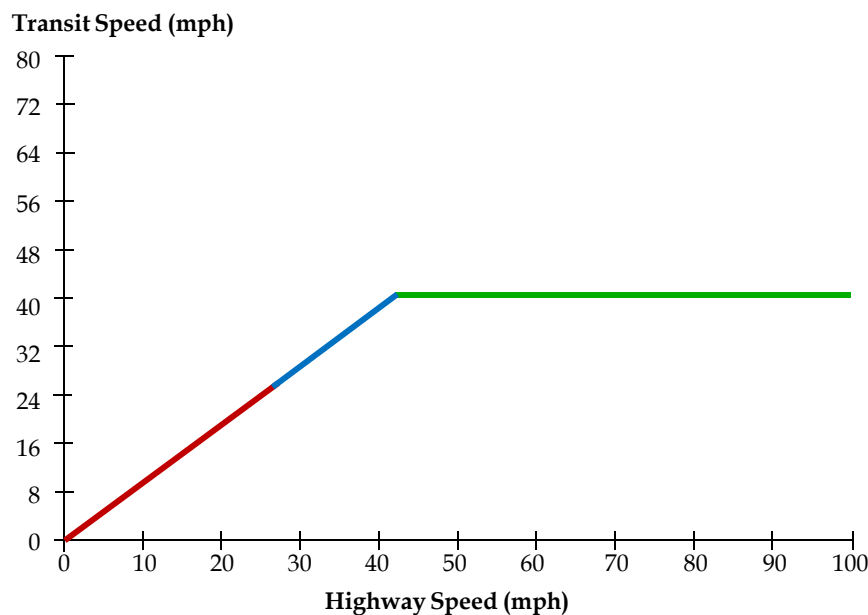
Table 6. SEMCOG Auto/Bus Speed Definition Points (Miles per Hour)

Curve	Point 1		Point 2		Point 3	
	Highway	Transit	Highway	Transit	Highway	Transit
Curve 2	0	0	30	30	70	65
Curve 3	0	0	12	5	36	15
Curve 5	0	0	18	7	40	15
Curve 7	0	0	12	6	30	12
Curve 8	0	0	18	11	45	20
Curve 10	0	0	24	15	50	24

Source: Cambridge Systematics, 2002, Page 3-22.

This procedure also is widely used in regional models in Florida, and is adopted in the Florida Standard Urban Transportation Model Structure (FSUTMS). A series of speed curves may be used for different combinations of modes, area types, and facility types. Different time periods may also be considered, such as peak-period transit speed based on peak-period auto speed and off-peak transit speed based on off-peak or free-flow auto speed. In developing the speed curves, dwell time, stop density, and acceleration and deceleration rates were considered. Figure 1 shows an example curve from the Tampa Bay Regional Planning Model.

Figure 1. Tampa Bay Bus Speed Curve



Source: Gannett Fleming, 2007.

Strengths of the bus speed curve method include:

- A direct linkage is established between transit speed and highway speed, along with implicit incorporation of stop density, dwell time, and acceleration and deceleration; and
- Bus speed is linked to area types and facility types.

Potential limitations of this method include:

- It does not usually account for delays at bus stops due to passenger loading and alightings;
- It is time-consuming to calibrate transit travel times at the route level, and indeed this is rarely done; usually these curves are calibrated at a much more aggregate level based on operator, service type (e.g. local versus express), and occasionally sector-to-sector movement information; and
- A bus route may traverse multiple area types and facility types.

Regression Model

In this method, an empirical relationship between bus speed and congested highway speed is developed using a regression model, having the following form:

$$\text{Bus Speed} = a \times (\text{congested highway speed}) + b$$

Where both a and b are parameters resulting from the model estimation. Usually, bus schedules from transit agencies are used to estimate and calibrate a and b parameter values, which can differ by area type and facility type. Congested highway speed data usually come from the model output. In this method, delays from stopping, deceleration, and acceleration are implicitly represented in the model.

Table 7 shows calibrated parameter values for parameter a by facility type and area type, while Table 8 show values for parameter b , adopted in the regional travel demand model of the Atlanta Regional Commission (2008). It is not necessary to use a single set of parameters; it also is possible to calculate separate sets for express and local bus routes. Alternatively, factors may only be calculated for bus routes using highway facilities (typically express routes), as local bus routes on collectors and local streets are very unlikely to see a benefit from the types of system capacity improvements under study in a regional model. Furthermore, these local routes are almost entirely dominated by bus stop dwell time, which is how these routes are often calibrated.

Table 7. Calibrated Bus Speed Parameter “a”

Area Type	Centroid Connectors	Freeway	Principal Arterial 1	Principal Arterial 2	Minor Arterials	Collector
1	1.0000	0.9000	0.3605	0.5376	0.5923	0.7379
2	1.0000	0.9000	0.4405	0.4668	0.5385	0.5290
3	1.0000	0.9000	0.4085	0.4569	0.4899	0.5067
4	1.0000	0.9000	0.4290	0.4807	0.5356	0.5449
5	1.0000	0.7000	0.5861	0.6116	0.6267	0.6783
6	1.0000	0.9000	0.5477	0.6030	0.6085	0.6287
7	1.0000	0.9000	0.5385	0.6697	0.7105	0.6926

Source: ARC, 2008.

Table 8. Calibrated Bus Speed Parameter “b”

Area Type	Centroid Connectors	Freeway	Principal Arterial 1	Principal Arterial 2	Minor Arterials	Collector
1	0.0000	0.0000	5.9993	2.6072	2.0696	0.6383
2	0.0000	0.0000	5.6740	5.1127	3.5933	3.7366
3	0.0000	0.0000	6.5819	5.5321	4.9054	4.3896
4	0.0000	0.0000	5.9857	4.8842	3.8184	3.4951
5	0.0000	0.0000	4.6927	4.1883	3.7634	2.7144
6	0.0000	0.0000	5.7155	4.4121	4.2765	3.6284
7	0.0000	0.0000	9.4463	4.4402	3.1615	3.6658

Source: ARC, 2008.

Sacramento Area Council of Governments (SACOG) did a regression analysis of highway times to the observed (scheduled) bus speeds for peak and off-peak service, respectively, which yielded time factors of 1.86 for the peak period and 1.89 for the off-peak. In this regression analysis, bus lines with non-stop, limited stop, or freeway portions were excluded, and the intercept was constrained to zero.

Strengths of this regression model methodology include:

- It is easy to estimate and calibrate the model parameters if accurate bus run times are available; and

- A direct linkage is made between transit speed and congested highway speed.

Potential weaknesses of this method include:

- Scheduled run time may not accurately represent actual transit run time; and
- Congested highway speed values from the model are assumed to be an accurate representation of the actual congested highway speeds on the roadway network. Although the base-year model is usually calibrated and validated at the regional level, congested speeds at a link level are seldom calibrated. In a congested metropolitan area it is possible that the model highway assignment results in some links have unusually low congested speeds, resulting in unrealistically low transit speeds for routes utilizing those particular links.

Highway Speed/Time and Bus Delay

In addition to congested highway time, some MPOs use delayed time due to bus operation to estimate bus travel time. For example, SANDAG (2008) determines bus travel time as a function of highway travel time and bus delay time due to stops:

$$bt_{tm} = ht_{tm} + bs \times dt_m$$

where:

bt = bus travel time on link during time period tm

ht = congested highway travel time

bs = number of bus stops on link

dt = per stop delay time by mode m

Stop delay times by modes are assumed to be 40 seconds, 30 seconds, and 18 seconds per stop for bus rapid transit (BRT), express bus, and local bus service, respectively. Express and local bus stop delays were calculated from observed data and include the effects of acceleration/deceleration, dwell time for boarding passengers, and likelihood of stopping at an individual stop. Longer delay times are assumed for BRT service since fixed-stop operation is proposed for these routes.

Highway travel times are modified for the following special conditions before computing bus times:

- Freeway high-occupancy vehicles (HOV) lane speeds are assumed to be no lower than a level-of-service “D” speed of 62 miles per hour (mph);
- Ramp meter delays at meters with HOV bypass ramps are assumed to be one-third of single occupant vehicle times;

- The maximum legal speed limit is used for the free-flow bus speed on freeways, whereas highway free-flow freeway speeds are set at five mph above the speed limit to reflect observed speeds from survey data;
- Bus speeds of 35 mph are assumed on selected freeways that allow buses to run on shoulder lanes when speeds on adjacent general lanes fall below 35 mph; and
- Travel times on arterial streets used by BRTs are reduced by 10 percent to reflect the effects of bus priority treatments.

The NCTCOG (2000) model methodology treats transit speed/time as a function of highway speed/time and stop delays. Way-type codes of the link segments are used to decide whether to use free-flow speeds or estimated loaded speeds:

- Way-type = 1 (Mixed-flow traffic; use estimated loaded speeds)
 2 (Reserved, Contraflow, or HOV lanes; use free speeds)
 3 (Exclusive guideway; use free speeds)

These non-stop speeds are then reduced by stop delays due to deceleration, dwell time, and acceleration:

$$\text{Delay} = \text{number of stops} \times (\text{dwell time per stop} + \text{acceleration and deceleration time per stop})$$

The number of stops per link is estimated by applying a stop-density factor to the link length. Stop density factors are defined as the number of stops per mile and are stratified by time-of-day, area type, and line haul mode (see Table 9). Dwell times are defined as passenger loading and unloading time, and also stratified by time-of-day, area type, and line-haul mode (see Table 10). Acceleration and deceleration rates vary by technology codes (see Table 11).

Table 9. Stop Density Factors (Stops per Mile)

Peak Period	Mode			
	Local Feeder Bus (Fort Worth)	Local Feeder Bus (Dallas)	Express Bus	Rail
Central Business District	8	8	8	5
Outer Business Districts	7	6	5	3
Urban Residential	5	4	3	2
Suburban Residential	4	3	2	1
Rural	2	2	1	1

Table 9. Stop Density Factors (Stops per Mile) (continued)

Off-Peak Period	Mode			
	Local Feeder Bus (Fort Worth)	Local Feeder Bus (Dallas)	Express Bus	Rail
Central Business District	8	8	6	3
Outer Business Districts	5	4	3	2
Urban Residential	4	3	2	1
Suburban Residential	3	2	1	1
Rural	2	2	1	1

Source: NCTCOG, 2000.

Table 10. Dwell Time per Stop (Seconds)

Peak Period	Mode			
	Local Feeder Bus (Fort Worth)	Local Feeder Bus (Dallas)	Express Bus	Rail
Central Business District	17	17	20	20
Outer Business Districts	12	12	12	20
Urban Residential	10	10	12	20
Suburban Residential	10	10	20	20
Rural	10	10	20	20

Off-Peak Period	Mode			
	Local Feeder Bus (Fort Worth)	Local Feeder Bus (Dallas)	Express Bus	Rail
Central Business District	15	15	20	20
Outer Business Districts	10	10	12	20
Urban Residential	8	8	12	20
Suburban Residential	8	8	20	20
Rural	8	8	20	20

Source: NCTCOG, 2000.

Table 11. Acceleration/Deceleration Rates (Miles per Hour per Second)

	Technology Code				
	1	2	3	4	5
ACC	2.50	2.50	3.10	3.30	3.00
DEC	2.50	2.50	3.10	3.30	3.00

Source: NCTCOG, 2000.

Note: Technology Code 1 = Diesel bus, 2 = Gasoline Bus, 3 = Light Rail Transit - Electrical, 4 = Heavy Rail Transit-Electrical, and 5 = Heavy Rail Transit-Diesel.

Strengths of this method include:

- Transit time/speed are explicitly represented by individual factors that affect transit time/speed, which are easy to understand and are responsive to changes in operational characteristics of bus systems; and
- It is relatively easy to implement in the modeling process, and calibration can be done by adjusting individual factors.

Potential weaknesses of this method include:

- Delay does not explicitly reflect the impact of transit demand, such as the amount of boardings and alightings activities which vary widely by lines.

4.3 Recommendations

While the current TPB bus travel time approach was designed to represent increasing congestion in the Washington region, it does not provide a direct linkage between bus travel speeds and the level of roadway congestion. In addition, since congestion varies in different parts of the region, differential impacts on transit travel time would be expected. That is, some routes might be affected more by increasing congestion than others. TPB may wish to now consider establishing a tighter linkage between bus speeds and highway congestion. An enhanced connection would permit significant improvements to regional highway capacity to accrue benefits to transit riders as well as auto drivers and reduce the potential for overstatement of the shifts to auto modes in future model runs.

It is recommended that TPB consider establishing an explicit relationship between bus speed and highway speed, along with bus delay. It can take the following form:

$$Time_{bus} = Time_{auto} + Bus\ Delay$$

Highway travel time reflects traffic congestion on the highway system. Bus delay captures all sorts of delays caused by bus operations, including dwell time, acceleration and

deceleration, loading time, and recovery time. Bus delay can be formulated as a delay factor multiplied by the number of stops (alternatively, link length). The delay factor can be empirically estimated using actual or scheduled bus run times and best estimates for auto travel time on the highway system. The relationship should be established by bus submodes, peak and off-peak periods, area types, and facility types.

■ 5.0 Mode Choice Model Development

Mode choice models have received a lot of attention in recent years because of FTA requirements for New Starts applications. In the regional model update, a key issue is how best to develop a mode choice model. CS reviewed FTA guidance and the current MPO practice on the methods used to update mode choice models. Motivations and issues in each method are discussed below.

5.1 Model Development Approaches

Three major approaches have been employed to develop mode choice models during the latest mode choice model updates in large MPOs. The following sections summarize MPO experiences adopting the “estimation” approach, the “assertion” approach, and the hybrid approach.

“Estimation” Approach

Several MPOs relied on the “estimation” approach to develop their mode choice models, especially in a multinomial logit model framework. In the 2008 trip-based demand model used in the Portland-Vancouver metropolitan area, METRO updated mode choice models, which were formulated in a multinomial logit framework. Central Transportation Planning Staff (2009) developed multinomial logit mode choice models by trip purpose for the Boston regional travel demand model based on the 1991 household travel survey data, travel impedances from highway and transit networks, 1990 and 2000 Census data, and other data sources.

Nested logit models were estimated to develop the mode choice models of the Metropolitan Transportation Commission for the San Francisco Bay Area (Purvis, 1997). Six trip purposes were modeled in the nested logit structure: home-based work, home-based shop/other, home-based social/recreation, non-home-based trips, home-based school/high school, and home-based school/college. Non-motorized modes are part of the nesting structure. Independent variables include tripmaker demographics (auto ownership, income, household size, workers in the household); trip characteristics (travel time and trip cost); and neighborhood characteristics. Estimated coefficients were evaluated in terms of the value of time by trip purpose.

“Assertion” Approach

Several large MPOs take an “assertion” approach to mode choice model development, either by using coefficient values reflecting FTA guidance or borrowing coefficients from another region that went through the New Starts process with FTA. Some of the regional models taking this approach include Baltimore Metropolitan Council (BMC), Central Florida Regional Planning Model (CFRPMv41) covering Orlando, Southeast Regional Planning Model VI covering Miami, Puget Sound Regional Council covering Seattle, Regional Planning Council covering New Orleans, and the Tampa Bay Regional Planning Model.

Some of the MPO mode choice models were estimated from survey data prior to the latest model update, but MPO staff decided to replace these models in favor of the assertion approach. BMC undertook a major model update on the mode choice model around 2000, using the 1993 household travel survey and testing a variety of model specifications including multinomial and nested logit models. During the process of the Red Line New Starts application, FTA raised some concerns about the mode choice model. Subsequent review of the model indicated inconsistencies with FTA requirements, including use of mode-specific in-vehicle travel time (IVTT) coefficients, use of thresholds or “cliffs,” and overspecified alternative-specific constants. In the 2006 model update, it was decided that the mode choice structure and parameters would be borrowed from the New Orleans mode choice model, which was believed to be acceptable to FTA.

In the Seattle model update for congestion relief analysis, mode choice model parameters were developed based on FTA guidelines and 2005 documentation of mode choice model parameters from 26 urban areas produced for the U.S. Environmental Protection Agency (EPA) (Cambridge Systematics, 2007). The 1999 PSRC household travel survey mode shares for each market segment were also used for informing the selection of mode choice model parameters for market segmentation.

Hybrid Approach

The hybrid approach to developing a mode choice model combines the estimation and assertion methods. Mode choice models are still estimated with statistical software using locally-collected household survey data, but values for some coefficients are “asserted” or “constrained” in the model estimation process. Constrained coefficients may include time, cost, and nesting coefficients. For example, in the initial mode choice model estimation of the SCAG model in July 2005, coefficients for level of service variables and nesting coefficients were outside FTA ranges. It was later decided to constrain some model parameters such as walk access time, wait time, transfer time, and HOV time savings. A set of nested logit models was estimated with additional variables such as distance for walk and bike, ratio of vehicles to persons in household, travel cost, and Central Business District (CBD) dummy, along with the set of constrained and nesting coefficients.

In the development of the NCTCOG mode choice model, it was found necessary to constrain the transit fare coefficient and auto-cost coefficient for the home-based work model. In-vehicle time coefficients (for auto and transit) were constrained for the home-based

non-work model, and both cost and in-vehicle travel time coefficients were constrained for the non-home-based trip model. A number of nesting structures and model specifications were tested.

CS reviewed mode choice models from a sample of the largest MPOs in the country. This review showed that many MPOs have asserted coefficients or have used a hybrid approach of combining statistically estimated coefficients with asserted coefficients based on rules of thumb. Out of the 25 regional models reviewed, approximately half of the mode choice models were estimated, and the remaining half took either the assertion approach or the hybrid approach.

5.2 FTA Requirements and Mode Choice Model Development Issues

The motivation for using a non-estimation approach appears to reflect FTA requirements for New Starts, data limitations, and difficulties associated with mode choice estimation. FTA provides procedural and technical guidance on travel forecasting for New Starts proposals and has conducted several workshops on Travel Forecasting for New Starts Proposals over the past years.

FTA has identified a series of issues related to travel forecasting in support of New Starts applications. These issues were largely identified during reviews of ridership forecasting for past and newly proposed transit projects. The latest assessment conducted in 2007 still indicates “systematic overestimation of ridership,” with average actual ridership representing only 74.5 percent of forecast (Lewis-Workman, 2008). This 2007 study suggests that more recent forecasting is not significantly more accurate than in the projects examined in a prior study performed in 2003, although both studies show much better forecasting than in projects evaluated in a similar 1990 study.

Over the years, FTA has identified a number of “problematic characteristics of transit forecasting methods” (FTA, 2004), including the following:

- Unusual coefficients;
- Bizarre alternative-specific constants;
- Non-logit decision rules;
- Problems in choice-set formation;
- Transit path-builder inconsistencies with mode choice models;
- Inaccuracy of bus running times; and
- Instability of highway assignment results.

FTA requires **compelling evidence** if the in-vehicle travel time (IVTT) or out-of-vehicle travel time (OVTT) utility expression coefficients in the mode choice model (C_{ivtt} and C_{ovtt} , respectively) are outside a certain range:

- $-0.03 < C_{ivtt} < -0.02$; and
- $2.0 < (C_{ovtt}/C_{ivtt}) < 3.0$.

FTA also requires **compelling evidence** if mode-specific IVTT coefficients are used instead of “generic” IVTT coefficients for all modes. If mode-specific IVTT coefficients are used, FTA requires **compelling evidence** if the relative magnitudes of mode-specific IVTT coefficients do not follow appropriate relationships:

- C_{ivtt} for transit less negative than C_{ivtt} for automobile.
- C_{ivtt} for commuter rail less negative than C_{ivtt} for transit.

It is not clear what constitutes compelling evidence, but simply stating that the “unusual” coefficients were estimated from the observed data is not satisfactory to FTA.

FTA identified a wide variation of model coefficients such as $-0.045 < C_{ivtt} < -0.007$, $0.25 < (C_{ovtt}/C_{ivtt}) < 16.0$, and expressed concern that these unusual coefficients may not reflect real travel behavior and may be due to estimation errors or distortion.

A wide range of variations in some mode choice coefficients was also documented from different reviews of mode choice models across the country. NCHRP Report 365 summarizes an early review of home-based work (HBW) mode choice models around the country, most of which were multinomial logit models developed before 1990 (Martin and McGuckin, 1998). Out of 13 HBW mode choice models reviewed across the country, the IVTT coefficients range from -0.04 to -0.015, with 10 being -0.034 to -0.019. But OVTT coefficients have a wide range from -0.114 to -0.028, and only five models fall in the range of $2.0 < (C_{ovtt}/C_{ivtt}) < 3.0$.

In 1999, Rossi and Outwater (1999) reviewed mode choice models from 11 MPOs across the country, which were developed mostly based on survey data in the 1980s and 1990s. This review shows similar values for HBW models; the IVTT coefficients range from -0.046 to -0.016, with an average of -0.029. OVTT coefficients for auto ranged from -0.093 to -0.038, with an exception of -0.26 from a model based on a 1965 survey. OVTT coefficients for walk had a narrower range of -0.069 to -0.032. Home-based other and non-home-based trip models have wider ranges than home-based work models.

In 2005, EPA documented mode choice model coefficients for HBW models from 26 metropolitan areas across the country, most of which were derived from survey data conducted in the 1990s. Similar to the earlier review, the IVTT coefficients range from -0.045 to -0.0113, averaging -0.0253, and OVTT coefficients have a wider range with walk time coefficients from -0.0931 to -0.0186 and transit wait time coefficients from -0.0978 to -0.155.

More recently, Cambridge Systematics (2008) reviewed mode choice model coefficients from 13 urban areas. These model coefficients were estimated without constraints, using survey data largely collected in the 1990s and 1980s. For home-based work mode choice models, the average values of mode choice coefficients are similar to those from previous reviews.

As described above, some MPOs certainly encountered estimation difficulties during mode choice model estimation. It is hard to untangle the root causes of those “unusual” coefficients, which might arise from poor accuracy, limited sampling size, (non)representativeness of the survey data, measures of the level of service variables derived from the model, interactions among independent variables, locality-specific travel behaviors, the nature of the nested logit model, and limitations of the estimation software.

One of the difficulties in estimating a nested logit model is that a unique optimal solution for a set of parameters may not exist (Koppelman and Bhat, 2006). There could be multiple optima in the nested logit model parameters, depending on the initial values used for estimating these parameters. Koppelman and Bhat (2006) show that very different initial nesting parameters may or may not reach similar solutions for estimated model parameters. It also is common in nested logit model estimations that some model parameters are out of reasonable ranges or estimated nesting coefficients are greater than one or greater than the nesting parameter in a higher-level nest. Although this theoretically means rejection of the proposed model structure, there are practical reasons for such estimation results, including imperfect data. Koppelman and Bhat (2006) believe that “it is a matter of judgment whether to eliminate the proposed structure based on the estimation results or to constrain selected nest parameters to fixed values that ensure that the structure is consistent with utility maximization.”

FTA recognizes the limitations of mode choice model estimations and in particular, believes that too many resources were spent on model estimation and too little resources on model calibration and validation. Therefore, FTA recommends that resources be better spent in careful mode choice model calibration and validation, with asserted coefficients for time and costs.

5.3 Recommendations

A careful approach to model estimation is the first critical step in the development of a policy-sensitive model of travel behavior. First, a model that is based on local survey data will more properly reflect the trade-offs between “level of service” and cost that are faced daily by area residents than one simply imported from another region. Second, a necessary building block for model estimation is the quality, accuracy, and timeliness of the travel data in the household survey. A sample rich in observations for all public transit and non-motorized modes is a requirement to properly reflect all modes available in the region. Third, another building block is the accuracy of the spatial data, specifically whether origins and destinations as reported by the respondents can be extracted from the survey and geocoded with reasonable precision. It is just as important that the underlying highway and transit networks provide realistic travel time components and paths for all available origin-destination pairs. Fourth, sampling/weighting is a critical issue. The final weighted survey data should result in a proper balancing of all observations that can be considered representative of travel decisions and available travel options in the region. Fifth, careful evaluation and validation of the resulting models is necessary.

There is a considerable body of literature over the last four decades, of which FTA guidance is a critical component, that should be used to evaluate model reasonableness. We recommend using FTA recommended parameters as one of the guiding principles in evaluations of the performance of the estimated models. We recommend engaging FTA in this process early to jointly evaluate the properties of the household survey and on-board survey data sets, the quality of the networks and the trade-offs that travelers face. If FTA is engaged in this evaluation process, the estimation process can be approved.

In the broader context, it is recommended that a hybrid approach be taken to develop a new mode choice model for the next model update, namely that standard mode choice models will be estimated but if (or when) unreasonable results emerge from the estimation, then FTA approved values and relationships between coefficients will be asserted. This approach will ensure compliance of the resulting mode choice model with FTA requirements, and at the same time, allow the flexibility of incorporating additional variables that are of interest to the TPB. These additional variables may be important in explaining travel behaviors in the Washington metropolitan area, and also be important to analysis of some policies.

References

Allen, William G. (2006). Baltimore Metropolitan Council Travel Forecasting Model Calibration Report.

Atlanta Regional Commission (ARC) (2008). The Travel Forecasting Model Set for the Atlanta Region: 2008 Documentation.

Cambridge Systematics (1997). Enhancement of DVRPC's Travel Simulation Models - Task 4: Transit Assignments. Prepared for Delaware Valley Regional Planning Commission, April 1997.

Cambridge Systematics (2002). SEMCOG Travel Model Documentation.

Cambridge Systematics (2004). NCTCOG Mode Choice Model Estimation.

Cambridge Systematics (2007). PSRC Travel Model Documentation (for Version 1.0): Updated for Congestion Relief Analysis. Final Report.

Cambridge Systematics (2008). FSUTMS-Cube Framework Phase II Model Calibration and Validation Standards Final Report, Prepared for Florida Department of Transportation Systems Planning Office, October 2, 2008.

Cambridge Systematics (2009). Fiscal Year 2009 Task Reports. Final Report. National Capital Region Transportation Planning Board, November 20, 2009.

Central Transportation Planning Staff (2009). Methodology and Assumptions of Central Transportation Planning Staff Regional Travel Demand Modeling: North Shore DEIR Support.

Delaware Valley Regional Planning Commission (DVRPC) (2008). 2000 and 2005 Validation of the DVPRC Regional Simulation Models.

DKS Associates (2002). Sacramento Regional Travel Demand Model: Version 2001 (SACMET 01) Model Update Report.

Environmental Protection Agency (EPA) (2005). Commuter v2.0 Model Coefficients. EPA420-B-05-019. Office of Transportation and Air Quality.

Federal Transit Administration (FTA) (2004). Travel Forecasting for New Starts: The FTA Perspective.

Gannett Fleming (2007). Tampa Bay Regional Planning Model, Cube Voyager Conversion, Technical Report No. 1, Validation Report, 2007.

Koppelman, F.S. and C. Bhat (2006). A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models. Prepared for Federal Transit Administration.

Kuppam, A., M. Outwater, M. A. Bradley, L. Blain, R. Tung, and S. Yan (2005). "Application of Time-of-Day Choice Models Using EMME/2: Washington State DOT Congestion Relief Analysis" Presented at the 19th International EMME/2 Users' Conference, Seattle, Washington, October 19, 2005.

Martin, W. A. and N. McGuckin (1998). Travel Estimation Techniques for Urban Planning. NCHRP Report 365. Washington, D.C.: National Academy Press.

Kim, K. (2008). Metro Travel Forecasting 2008 Trip-Based Demand Model Methodology Report. Portland METRO.

Lewis-Workman, S. (2008). Predicted vs. Actual Costs and Ridership of New Starts Project, Predicted vs. Actual Ridership. Presented to the Annual Meeting of Transportation Research Board, Session 371, January 14, 2008.

Milone, R. (2004). "Methodology for linking future bus speeds to highway congestion in the Version 2.1/D model." Memorandum to Files, July 14, 2004.

Milone, R., H. Humeida, M. Moran, and M. Seifu (2009). FY 2009 Development Program for TPB Travel Forecasting Models. Draft report. Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board (COG/TPB), June 30, 2009.

Milone, R., H. Humeida, M. Moran, and M. Seifu (2008a). TPB Travel Forecasting Model, Version 2.2: Specification, Validation, and User's Guide. Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board (COG/TPB), March 1, 2008.

Milone, R., H. Humeida, M. Moran, and M. Seifu (2008b). TPB Travel Forecasting Model, Version 2.3: Specification, Validation, and User's Guide. Draft report. Washington, D.C.: Metropolitan Washington Council of Governments, National Capital Region Transportation Planning Board (COG/TPB), June 30, 2008.

Milone, R., H. Humeida, M. Moran, M. Seifu, and M. Martchouk (2010). FY-2010 Development Program for TPB Travel Forecasting Models: DRAFT Report. Washington, D.C.: National Capital Region Transportation Planning Board (TPB), June 30, 2010.

Milone, R., R. Snead, J. Posey, W. Hamlin, J. Davis, W. Bacon, C. Howard, M. Kile, M. Seifu, M. Moran, and J. Park (2009). FY-2009 Network Documentation: Highway and Transit Network Development. Washington, D.C.: COG/TPB, June 30, 2009.

Moran, M. S. (2009). Memorandum to John (Jay) Evans. "Proposed Task Order #7 of contract 08-056 (FY 2010): Further investigation of convergence in user equilibrium traffic assignment and speed feedback." Memorandum, November 30, 2009.

North Central Texas Council of Governments (NCTCOG) (2000). Dallas-Fort Worth Regional Travel Demand Model (DFWRTM): Description of the Multimodal Forecasting Process.

North Central Texas Council of Governments (NCTCOG) (2009). Dallas-Fort Worth Regional Travel Demand Model (DFWRTM): Model Description.

Pu, W., and A. Meese (2009). "Update on the Congestion Management Process (CMP)" Presented at the Travel Management Subcommittee, Washington, D.C., November 24, 2009.

Purvis, C. L. (1997). Travel Models for the San Francisco Bay Area (BAYCAST-90), Technical Summary, Metropolitan Transportation Commission, June 1997.

Rossi, T. F. and M. L. Outwater (1999). How Valid is It to Transfer Mode Choice Model Parameters? Proceedings of the Seventh National Conference on Transportation Planning Applications, Transportation Research Board, Boston, Massachusetts.

Rossi, T., B. Winkler, T. Ryan, K. Faussett, Y. Li, D. Wittl, and M. Abou Zeid (2009). "Deciding on Moving to Activity-Based Models (or Not)." TRB 88th Annual Meeting Compendium of Papers DVD, Washington, D.C.: Transportation Research Board.

San Diego Association of Governments (SANDAG) (2008). 2030 Regional Growth Forecast Update: Process and Model Documentation. April 2008.

Southern California Association of Governments (SCAG) (2003). SCAG Regional Transportation Model: Model Validation and Summary.

URS Corporation (2008). North Jersey Regional Transportation Model Enhanced. North Jersey Transportation Authority (NJTPA).

Vanasse Hangen Brustlin, Inc. (VHB), Resource Systems Group, Inc., Shapiro Transportation Consulting, LLC, and Urban Analytics (2009). "Task 1 Report: Identify MPO Experiences with Advanced Travel Models That Would Be of Interest to Other Practitioners." Memorandum to Association of Metropolitan Planning Organizations.

Vanasse Hangen Brustlin, Inc. (VHB) (2007). Determination of the State of the Practice in Metropolitan Area Travel Forecasting: Findings of the Surveys of Metropolitan Planning Organizations: Prepared for Committee B0090 TRB/National Research Council.

Virginia Department of Transportation (VDOT) (2009). Implementing Activity-Based Models in Virginia. VTM Research Paper. Richmond, Virginia: Virginia Department of Transportation (VDOT), July 2009.

This page intentionally left blank.