SHRP 2 Capacity Project C10A

Partnership to Develop an Integrated, Advanced Travel Demand Model and a Fine-Grained Time-Sensitive Network

PREPUBLICATION DRAFT • NOT EDITED

SHRP 2
STRATEGIC HIGHWAY RESEARCH PROGRAM
Accelerating solutions for highway safety, renewal, reliability, and capacity

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

© 2013 National Academy of Sciences. All rights reserved.

ACKNOWLEDGMENT
This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials. It was conducted in the second Strategic Highway Research Program, which is administered by the Transportation Research Board of the National Academies.

NOTICE
The project that is the subject of this document was a part of the second Strategic Highway Research Program, conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council.

The members of the technical committee selected to monitor this project and to review this document were chosen for their special competencies and with regard for appropriate balance. The document was reviewed by the technical committee and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the Governing Board of the National Research Council.

The opinions and conclusions expressed or implied in this document are those of the researchers who performed the research. They are not necessarily those of the second Strategic Highway Research Program, the Transportation Research Board, the National Research Council, or the program sponsors.

The information contained in this document was taken directly from the submission of the authors. This document has not been edited by the Transportation Research Board.

Authors herein are responsible for the authenticity of their materials and for obtaining written permissions from publishers or persons who own the copyright to any previously published or copyrighted material used herein.

The Transportation Research Board of the National Academies, the National Research Council, and the sponsors of the second Strategic Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

SHRP 2 C10A:

Partnership to Develop an Integrated, Advanced Travel Demand Model and a Fine-Grained Time-Sensitive Network

Final Report

DRAFT

August 2013

Resource Systems Group
AECOM
Mark Bradley Research & Consulting
John Bowman Research & Consulting
Mohammed Hadi, FIU
Ram Pendyala, ASU
Chandra Bhat, UT Austin
Travis Waller, UT Austin
North Florida Transportation Planning Organization

DATA ■ ANALYSIS ■ SOLUTIONS

Table of Contents

List of Figures

LIST OF TABLES

1.0 EXECUTIVE SUMMARY

1.1 Purpose and Need

1.1.1 Modeling Travel

Travel models are used to support decision-making by providing information about the impacts of transportation and land use investments and policies, as well as demographic and economic trends. When applied properly, they can provide a consistent framework for evaluating different alternative scenarios. Transportation decision-makers need to have confidence that the tools they use to inform policy and investment decisions, including travel demand forecasting models, produce reasonable results that are appropriately sensitive to the questions at hand.

Most travel models are comprised of a set of components that address different aspects of traveler choices, such as determining the number and purpose of trips to be made, the origins and destinations of these trips, the travel mode (such as driving alone, or riding transit), and the specific network routes used. These "4 steps" of travel models can be broadly grouped into "demand" and "supply" categories, with the first three steps (generation, distribution, mode choice) describing the demand components, and the last step (assignment) describing the supply components. Recent methodological advances have occurred with both model demand components as well as model supply components, and these advances provide the opportunity to develop more robust travel models for use in transportation decision-making.

1.1.1.1 Activity-Based Demand Models

On the demand side, activity-based models have become increasingly adopted by metropolitan planning organizations. Activity-based travel demand models produce estimates of daily activity patterns including tour- and trip-generation, destination choice, mode choice, and time-of-day choice. A tour is a chain of trips that begin and end at home or work, and is essential for representing the interrelationships between activities undertaken by travelers. Daily activity pattern models consider the coordinated aspects of travel made by an individual across the entire day, as well as activities potentially coordinated across individuals within a household. In addition, these models typically incorporate accessibility measures that allow changes in network performance to influence demand generation.

Agencies typically develop and apply activity-based models in order to have sensitivities to policies that are may be challenging to represent within the context a traditional "trip-based" model system. For example, the impacts of pricing policies on demand generation, destination and mode choices can be better captured using activity-based models than using more trip-based methods. A number of features distinguish activity-based approaches to modeling demand from traditional trip-based approaches, such as:

- Activity-based models represent travel demand in a more intuitively correct manner than traditional trip-based demand models because they simulate individual- and household-level travel choices, such as whether to make a tour or stop to participate in an activity where this activity will take place (such as whether to work-at-home or journey to work), and when and how to get there, in an intuitive way that captures opportunities and constraints;

- Activity-based models provide more consistency, and potentially more detail, across all dimensions of travel behavior, especially space and time, which in turn results in more realistic representations of transportation system performance by the network supply model. Significantly, activity-based models do not include "non-home-based trips," which frequently comprise a large proportion of the demand in traditional trip-based models but for which trip-based models cannot include potentially relevant information such as prior trip mode choices and traveler income;
- Activity-based models include significantly more detail on traveler attributes and spatial and temporal constraints, which provides better estimates of the transportation impacts of a given alternative scenario. For example, activity-based models can assign person-specific and purpose-specific values of time to different individuals, which is important for modeling pricing alternatives. Such detailed market segmentation is possible due to the disaggregate nature of most activity-based model implementations, and is often intractable in the context of aggregate trip-based travel models;
- Activity-based models produce a wider range of performance measures, with greater detail.

Perhaps most significantly for the C10A project, activity-based demand models explicitly include a detailed representation of time-of-day using temporal units of half-hours and minutes rather than broad multi-hour time periods. This temporal resolution facilitates the incorporation of changes in network performance by time-of-day produced by the dynamic supply model, as well as provides an explicit method for reflecting availability constraints (such as time window accounting) that produces activity patterns that are logically consistent in both time and space.

1.1.1.2 Dynamic Traffic Assignment Models

On the supply side, dynamic traffic assignment approaches are also increasingly being adopted by metropolitan planning organizations. Traffic assignment is the fourth and final step of the traditional "4 step" planning process. Until the past decade, virtually all travel models have incorporated "static" traffic assignment methods, which produce estimates of travel times, costs, and volumes across relatively broad time periods. However, transportation policies increasingly incorporate time-varying assumptions, and the analysis and management of transportation network performance requires information about time-varying network times, costs and flows. Static based assignment approaches are unable to represent with sufficient detail time varying flows and congestion, and the impacts on travel times and costs. In contrast, dynamic network models provide the ability to represent time-varying network time and costs, and can provide more information on network performance by detailed time of day, which can be used as input into travel model demand components. Features which distinguish dynamic network methods from static network approaches include:

- DTA models incorporate more complete representations of transportation network attributes and configurations, including the ability to better represent intersection controls such as signal synchronization and or other advanced network control schemes;
- DTA models use more realistic flow models to propagate traffic on links, rather than simplified volume delay functions which may produce unrealistic estimates of network times and volumes;

- DTA models provide more detailed estimated of network system performance, which is essential for accurately to evaluating the impacts of different transportation policy, systems management, and funding alternatives.

1.1.1.3 Integrated Models

As transportation policy and investment questions have become more complex, and as additional models have been developed and previous models made more behaviorally descriptive, separate models often viewed independently are increasingly viewed as inter-dependent. The primary project objective of this project is to make operational a dynamic integrated model–an integrated, advanced travel-demand model with a fine-grained, time-dependent network, and to demonstrate the model's performance through validation tests and policy analyses. This integrated model system is essential because most current travel models are not sufficiently sensitive to the dynamic interplay between travel behavior and network conditions, and are unable to reasonably represent the effects of transportation policies such as variable road pricing and travel demand management strategies. The availability and capabilities of activity-based demand models and dynamic network supply models provide the opportunity to address the shortcomings of current tools and provide decision-makers with more complete information.

1.1.2 Project Objectives

The primary objective of the C10A project is to test the principal of making operational a dynamic integrated model–an integrated, advanced travel-demand model with a fine-grained, time-dependent network, and to demonstrate the model's performance through validation tests and policy analyses. This integrated model system is necessary because most current travel models are not sufficiently sensitive to the dynamic interplay between travel behavior and network conditions, and are unable to reasonably represent the effects of transportation policies such as variable road pricing and travel demand management strategies. Secondary project goals include producing a transferrable process and sample data that can be used in other regions, demonstrating an effective interface with EPA's MOVES model, incorporating knowledge from other SHRP 2 efforts such as C04 (pricing) and C05 (operations) and addressing travel time reliability in travel models. This report describes the tools incorporated into the integrated model system, the data required to implement these tools, modifications to the tools that were necessary to achieve this integration, and results of a set sensitivity tests of the integrated model system.

The C10A implementation was envisaged to be in a region in which choices of non-highway modes are limited, and as such the dynamic integrated model represents behavioral changes in response to roadway conditions. To meet this objective, the model system was designed to capture changes in demand such as time of day choice (i.e. peak spreading) and route choice in response to capacity and operational improvements such as signal coordination, freeway management, variable tolls, and capacity improvements.

While the primary project objective called for the development of a dynamic integrated model with advanced policy analysis capabilities, it was also important that it be feasible for advanced practitioners to be able to implement the model system in other regions without excessive costs or undue complexity. This resulted in a model system with the following features:

- The model is scalable. While the model system implemented for C10 can exploit distributed computing in order to reduce model system runtimes, it does not require a large hardware cluster or complex computing environment to run.

- The model is relatively easy to implement and maintain. Although the model system is inherently complex due to its advanced capabilities, it can be easily and flexibly configured to operate with different levels of temporal and spatial detail and in different computing environments.
- The model system does not require a multi-year, multi-million dollar implementation and maintenance effort. The model system was implemented in two regions, and subjected to a set of initial calibration and sensitivity tests in approximately one year.

The purpose of this report is to document the implementation of the model system. The report describes the implementation of the model system in both Burlington, Vermont and in Jacksonville, Florida, the calibration and validation of the model system, and the application of the model system to a set of initial sensitivity tests.

1.2 Model System Components

The proposed model system is comprised of three primary components: DaySim, the TRANSIMS Router and Microsimulator, and MOVES. DaySim is a travel demand forecast model that predicts household and person travel choices at a parcel-level on a minute-by-minute basis. The TRANSIMS Router and Microsimulator are dynamic traffic assignment and network simulation software that can perform regional traffic microsimulation on a second-by-second basis. MOVES is the EPA's latest software for estimating emissions and air quality impacts. The integrated model links these model components in an equilibrated model system that provides enhanced policy sensitivities at significantly higher levels of spatial and temporal resolution than found in a traditional regional travel demand forecasting system.

1.2.1 DaySim

The travel demand model to be used for this project is coded in a software framework called DaySim. DaySim is one of the two main "families" of activity-based model systems now being used by MPO's in the U.S. DaySim was initially implemented in Sacramento, CA, and has been enhanced to interface effectively with the TRANSIMS tools.

DaySim simulates 24-hour itineraries for individuals with spatial resolution as fine as individual parcels and temporal resolution as fine as single minutes, so it can generate outputs at the level of resolution required as input to dynamic traffic simulation. DaySim's predictions in all dimensions (activity and travel generation, tours and trip-chaining, destinations, modes, and timing) are sensitive to travel times and costs that vary by mode, origin–destination (OD) path, and time of day, so it can, in turn, effectively use as inputs the improved network travel costs and times output from a dynamic traffic simulator. DaySim is structured as a series of hierarchical or nested choice models. The general hierarchy places the long term models such as auto availability at the top of the choice hierarchy, and the short term models such as trip mode and time-of-day choice at successively lower levels in the hierarchy. A more complete description of the DaySim structure and capabilities can be found in the SHRP 2 C10A Task 1 Report: Project Approach and Industry Synthesis. A description of the DaySim-TRANSIMS linkage is provided in Section 2.8.10 of this document.

1.2.2 TRANSIMS

TRANSIMS network and travel assignment processes are used to represent the performance of the transportation networks in the integrated model system. TRANSIMS assigns a sequence of trips or tours for individual household persons between specific activity locations (smaller than travel analysis zones but larger than individual parcels) to roadways, walkways, and transit modes on a second-by-second basis for a full travel day. The TRANSIMS networks include detailed information regarding the operational characteristics of the transportation facilities that may vary by time of day and by vehicle or traveler type. This includes the number of lanes, the lane use restrictions, the traffic controls and signal timing and phasing plans, turning restrictions, and tolls.

TRANSIMS implements a dynamic user equilibrium network assignment for trip and activity files that define the demand by detailed time of day. The primary demand input to TRANSIMS is an activity file produced by DaySim that contains information on each individual's activity locations, timing, and mode of travel. In addition, trip list files are used to represent non-household-related travel such as trucks, external trips, and other commercial travel in the network demand. TRANSIMS tools also used to generate zone-to-zone network impedance measures by detailed time of day for use in subsequent DaySim demand simulations. A description of this TRANSIMS-DaySim linkage can be found in Section 2.8.12.

1.2.3 MOVES

The MOVES software was developed by the Environmental Protection Agency to provide estimates of emissions and greenhouse gases. MOVES uses detailed information about the distribution of VMT by source type, facility type, area type, time of day, week day, and 5 mph average speed bins to calculate emissions for an array of pollutants. In addition to travel data, MOVES uses information about the fleet age and fuel type distributions, inspection maintenance programs and monthly temperatures and humidity for each county in the analysis area. These are used to calculate county-based emissions inventories or custom domains that combine counties into aggregate estimates. TRANSIMS tools have been developed to interface with MOVES to support the generation of these estimates. A detailed description of the TRANSIMS-MOVES linkage is provided in Section 2.8.13 of this document.

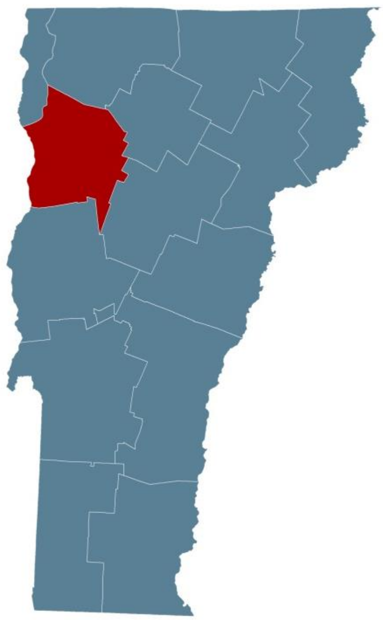
1.3 Regional Implementations

As part of the C10A project, the DaySim-TRANSIMS-MOVES model system was implemented in two regions: Burlington, Vermont and Jacksonville, Florida.

1.3.1 Small-scale Regional Testbed

The integrated model system was first implemented in the Burlington, Vermont region. The purpose of this implementation was to establish a test-bed for developing and refining model system capabilities and configurations. The Burlington modeling area is comprised of a single county (Chittenden) of approximately 620 square miles, and is home to 55,000 households in the base year of 2005. These households generate approximately 525,000 daily person trips.

Figure 1. Burlington Model Area



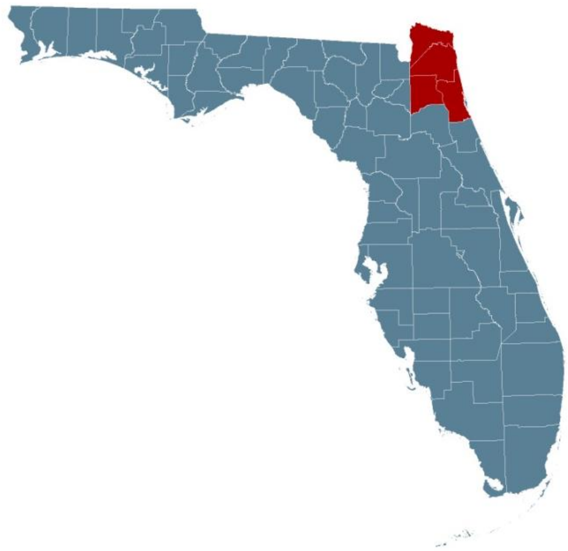
From a development perspective, the primary advantage of the implementation of the model system in a smaller region is that it allows for the ability to more rapidly test alternative model configurations and to debug model processes due to shorter model system runtimes. These shorter runtimes are associated with both the DaySim "demand" component of the model system as well as the TRANSIMS "supply" component of the model system.

In the DaySim model, runtimes are directly related to the amount of demand, so region with a smaller population can be simulated more quickly. However, the DaySim demand model is not the primary performance "bottleneck" in the model system. The overall model system runtimes are primarily driven by the performance of the TRANSIMS Router and Microsimulator network assignment tools. Like DaySim runtimes, TRANSIMS runtimes are related to the amount of demand being simulated, but they are also significantly influenced by the level of transportation network detail – specifically the number of links in the network. In the Burlington implementation, the TRANSIMS network is relatively coarse, and the small modeling area also limits the number of network links. However, this smaller region cannot support the full range of model system sensitivity testing required by the C10A project because the levels of congestion in the region are relatively low. As a result of this limited congestion, the model system's responses to a number of policies and improvements is limited.

1.3.2 Large-Scale Regional Demonstration

Subsequent to the initial model implementation in the Burlington region, the integrated model system was implemented in the Jacksonville, Florida region. The purpose of this implementation was to provide a more robust and challenging context for testing the model system capabilities. The Jacksonville region is comprised of four counties in northeast Florida covering 3,100 square miles. The regional population includes over 525,000 households and 1.25 million people, and generates more than 4 million daily person trips.

Figure 2. Jacksonville Model Area



From a model application and sensitivity testing perspective, the purpose of implementing the model system in this larger region was to provide a platform for subjecting the model system to a broader and more rigorous set of policy sensitivity tests, due to higher levels of network congestion observed in this region. The primary disadvantage of using this larger model region is that the additional demand and network detail results in significantly longer model system runtimes, primarily attributable to the TRANSIMS Microsimulator. Use of this larger area for model development would have resulted in a longer model development phase. The longer runtimes associated with the Jacksonville integrated model implementation also necessitated the development and testing of a number of alternative application modes, as described in the following section.

1.4 Application Modes

The primary driving force behind the SHRP 2 C10 project is the need to address transportation policies that are being considered in metropolitan planning organizations around the U.S, but which are not adequately addressed by the current state of the practice travel forecasting models. Some of the broad categories of policies and strategies that the integrated, time-sensitive model developed for this project seeks to address include:

- Pricing policies
- Capacity enhancements
- Transportation System Management (Operations) improvements
- Travel demand management policies
- Greenhouse gas reduction strategies

The three primary components of the integrated model system, DaySim, TRANSIMS and MOVES, each provide unique capabilities and can be flexibly configured to address the different analysis needs associated with these different policies and strategies. The project team developed different methods of combining and linking the model system components in application as a result of practical experience in working with and testing the model system. Specifically, some policies or improvements such as roadway pricing require regional-scale analysis, but regional scale microsimulation can result in excessively long runtimes while adding little policy-specific sensitivity. Conversely, other smaller scale policies or improvements such as signal coordination in a corridor may not be expected to impact overall regional travel patterns but may require the local sensitivities of a traffic microsimulation model.

In order to balance the policy analysis needs against practical runtime considerations, the project team developed a set of model system application modes: planning, operations, and planning+operations.
Table 1 illustrates some typical Jacksonville model system runtimes for these application modes when implemented and distributed on the TRACC computing cluster at Argonne National Lab. Note that the runtimes are highly dependent on the particular hardware being used, the specific versions of the software tools employed (which are frequently updated), and the level of convergence required for a particular analysis. As computing power increases, it is expected that runtimes will decrease. The following sections describe the configuration of these application modes and identifies the types of policies or improvements that each might most effectively test.

Table 1. Application Mode Runtimes

	Planning (hours)	**Operations (hours)**	**Planning+Operations (hours)**
DaySim demand estimation	4.0	4.0	4.0
Assignment Iteration	0.5	5.0	5.0
Convergence checking	1.0	1.0	1.0
Skimming procedures	1.0	0.0	1.0
TOTAL	6.5	10.0	11.0
Assignment Iteration	40	40	40
System Iterations	3	1	3
Total Iterations	120	40	120
TOTAL SYSTEM RUNTIME			
Hours	195	244	735
Days	8	10	31
Weeks	**1.2**	**1.5**	**4.4**

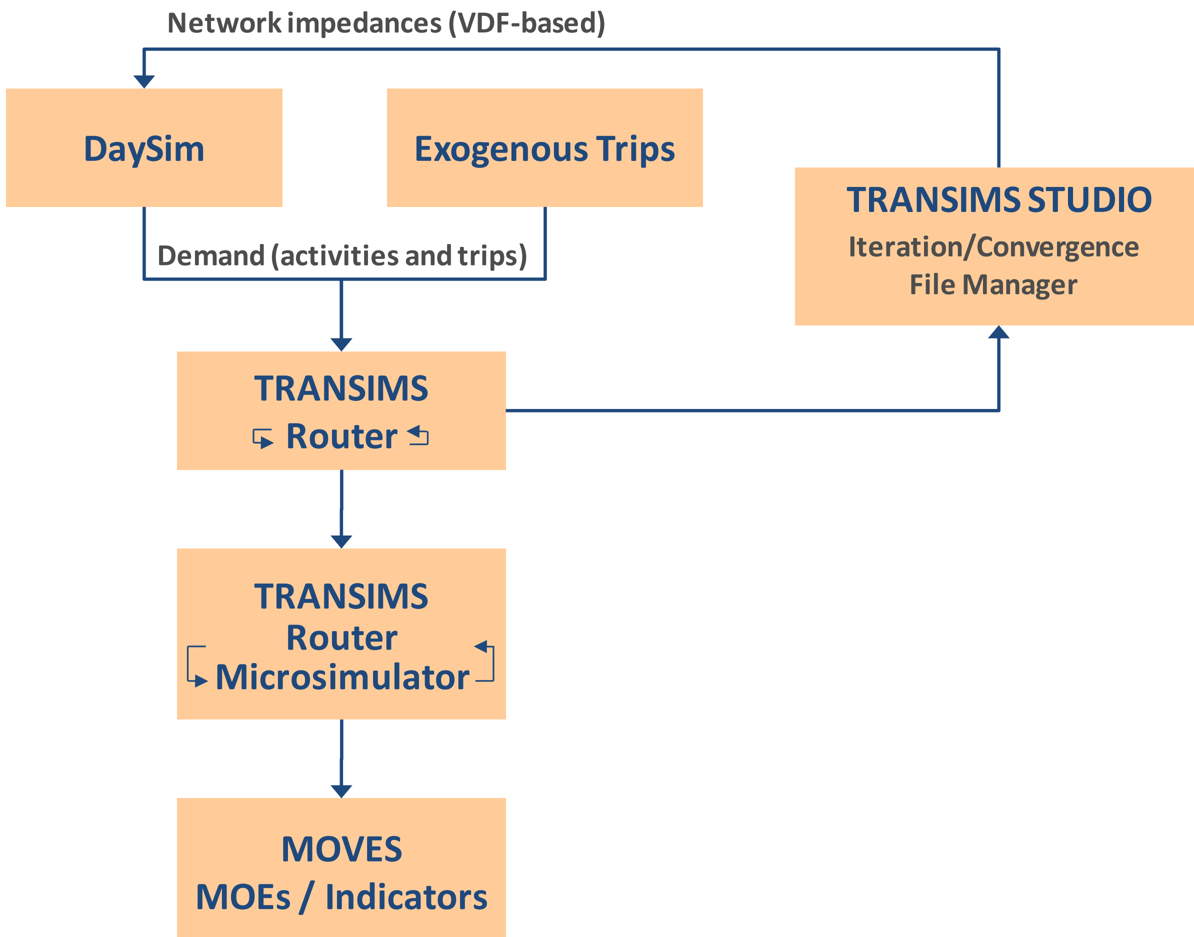
1.4.1 Planning Mode

The planning application mode can be used when the analysis needs are expected to result in regional-scale changes in overall level of travel demand, or changes in regional travelers' destination, mode, or time of day choices, but which are not expected to be significantly impacted by local-scale traffic dynamics. The planning mode integrates the DaySim demand model with the TRANSIMS supply model in an iterative feedback loop in which DaySim outputs estimates of travel demand at the level of individual minutes for routing within the TRANSIMS Router. Temporally detailed network impedance skims based on these Router assignments are then generated and fed back as input to DaySim. A full scale regional TRANSIMS microsimulation may be optionally run as a "post-process" to the integrated DaySim-TRANSIMS Router application

1.4.1.1 Planning Mode Configuration

The distinguishing feature of the planning application mode is that only the TRANSIMS Router is used in an integrated way with DaySim, with the TRANSIMS Microsimulator available as a post-process. The TRANSIMS Router operates at detailed temporal resolutions such as 5-minutes or 15-minutes and incorporates important features such as the use of time-dependent shortest pathbuilding, but uses traditional volume-delay functions to convert assigned volumes into 5-minute or 15-minute measures of link delay. Use of these volume delay functions is inferior to the use of Microsimulator-based delays in which the travel times and costs experienced by individual travelers are used to directly generate the times and delays used in pathbuilding. The Router also lacks some key functionalities of the Microsimulator such using actual signal timings to estimate delays at intersections, instead relying of fixed delays derived from prior assignment iterations. A critical advantage of using the Router only in the planning mode is that it runs relatively quickly even at a regional scale because it can be partitioned across multiple processing cores. By incorporating the Microsimulator as a post-process, the impact of strategies and policies on regional and local traffic dynamics may also be assessed, albeit not in an integrated way. Figure 3 illustrates the configuration of the model system components in the planning mode.

Figure 3. Planning Mode System Configuration



1.4.1.2 Planning Mode Applications

The planning mode can be used when the policies or strategies being considered are expected to result in regional-scale changes in overall level of travel demand, or changes in regional travelers' destination, mode, or time of day choices, but which are not expected to be significantly impacted by local-scale traffic dynamics. The planning mode can be applied to the following primary policy and strategy analyses:

- Pricing: Pricing strategies are costs that are imposed on travelers using certain roads, traversing certain screenlines, or travelling to certain areas (tolling, cordon pricing or area pricing). These costs may be either fixed, or vary by time-of-day or in response to congestion. Additionally, these costs may vary by user to reflect discounts or subsidies provided to some users. Pricing strategies are most effectively addressed in the context of a regional scale model due to the nature of the potential responses to pricing strategies, which may include changes in the overall level of activity and trip generation, changes in the destinations for these activities, changes in the travel modes used to access these destination, and changes in the specific routes on the roadway or transit networks given the selected mode.

- Capacity: Capacity strategies involve adding, modifying or deleting capacity on the roadway system. This may include the addition of new roads or lanes to the travel model networks, or may involve adjusting existing capacity, such as the implementation of reversible lanes, auxiliary lanes or turn lanes at intersections. While the impacts of local capacity enhancements may be better captured using traffic microsimulation tools, significant increases in capacity such as the addition of new roads or lanes are better addressed using regional scale models because of their potentially broad impacts on regional network level of service, which would influence, generation, distribution, mode choice, route assignment.
- Travel Demand Management: Travel demand management strategies are typically aimed at changing travel behavior to reduce congestion and improve mobility. For example, these policies may seek to increase the frequency and numbers of people who work at home, induce workers to adjust their schedules to travel in off-peak, less congested conditions, or increase the number of people who carpool to work. These policies are most appropriately addressed at a regional scale because of their expected impact on performance on regionally significant / congested facilities. However, use of a detail model of traffic dynamics is not necessarily required to capture the impact of these policies

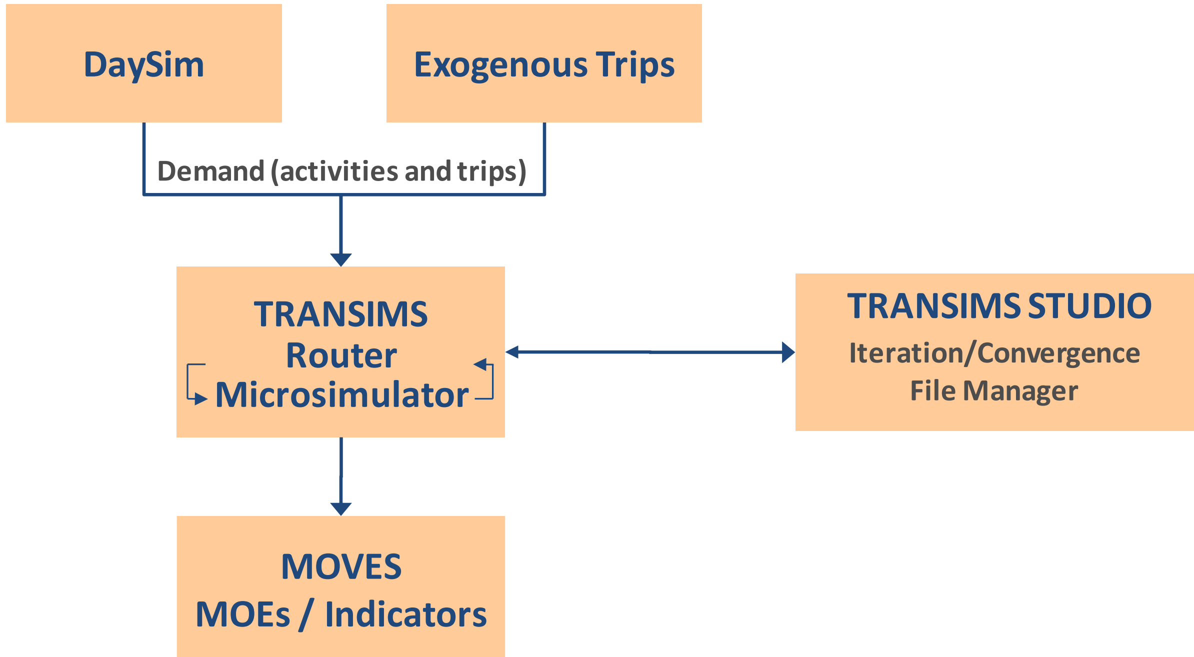
1.4.2 Operations Mode

The operations mode can be used when the analysis requires an assessment of the impacts of a policy or strategy on local traffic dynamics, and where these improvements are not expected to result in changes in overall level of travel demand, or in destination, mode or time of day choices.

1.4.2.1 Operations Mode Configuration

The distinguishing features of the operations mode are that it incorporates a full regional traffic microsimulation, but that it does not include an iterative feedback loop in which Microsimulator-based network simulation impedance measures are fed back to DaySim. The elimination of this feedback loop reflects the fact that some operational improvements may greatly improve local traffic dynamics but have only marginal effects on the other travel dimensions, as well as the fact that regional microsimulation is computationally intensive and results in extremely long runtimes. Figure 4 illustrates the configuration of the model system components in the operations mode.

Figure 4. Operations Mode System Configuration



1.4.2.2 Operations Mode Application

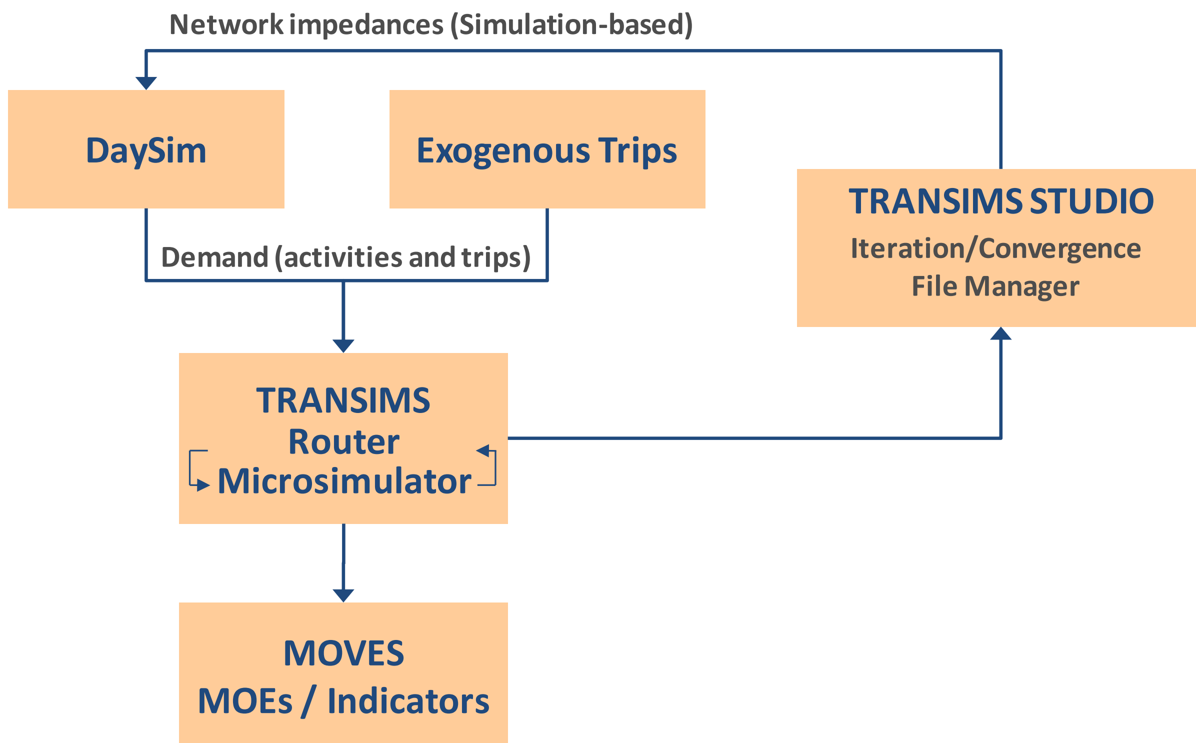
The operations mode can be used when the policies or strategies under consideration are not expected to result in significant changes in overall level of travel demand, or in destination, mode or time of day choices. The operations mode can be applied to the following primary policy and strategy analyses:

- Capacity: Some capacity improvements or enhancements to existing capacity may be evaluated using more local, operationally focused tools. These may include changes such as turn lanes at intersections, other geometric changes such as lane connectivity, lane widths, and the presence of shoulders.
- Operations: The operations model may support the analysis of bottleneck improvements, such as the addition of new signals or signs, adjustment of signal timing and phasing, or implementation of ramp meters. These are often most appropriately tested at a local scale using network assignment tools, while holding the other choice dimensions fixed. However, more extensive bottleneck or other operational improvements, such as those applied across an extensive coordinated signal systems within a corridor or at a regional scale, may be appropriately tested using the full integrated planning+operations application mode described below.

1.4.3 Planning + Operations Mode

The planning+operations mode represents the fully integrated DaySim and TRANSIMS model system. In this application mode, the TRANSIMS Router and Microsimulator are used to perform a regional traffic microsimulation as part of every model system global iteration, and microsimulation-based network impedance measures are fed back and used as input to DaySim. The advantage of this application mode is that it provides the full range of sensitivities to both changes in regional demand and to local and regional traffic dynamics. However, these extensive sensitivities come with the significant disadvantage of extremely long model system runtimes. Figure 5 illustrates the configuration of the model system components in the planning mode.

Figure 5. Planning+Operations Mode System Configuration



1.4.3.1 Planning + Operations Mode Applications

The planning+operations mode can be used when the proposed policies or strategies are expected to result in regional-scale changes in overall level of travel demand, changes in regional travelers' destination, mode, or time of day choices, and which expected to be significantly impacted by local-scale traffic dynamics. The planning mode can be applied to the following primary policy and strategy analyses:

- <u>Pricing:</u> Pricing strategies are costs that are imposed on travelers using certain roads, traversing certain screenlines, or travelling to certain areas (tolling, cordon pricing or area pricing). These costs may be either fixed, or vary by time-of-day or in response to congestion. Additionally, these costs may vary by user to reflect discounts or subsidies provided to some users. Pricing strategies are most effectively addressed in the context of a regional scale model due to the nature of the potential responses to pricing strategies, which may include changes in the overall level of activity and trip generation, changes in the destinations for these activities, changes in the travel modes used to access these destination. Use of the planning+operations mode is necessary when the pricing policies are also expected to have regionally significant impacts of traffic dynamics.

- Capacity: Capacity strategies involve adding, modifying or deleting capacity on the roadway system. This may include the addition of new roads or lanes to the travel model networks, or may involve adjusting existing capacity, such as the implementation of reversible lanes, auxiliary lanes or turn lanes at intersections. Significant increases in capacity such as the addition of new roads or lanes are most effectively addressed using regional scale models because of their potentially broad impacts on regional network level of service, which would influence, generation, distribution, mode choice, route assignment, and the use of regional traffic microsimulation within the integrated model provides a more robust platform or estimating these impacts.
- Travel Demand Management: Travel demand management strategies are typically aimed at changing travel behavior to reduce congestion and improve mobility. For example, these policies may seek to increase the frequency and numbers of people who work at home, induce workers to adjust their schedules to travel in off-peak, less congested conditions, or increase the number of people who carpool to work. These policies are most appropriately addressed at a regional scale because of their expected impact on performance on regionally significant / congested facilities. Although use of a detailed regional-scale model of traffic dynamics is not necessarily required, the microsimulation may provide a better tool for assessing the impact of these policies.
- Operations: The planning+operations model can support the analysis of bottleneck improvements, such as the addition of new signals or signs, adjustment of signal timing and phasing, or implementation of ramp meters. Although these are often tested at a local scale using network assignment tools while holding the other choice dimensions fixed, more extensive bottleneck or other operational improvements, such as those applied across an extensive coordinated signal systems within a corridor or at a regional scale, may be appropriately tested using the full integrated planning+operations application mode described below.
- Greenhouse Gas: Strategies to reduce greenhouse gases (GHG) may include both land use and transportation improvements. For this project, we focus on the transportation strategies. Use of the Microsimulator integrated with DaySim in the planning+operations model provides the greatest sensitivity to GHG reduction strategies such as using pricing strategies to reduce VMT, or increased fuel efficiency standards, as well as produces the most detailed estimates transportation measures used as input to the MOVES emissions and greenhouse gas estimation tools.

Prior to using the model system in an application context, it was necessary to first implement and link the DaySim and TRANSIMS model components. In the Jacksonville region, the model system was then calibrated and validated. These efforts, as well as the initial sensitivity testing of the model system in Burlington, are the focus of this document.

1.5 Conclusions

1.5.1 Model Implementation

1.5.1.1 Demand model data development

Developing the parcel-level inputs to the ABM was relatively straightforward. The cleaning of the employment data by NFTPO and FDOT significantly reduced the amount of time required to implement the model, although it was still necessary to make relatively crude updates to the employment data in one of the counties. The parcel file required some additional cleaning to establish reasonable totals of housing units and to address inconsistencies in the parcel geography. School enrollment, transit stops, intersection types, and parking data were all relatively easy to assemble from existing data sources. In addition, developing the synthetic population was relatively straightforward given the availability of the data and tools; however, the overall effort still required approximately six months.

Accommodating auxiliary demand within the integrated DaySim-TRANSIMS model system was achieved using readily available static methods from the region's trip-based model; however, revisions to these auxiliary demand components are necessary for a more spatially and temporally consistent integrated demand-supply model system. A drawback of the current implementation is that the auxiliary demand is currently "fixed" for each forecast year. That is, although this demand varies by forecast year, it is not affected by changes in network impedances. Ideally, the auxiliary models would be revised to provide sensitivity to changes in network performance.

1.5.1.2 Network model data development

Developing detailed and usable networks for microsimulation requires a significant level of effort. The TRANSIMS software comes with a wide array of tools to perform many network development tasks, and spatially detailed network data are widely available; however, users should expect to spend on the order of hundreds of hours debugging simulation networks: correcting topological errors, resolving attribute discontinuities, and coding intersection controls. The time-consuming effort involves iteratively evaluating, adjusting and testing the networks by running simulations. In addition, users face numerous challenges when attempting to develop future year or alternative network scenarios, a topic which is discussed in a subsequent section of this document.

1.5.1.3 Model Integration

Configuring Daysim to generate temporally, spatially, and behaviorally-detailed travel demand information for use in TRANSIMS was straightforward. Configuring TRANSIMS to generate the skims for input to DaySim was also straightforward. More sophisticated methods of providing TRANSIMS-based impedances to DaySim, such as implementing efficient multi-stage sampling of destinations (and corresponding impedances) at strategic points in the DaySim looping process, or tightly integrating DaySim and TRANSIMS so that DaySim can call TRANSIMS to extract the required measures quickly could potentially be implemented, though the runtime implications and required resource for development were felt to be prohibitive. It was felt that the current methods provide sufficient spatial and temporal detail. The network convergence equilibration effort revealed that the most effective convergence strategies were often the least acceptable to the larger DTA community, but were necessary in order to ensure sufficiently converged assignments within reasonable runtimes. Schedule consistency was identified as another measure of the soundness of a model solution. Extensive testing of the model system was necessary to determine the number of network assignment and model system iterations required to ensure that differences between alternative scenario model results were attributable to these policy and investments and not obscured by "noise" in the model system.

1.5.1.4 Model Enhancements

The enhancements made to the model system were necessary to improve the model system's sensitivity and to fulfill the goals of the SHRP2 C10A project. Updates to the DaySim model system were relatively straightforward to implement, although these updates were not fully completed until a new DaySim software architecture was implemented, which took significantly longer than expected. The updates to the TRANSIMS model components were much more extensive and involved much more time to implement, and many of these enhancements were under development during the C10A project. While these enhancements were necessary to fulfill project goals, they undoubtedly also resulted in schedule delays.

1.5.1.5 Model Application

The challenges in interacting with the model are primarily associated with debugging the model system. As has been repeatedly mentioned, the network simulation model is very sensitive to small scale network coding and parameter assumptions, and the network simulation is subject to frequent failures as input assumptions are being refined. It is critical that staff have the ability to understand and mine which data generated by the model system can illuminate the source of simulation problems, but also to make informed decisions about how to modify model inputs to achieve the proper model sensitivity. It is essential that model users have a basic understanding of Python in order to understand the overall model system flow, as well as robust data manipulation, statistical analysis, and GIS skills. It is not essential that model users know C# or C++, the development platforms used for DaySim and TRANSIMS, respectively.

The types of analysis that can be performed with the new model system are fundamentally different and more expansive than can be performed with a traditional model system, and the application and interpretation of model outputs must be thoughtfully considered. Use of the fully integrated model system would be most valuable when the proposed policies or strategies are expected to result in regional-scale changes in overall level of travel demand, changes in regional travelers' destination, mode, or time of day choices, and which expected to be significantly impacted by local-scale traffic dynamics.

The model system software can be flexibly deployed on hardware running either Windows or Linux, and the implementation can be scaled or configured to reflect available hardware resources. In order to avoid these long runtimes it also is possible to use the model in different application modes, as described earlier in this chapter. Although many DaySim and TRANSIMS tools exist to assist in data preparation and coding, the model system is highly sensitive to alternative configurations of the model system and to small scale coding issues, and it can take anywhere from an hour to many weeks to generate plausible alternative scenarios.

1.5.2 Model System Calibration / Validation

Transferring the DaySim activity-based demand component from Sacramento to Jacksonville radically reduced the amount of time required to implement the activity-based demand model component of the model system. Additional calibration and validation of some of the subcomponents of the model, such as the daily activity pattern component of DaySim, or the refinement of TRANSIMS networks was necessary in order to improve model performance. However, a number of the models required little, if any, recalibration. Use of the NHTS as the primary observed data source for developing demand model calibration targets was somewhat challenging given the limited weekday sample size in the Jacksonville region and other data completeness issues. Ultimately, more time was spent refining and validating the roadway networks than refining the calibration of the DaySim demand model components. This reflects the fact that network microsimulation models are significantly more sensitive to network coding assumptions and it simply requires more time to identify and resolve these issues.

1.5.3 Model Sensitivity Testing

Travel demand forecasting model systems are only able to test the effects of policies and assumptions which have been explicitly included when designing and implementing the model system, and are not intrinsically sensitive to the increasingly broad range of transportation policies and improvements of interest to decision-makers. While most regional models are sensitive to large-scale assumptions about land use and demographics, few are sensitive to more detailed assumptions about pricing policies, or to traffic or travel demand management strategies. Even where models have the capability to address these types of policies, they are typically not sufficiently sensitive to the dynamic interplay between travel behavior and network conditions by time-of-day, and are unable to reasonably represent the effects of road pricing, travel demand management, and other policies. Sensitivity testing of model systems involves the evaluation of the effects of changes in model inputs on model outputs. The Burlington implementation of the C10A model system was subjected to a set of sensitivity tests designed to illustrate the unique capabilities of the model system and included pricing, travel demand management, and operations.

1.5.3.1 Pricing

Two types of pricing tests were evaluated as part of this effort. In the first, a number of scenarios were defined in which freeway tolls varied by time of day. In the second, a number of scenarios were defined in which auto operating costs were modified from the “baseline” condition. For the first type of sensitivity tests, a set of three scenarios were evaluated and compared to the baseline alternative. In the baseline alternative, no costs were assessed at any time, while in the three scenarios different fixed per miles charges that varied by time-of-day were evaluated. The expected responses to these policies – that travel would decline during tolled periods and on tolled facilities, and that there would be differences in these changes by activity purpose - were all observed in the model system outputs. Interestingly, in all three pricing scenarios there were pronounced increases in travel demand during the evenings, suggesting that travelers are rescheduling activities to occur when there are no tolls, as well as fewer scheduling constraints such as are present during the midday. It was also observed in these tests that total trips by time-of-day affects different purposes. For example, work-related travel was relatively unaffected, but the social/recreation-related travel shifted noticeably out of the peaks and into the evening. Finally, the network-based total delay was higher than the base in all scenarios, as the tolling induces travelers to shift onto more capacity-constrained surface facilities.

For the second set of pricing sensitivity tests, a set of 3 auto operating cost scenarios were evaluated and compared against the baseline alternative. The baseline alternative assumed a cost of $0.12/mile, while the alternatives assumed charges as low as $0.06/mile and as high as and $0.60/mile. These tests confirmed that when auto operating costs decline the share of households choosing to maintain zero vehicles also declines, and as the costs increase, the share of zero-vehicle households also increases. However, these changes were relatively modest. The results also showed small changes in regional tour frequency by purpose, although these shifts did not result in significant changes in network performance or congestion.

1.5.3.2 Travel Demand Management

TDM approaches incorporate a wide range of strategies aimed at changing travel behavior to reduce congestion and improve mobility. The sensitivity testing for C10A focused on assessing the impacts of a flexible work schedule in which all workers worked fewer days but longer hours on those days. The overall time spent in work activities was held fixed. The model results were consistent with expectations based on the structure and linkages of the DaySim and TRANSIMS model. In general, overall levels of activity generation were lower, although the declines in work-related travel were offset by increases in travel for discretionary purposes. The model produced shifts in the distribution of travel by time-of-day due to the lengthened workday, although as expected, changes in the destination and mode choices were relatively small. This test did reveal noticeable changes in network performance, with reduced congestion across all facility types throughout most of the day, with a slight increase in congestion in the evening, which reflects both the later return times from work, as well as increased participation in discretionary activities in the evening.

1.5.3.3 Operations

The sensitivity testing was focused on a scenario in which signals were coordinated using TRANSIMS tools along three primary regional corridors, with the goal of reducing bottlenecks and improving the overall traffic flow. The DaySim-TRANSIMS model system provides sensitivity to these improvements, which traditional travel demand forecast models cannot typically represent the due to their linkage with traditional static network assignment methods that lack detailed network operation attributes and have coarse temporal resolution. The initial model results showed some reductions in delay by facility type, particularly during the peak periods. However, closer inspection of the speed profiles along the three targeted corridor showed more mixed results, with the signal progression producing better speeds in some corridor directions, and worse speeds in other corridor directions. As has been noted by others, the sensitivity of DTA and traffic microsimulation models to these detailed inputs suggest distinct challenges when attempting to incorporate these assumptions in a forecasting mode, especially if at a regional scale. Of all the scenarios evaluated as part of this sensitivity testing, the signal progression scenario required the greatest amount of time, and resulted in the least interpretable results.

1.5.3.4 Disaggregate framework

Because both the demand and the supply components of the model system are fully disaggregate, it is possible to trace the impacts of policies and investments on individual travelers from long term choices such as usual work locations all the way down to the specific paths taken by each individual traveler on a second-by-second basis. Although disaggregate model results are not reported, this disaggregate framework provides tremendous flexibility for aggregating model results for specific travel markets or communities of concerns, and is useful for debugging, calibration and refining model sensitivity. Also, note that random simulation variation did not compromise the ability to use the model system, provided a sufficient degree of convergence was achieved both within the network assignment and for the model system overall.

Overall, the new model system is more sensitive to a wider range of policies than a traditional travel demand model system, and this sensitivity is further enhanced by the detailed representation of temporal dimension, as well as the increased behavioral and spatial detail. In addition, the model system produces a wider range of statistics of interest to decision-makers.

Extracting, managing, and interpreting these results was not difficult; however, the level of effort required to effectively test different types of improvements varied widely, from as little as an hour to as more than a week. It was relatively easy to use the model to evaluate the pricing and TDM scenarios, requiring straightforward adjustments to network coding or to model coefficients. Using the model to evaluate the operational scenarios required significantly more effort due to the sensitivity of the network simulation to different signal coordination and timing assumptions. This level of effort would undoubtedly increase if more extensive changes to operational assumptions were required. In addition, even with the additional effort the results produced by the model system did not seem as intuitive as the results of the other scenario tests.

2.0 MODEL IMPLEMENTATION

The following sections describe the data requirements and steps necessary in order to implement the integrated model system. First, all key data inputs and tools are identified and described, DaySim model requires data that reflect a wide variety of factors that influence travel decisions, including socioeconomic, employment and school information, transportation network level of service, and urban form attributes. Much of this data is developed and applied at the detailed parcel-level, which enhances the model's sensitivity

2.1 Data Development

The DaySim model requires data that reflect a wide variety of factors that influence travel decisions, including socioeconomic, employment and school information, transportation network level of service, and urban form attributes. Much of this data is developed and applied at the detailed parcel-level, which enhances the model's sensitivity but which also increases the data development, maintenance and update requirements. The DaySim data inputs are discussed in the following sections.

2.2 Synthetic Population

Prior to applying the DaySim models in Jacksonville and Burlington, it is necessary to first develop a "synthetic population" of these regions' residents. This synthetic population is a list of households and persons that is based on observed or forecast distributions of socioeconomic attributes and created by sampling detailed Census microdata. This list functions as the basis for all subsequent choice-making simulated in the model system. All base year 2005 data required to develop the synthetic population using the DaySim population generation component are available from the Census (American Community Survey Public Use Microdata Sample (PUMS) and Decennial PUMS), Northeast Florida Regional Planning Model (NERPM) model inputs, and the National Household Travel Survey (NHTS).

2.2.1 PopGen

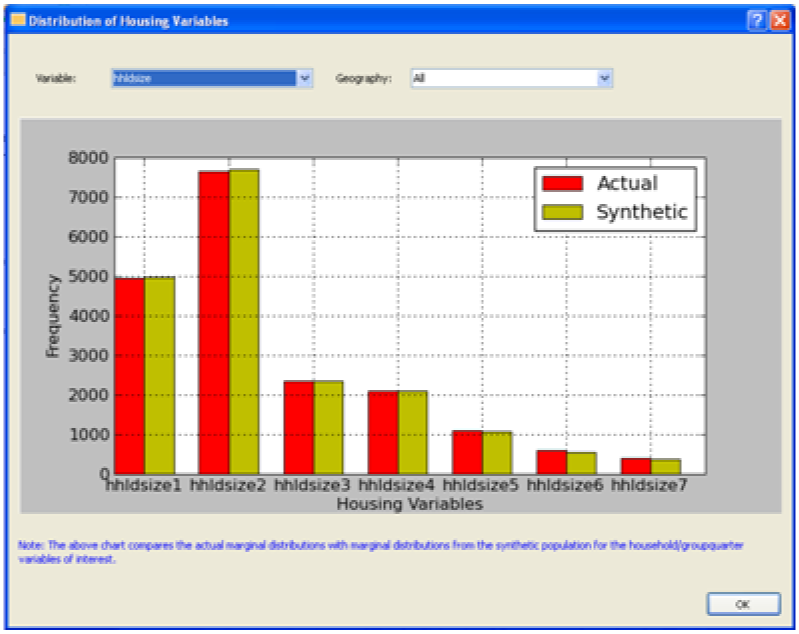
PopGen, a synthetic population generator developed at Arizona State University (ASU), was chosen for synthesizing the Jacksonville and Burlington populations. Synthetic population generators typically use census-based marginal distributions on household attributes to generate joint distributions on variables of interest using standard iterative proportional fitting (IPF) procedures. Households are then randomly drawn from an available sample in accordance with the joint distribution such that household-level attributes are matched perfectly. However, these traditional procedures typically operate at the household level and do not control for person-level attributes and joint distributions of personal characteristics. PopGen incorporates a heuristic approach which it uses to generate synthetic populations while matching both household-level and person-level characteristics of interest.

PopGen is a Python based software with an easy-to-use and flexible graphical user interface (GUI). It features a wizard-based project setup process using which a user can choose the region for population synthesis and specify the required inputs. Figure 6 shows the PopGen project setup wizard. It accommodates sample and control inputs from Census, ACS, and even region-specific sources such as household surveys and land-use model outputs. Populations can be synthesized with controls at various geographic resolutions such as Census block groups or travel analysis zones (TAZs). For Jacksonville and Burlington, the population is synthesized at the TAZ level and is subsequently allocated to individual parcels.

Figure 6: The PopGen Project Setup Wizard

Once the required inputs have been specified, PopGen imports them into tables in a MySQL database and works off of these tables to draw a synthetic population. After a population synthesis run the match between the synthetic population and control data can be checked using visualization features in PopGen. Figure 7 shows one such feature. If the synthetic population is found to be appropriate, tools in PopGen enable exporting it to specific file formats that can be used in travel demand microsimulation applications, such as DaySim.

Figure 7: PopGen Visualization



The following are the basic steps in preparing a synthetic population for microsimulation using DaySim:

1. Prepare the Control Data
2. Prepare the Sample Data
3. Synthesize the Population
4. Process the synthetic population for use in DaySim

Each of the above steps is described in detail in the following sections.

The Jacksonville synthetic sample population is comprised of three segments: permanent households and population, seasonal households and population, and group quarters population. These segments were established to reflect the differences in travel patterns associated with these sub-populations as well as to provide the ability to support seasonal analyses. For example, the seasonal population is generally older than the permanent population, has lower levels of workforce participation, and clusters in certain geographic areas. All of these attributes influence travel patterns and the demand for travel.

The Burlington synthetic sample population is comprised of two segments: permanent households and population and the group quarters population. Burlington does not have a significant seasonal population, and thus it was not necessary to develop a separate seasonal segment. However, Burlington does have a significant group quarters population comprised of University of Vermont students, so this segment was maintained.

2.2.2 Jacksonville Synthetic Population

2.2.2.1 Control Data

This section identifies the data sources and steps to prepare the control data for all three of the subpopulations: permanent resident households, seasonal households, and Non-institutionalized GQ residents.

Estimate the demographic distributions

The first step is to identify specific control variables of interest and derive demographic distributions for them. The control variables are attributes based on demographic distributions that are relevant to travel demand patterns. Control variables are specified for each of the three segments that comprise the synthetic population (permanent households and population, seasonal households and population, and GQ population).

Permanent Households: Table 2 and Table 3 respectively show the household and person control categories and the data sources used for PopGen synthesis of the permanent household population. For households, these categories include:

- The age of the head of household
- Household size
- Household workers
- Household income
- Presence of children

For persons, these categories include:

- Gender
- Age

The attributes are based primarily on Census Transportation Planning Package (CTPP) distributions, while the presence of children attribute is obtained from the Census Summary File 1 (SF1). Prior to working with the data, CTPP tables at the Census TAZ level were mapped to corresponding NERPM TAZs using these steps:

- A NERPM parcel centroid file was created from the parcel boundary shape file. This file also contains the NERPM TAZ for each parcel.
- The parcel centroid shape file was intersected with the Census TAZ shape file in ArcGIS and centroids of NERPM parcels were matched with Census TAZs. This creates a many-to-many correspondence between Census and NERPM TAZs.
- Using the total number of housing units in all the parcels in a NERPM TAZ and total number of housing units in all the parcels in a Census TAZ, the proportion of housing units from a Census TAZ that belong to a particular NERPM TAZ was calculated.

The number of households in various categories of the controls variables were aggregated at the Census TAZ level and distributed to NERPM TAZs based on the calculated proportions. The data were aggregated again at the NERPM TAZ level.

Similarly, since SF1 (Census Summary File) data is at the census block level, centroids for Census block polygons were mapped to NERPM TAZs using ArcGIS to obtain a block-TAZ correspondence table. By combining appropriate fields in the data tables, distributions of the various categories among the control variables chosen were obtained at the NERPM TAZ level.

Table 2. Household Control Data for Permanent and Seasonal Households

Household Attributes	Control Column	Categories	Source Data
Householder age	1	18-44	CTPP 1-70
	2	45-64	
	3	65+	
Household size, number of workers and income		Size cats 1-4: 1, 2, 3, 4+; Workers cats 1-3: 0,1,2+; Income cats 1-4: Under $30,000, $30,000-$59,999, $60,000-$99,999, $100,000 & over (Specified as joint distribution using a composite attribute):	CTPP 1-75
	4	Size1 Workers1 Income1	
	5	Size1 Workers1 Income2	
	6	Size1 Workers1 Income3	
	7	Size1 Workers1 Income4	
	8	Size1 Workers2 Income1	
	9	Size1 Workers2 Income2	
	10	Size1 Workers2 Income3	
	11	Size1 Workers2 Income4	
	12	Size2 Workers1 Income1	
	13	Size2 Workers1 Income2	
	14	Size2 Workers1 Income3	
	15	Size2 Workers1 Income4	
	16	Size2 Workers2 Income1	
	17	Size2 Workers2 Income2	
	18	Size2 Workers2 Income3	
	19	Size2 Workers2 Income4	
	20	Size2 Workers3 Income1	
	21	Size2 Workers3 Income2	
	22	Size2 Workers3 Income3	
	23	Size2 Workers3 Income4	
	24	Size3 Workers1 Income1	
	25	Size3 Workers1 Income2	
	26	Size3 Workers1 Income3	
	27	Size3 Workers1 Income4	
	28	Size3 Workers2 Income1	
	29	Size3 Workers2 Income2	
	30	Size3 Workers2 Income3	
	31	Size3 Workers2 Income4	
	32	Size3 Workers3 Income1	
	33	Size3 Workers3 Income2	
	34	Size3 Workers3 Income3	
	35	Size3 Workers3 Income4	
	36	Size4 Workers1 Income1	
	37	Size4 Workers1 Income2	
	38	Size4 Workers1 Income3	
	39	Size4 Workers1 Income4	
	40	Size4 Workers2 Income1	
	41	Size4 Workers2 Income2	
	42	Size4 Workers2 Income3	
	43	Size4 Workers2 Income4	
	44	Size4 Workers3 Income1	
	45	Size4 Workers3 Income2	
	46	Size4 Workers3 Income3	
	47	Size4 Workers3 Income4	
Presence of children under 18	48	yes	sf1-p19
	49	no	

Table 3. Person Control Data for Permanent and Seasonal Households

Person Attributes	Control Column	Categories	Source Data
Gender and age		gender cats 1&2: male/female; age cats 1-5: 0-15, 16-20, 21-44, 45-64, 65+	CTPP 1-51
	1	male age 0-15	
	2	male age 16-20	
	3	male age 21-44	
	4	male age 45-64	
	5	male age 65+	
	6	female age 0-15	
	7	female age 16-20	
	8	female age 21-44	
	9	female age 45-64	
	10	female age 65+	

Seasonal Households: For the seasonal population, the control categories are the same as for permanent households. However, the base year control values come directly from the seasonal households in the statewide NHTS survey sample, conducted in 2008 and 2009, which includes 530 households that reported living in Florida for 8 or less months per year. Of these households, 463 provided income information. The demographic distribution of this statewide sample is assumed to apply to all TAZs in the model area due to the unavailability of reliable seasonal population attributes at detailed geographic levels. However, the seasonal population is clustered in certain areas, such as along the coast.

For population synthesis, all dollars are normalized to represent 1999 dollars as closely as possible, since these were used in the 2000 Census that supplies PUMS and control table data. The NHTS survey data is recorded in categories of nominal 2007 or 2008 dollars (5K increments to 80K, then 80-100K, and above 100K), so it is necessary to place each one of these categories within one of the four 1999 income categories used for population synthesis (under 30K, 30 to under 60K, 60 to under 100K, 100K+). To do this the GDP deflator (0.817) was used to inflate the 1999 synthesis categories to 2007 values (under 36.7K to under 73.4K, 73.4K to under 122.4K, and 122.4K+), so that the recorded category of each household could be placed in the best synthesis category. Because the survey data's top category is only '$100K+ (2007 dollars)', it is impossible to make an accurate assignment of high income survey respondents between the top two synthesis categories, where the threshold between them is '100K (1999 dollars)'. The best option was to assign them all to the top income category '$100K+ (1999 dollars)'.

Non-institutionalized Group Quarters: Table 4 shows the proposed control categories for the GQ residents.

Table 4. Control data for Non-institutionalized Group Quarters residents

Household/Person Attributes	Control Column	Categories	Source Data
Age	1	under 18	2000 SF1-p38
	2	18-64	
	3	65+	

The distribution is extremely simple because of limited census table data for GQ residents. However, the age distribution will help PopGen to properly locate two important GQ subpopulations: college students and retirement center residents. The control information is so simple that an IPF procedure would not be necessary. However, in order to use PopGen to generate the sample, PopGen can be set up to run only the household level IPF, which will converge quickly, and avoid entirely the person-level IPF and the IPU procedures.

Estimate the number of households and persons in each TAZ

The number of households and people resident in each TAZ in 2005 are required as control totals for both the permanent and seasonal populations. The final control total required to synthesize the population is the number of GQ residents.

The total number of permanent households and seasonal households for 2005 at the TAZ level were obtained by combining NERPM model data on permanent and seasonal housing occupancy with parcel level estimates of housing units. The development of the parcel level estimates of housing units is described in section 2.3.1. The NERPM model demographic data includes TAZ level data on the number of housing units in a TAZ, the proportion of those households that are seasonally occupied or vacant and the proportion that are vacant. The following formulas were used to derive the number of households. They produced a total of 479,250 permanent households and 35,339 seasonal households.

PHHP = (SFDU * (1-SFSEAS/100) + MFDU * (1-MFSEAS/100))/(SFDU+MFDU)

PHH = ParcelHU * PHHP

SHHP = (SFDU * (SFSEAS-SFVAC)/100 + MFDU * (MFSEAS-MFVAC/100))/(SFDU+MFDU)

SHH = ParcelHU * SHHP

Where:

PHHP = Permanent household proportion

SFDU = Single family dwelling units (NERPM data)

SFSEAS = Percentage of seasonal or vacant single family dwelling units (NERPM data)

MFDU = Multi family dwelling units (NERPM data)

MFSEAS = Percentage of seasonal or vacant multi family dwelling units (NERPM data)

PHH = Permanent households

ParcelHU = Housing unit estimates from parcel data

SHHP = Seasonal household proportion

SFVAC = Percentage of vacant single family dwelling units (NERPM data)

MFVAC = Percentage of vacant multi family dwelling units (NERPM data)

SHH = Seasonal households

The permanent population controls are based on the total number of persons by county from the Census population estimated data. According to these data, the July 1 2005 population of the 4 county model area was 1,223,279. This includes GQ residents but is assumed to not include seasonal residents. GQ residents were separated out by estimating their 2005 population by interpolating between the number of GQ residents according to the 2000 Census, 20,122, and the number according to the 2006-2008 ACS, 21,047, which gives 20,783 GQ residents. The total permanent population according to the Census is 1,202,496. The county level permanent population totals were used to calculate an average household size number for the highest household size category (4+ people) for each county. This was applied to the TAZ level household size distribution, from CTPP Table 1-62, and the number of permanent households, to calculate the number of permanent residents in each TAZ.

The average seasonal household size was calculated from the NHTS data using the ratio of the total number of seasonal persons in the sample and households in the sample. To calculate the total number of persons in the seasonal population, this average household size was multiplied by the total number of seasonal households in each TAZ. This calculation resulted in 63,611 seasonal residents.

The total number of non-institutional GQ residents for the base year was estimated using the total GQ population estimated as described above and data from Census 2000 SF1 (Table P37) which identified the proportion of GQ residents who are classed as non-institutional (this distinction is important for travel modeling as institutionalized GQ residents, such as prisoners in jails, do not travel outside of their institution). The number of non-institutional GQ residents is 10,813. The county level estimates were assigned to TAZs based on the number of GQ housing units according to parcel data.

The demographics distributions described above were rescaled to match the estimated number of households and people resident in each TAZ.

Reformat control data to PopGen specifications

For permanent residents, two PopGen "Marginals Files" are needed with 49 household controls in one file and 10 personal controls in another, as shown in Table 2 and Table 3. The layout of the marginal input file required by PopGen is shown in Table 5. This file begins with four mandatory fields – state, county, tract, and bg, with bg interpreted as TAZ. After that there is a column for each control category, with entries representing the number of households (or persons) within the category for each TAZ. There are two header rows (the column name in row 1 and the data type of the field in row 2) followed by one of control data items for each TAZ.

Table 5. PopGen Household Marginal File Layout

state	County	Tract	Bg	<hhvar1cat1>	<hhvar1cat2>	…
<variabletype>	<variabletype>	<variabletype>	<variabletype>	<variabletype>	<variabletype>	
<data>	<data>	<data>	<data>	<data>	<data>	
…	…	…	…	…	…	

Household and person marginals files with the same control categories are needed for seasonal residents. For GQ population, the three controls in Table 4 need to be included in a household marginals file, but no person marginals file is required.

All of the above steps have been coded in an R-script. R[1] is open-source software for statistical computing and graphics. It is also an efficient tool for data manipulation and processing. The inputs required to the R-script are correspondences between Census TAZs, NERPM TAZs (for Jacksonville) and DaySim TAZs (DaySim TAZs are renumbered NERPM TAZs since the DaySim software requires external zones to be listed first), CTPP tables in CSV format, SF1 data with NERPM TAZ mapped (also in CSV format), NERPM zonal data (dbf format), demographic distributions of seasonal households from NHTS, and group quarter population totals by county from Census SF1.

[1] http://www.r-project.org/

On synthesizing the population it was found that there was a considerable overestimation in the total number of workers when compared to the employment in the region, accounting for in- and out-commuting. The household distribution by number of workers was found to be influencing the higher number of workers. Hence, the demographic distributions obtained from Census data were adjusted at the county level to match those obtained from ACS 2005-07 which had different proportions of households by workforce participation.

2.2.2.2 Sample Data

The household sample provides the households and person records that will be drawn into the synthetic population. In PopGen, it also provides the multi-dimensional attribute seed distribution for the iterative proportion fitting (IPF) procedures used in PopGen. Since the distribution does not depend primarily on the household sample, but rather on the controls described above, it is not necessary to have a sample that exactly represents the distribution. However, it is desirable for the sample to include many households of the types found in the region that are included in the synthetic distribution. So it is desirable to have a large sample from which to draw. Typically the census PUMS of each Public Use Microdata Area (PUMA) serves as the sample for all smaller geographical units included in the PUMA. Now that the ACS PUMS is available, it is possible to use either or both the 2000 census PUMS or ACS PUMS because the PUMA definitions and the definitions of the PUMS data items used by DaySim are essentially the same for the 2000 census and ACS PUMS. Both these PUMS data sources have been combined to create the sample file.

Table 6 lists the PUMAs that cover the model area. The PUMS samples for permanent and seasonal households include all occupied housing and person records from these PUMAs. The sample for Non-institutionalized GQ residents includes only those housing and person records, from these PUMAs, that represent the Non-Institutionalized GQ Population. For 2000 PUMS, these have housing record UNITTYPE=2, and for 2006-2008 ACS they have housing record TYPE=3. If the resulting sample is quite small, then it may be necessary to combine all these GQ records into a single sample that is used for all PUMAs, and perhaps even to add GQ records from other PUMAs in the state.

Table 6. Jacksonville PUMAs

PUMA ID	Description
1300	Clay County
1101	parts of Duval and Nassau Counties
1102	part of Duval County
1103	part of Duval County
1104	part of Duval County
1105	part of Duval County
1106	part of Duval County
1107	part of Duval County
1200	St. Johns County

Table 7 lays out the data elements needed for the PopGen Input Sample Files, including the items required by PopGen, the items corresponding to control variable categories, and the items needed by DaySim. As specified here, it includes data items corresponding to the controls required for the generation of all three synthetic subpopulations.

Table 7: The PopGen Input Sample File Data Items

Data item and Description	Values	Control var	ACS 2006-2008 item	Census 2000 5% PUMS item
Household Sample File				
State			ST	STATE
Pumano			PUMA	PUMA5
Hhid (same as serialno)				
Serialno				
Household size			NP	PERSONS
Number people in related family			NPF	NPF
Household income[1]	in dollars		HINCP (PINCP for GQ)	HINC (INCTOT for GQ)
CTPP1-70 category: age of householder		HH		
CTPP 1-75 category: household size, income[2] and number of workers		HH		
SF1-P19 category: presence of children in HH		HH		
SF1-P38 category: age		GQ		
Person Sample File				
State			ST	STATE
Pumano			PUMA	PUMA5
Hhid				
serialno (same as serialno)				
Pnum				
Gender	1-male, 2-female		SEX	SEX
Age	years		AGEP	AGE
Grade in school	1:pre-K 2:K 3:Grade 1-4 4:Grade 5-8 5:Grade 9-12 6:Undergrad 7:Grad/Prof school		SCHG	GRADE
Hours worked per week			WKHP	HOURS
CTPP1-51 category: gender and age		HH		

1. Income data from separate years is deflated to 1999 dollars, used by 2000 CTPP 1-75 and by DaySim. Household income is not available in PUMS for GQ residents, but total personal income is available in the person record and was used instead.

Table 8 shows the exact format of the input sample file as required by PopGen. An input sample file contains four mandatory fields (State, Pumano, Hhid, Serialno) followed by population attributes. The "Serialno" and "Hhid" are identical IDs for the sample housing unit indexed at 1[1].

[1] The duplication is a legacy of an earlier version of PopGen. The current version only requires one unique ID but the code still requires two fields to be present in the input file.

Table 8: PopGen Household Sample File Layout

state	pumano	hhid	Serialno	<hhvariable1>	<hhvariable2>	…
<variabletype>	<variabletype>	<variabletype>	<variabletype>	<variabletype>	<variabletype>	
<data>	<data>	<data>	<data>	<data>	<data>	
…	…	…	…	…	…	

2.2.2.3 Population Synthesis

In this step, PopGen is run separately for each of the three subpopulations, using the above specified input control and sample files. In addition to the control and sample files, PopGen requires a geography correspondence file (shown in Table 9). This file has one row per TAZ, and associates the TAZ with the PUMA (and other larger geographies) to which it belongs. For Jacksonville, the correspondence file has been prepared and is named Geocorr.csv.

Table 9: PopGen Geographic Correspondence File Layout

county	tract	bg	state	pumano	stateabb	countyname
<vartype>	<vartype>	<vartype>	<vartype>	<vartype>	<vartype>	<vartype>
<data>	<data>	<data>	<data>	<data>	<data>	<data>
…	…	…	…	…	…	…

After PopGen is run, output population files are exported. Table 10 lays out the data elements needed for the PopGen Synthetic Population Output Files. Compared to the sample input data items, it drops the items required for sampling, and adds identification numbers required by PopGen.

Table 10: PopGen Synthetic Population Output File Data Items

Data item and Description	Values
Household Characteristics	
State	
County	
Tract	
bg (TAZ)	
Hhid	
Serialno	
Frequency	
hhuniqueID	
Household size	
Number people in related family	
Household income	in dollars
Person Characteristics	
State	
County	
Tract	
bg (TAZ)	
hhid	
serialno	
pnum	
frequency	
personuniqueID	
Gender	1-male, 2-female
Age	years
Grade in school	1:pre-K (age 3+) 2:K 3:Grade 1-4 4:Grade 5-8 5:Grade 9-12 6:Undergrad 7: Grad/Prof school
Hours worked per week	

2.2.2.4 Synthetic Population DaySim Integration

DaySim currently generates and reads the synthetic population in the form of person records, with household data repeated in every person record. In addition, DaySim operates at the parcel-level, while PopGen creates the synthetic population at the larger TAZ level. Therefore, a DaySim population conversion/parcel allocation utility has been created that reads in PopGen population files, associates household attributes with persons, and allocates the households in the synthetic population to parcels. It then outputs a combined synthetic population file (dbf) in a format required by DaySim. The primary inputs to this utility are six PopGen output files (household and person files for three population groups) and a TAZ controls file along with DaySim's regular parcel data input file (see Section 2.3). The TAZ controls file is an input of permanent households, seasonal households and non-institutionalized GQ residents living in each TAZ. The file format is shown in Table 11.

Table 11: DaySim Synthetic Population Input Data Items

Label	Format	Definition
TAZ	Integer	Zone number
HHPerm_ZC	Float	Permanent households living in TAZ
HHSeas_ZC	Float	Seasonal households living in TAZ
GQUnitsZC	Float	Non-institutionalized GQ residents in TAZ

Note: These data must be in comma delimited text format (.csv), with a header row, in the order specified. DaySim reads them in as integers.

Table 12 shows the format of the PopGen population household files that are need by the DaySim utility. Whether 1, 2 or 3 input files are required depends on the segmentation of the synthetic population and the associated settings in the control file. For Jacksonville, three files are used, one for each population segment.

Table 12: Input files format of households generated from PopGen: permanent, seasonal and/or GQ

Label	Definition
state	State of residence
county	County of residence
tract	Tract of residence
bg	TAZ of residence
hhid	Household ID (generated by PopGen; DaySim reassigns household number)
serialno	Serial Number (generated by PopGen)
frequency	Number of households represented (DaySim assigns this number of households)
HINC	Household income ($)
Hhsize	Number of persons in household
NPF	Number of persons part of family

Note: These data must be in comma delimited text format (.csv), WITHOUT a header row, in the order specified. DaySim reads them in as integers.

Table 13 shows the formats of PopGen population person files that are need by the DaySim utility. A PopGen population person file is required for each of the population segments; three in the case of Jacksonville.

Table 13: Input files format of persons generated from PopGen: permanent, seasonal and/or GQ

Label	Definition
pstate	State of residence
pcounty	County of residence
ptract	Tract of residence
pbg	TAZ of residence
phhid	Household ID (generated by PopGen; DaySim reassigns household number)
pserialno	Serial Number (generated by PopGen)
pnum	Person number within household (DaySim reassign person number)
pfrequency	Number of households represented (DaySim assigns this number of households) *
Age	Age in years
Gender	Gender: 1-male, 2-female
GradeCat	0:non-student 1:pre-K (age 3+) 2:K 3:Grade 1-4 4:Grade 5-8 5:Grade 9-12 6:Undergrad 7:Grad/Prof school
Hours	Hours worked per week

Note: These data must be in comma delimited text format (.csv), WITHOUT a header row, in the order specified. Each person file must be in the same household order as its corresponding household file. DaySim reads the data as integers.

The utility also creates three additional data items (shown in Table 14) required by DaySim during microsimulation. It creates binary variables for each person, indicating whether or not they are a worker and/or a student. It then assigns the household to a specific parcel within the TAZ to which the synthetic household was assigned by PopGen. Permanent and seasonal households in a TAZ are combined into one large group and allocated to parcels based on the availability of dwelling units. GQ residents are allocated to parcels containing GQ dwelling units, which are identified in the parcel data input file separately from dwelling units for permanent and seasonal households.

Table 14: DaySim Synthetic Population Derived Data Items

Data item and Description	Values	ACS 2006-2008 item	Census 2000 5% PUMS item
Household Characteristics			
Parcel	parcelid		
Person Characteristics			
Worker indicator	yes/no	WKHP>0	HOURS>0
Student indicator	yes/no	SCHG not blank	grade>0

2.2.2.5 Synthetic Population Validation

For Jacksonville, the synthetic population generated using PopGen was validated across the different control dimensions (both household and person) used in order to ensure that the population matched the control variables. The validation was done separately for the three population groups that have been synthesized. Table 15 summarizes the differences in the total number of households and persons in the synthetic population and observed controls. Overall, the synthetic population has about 1.4% less persons. This can be considered reasonable given the total population of about 1.2 million persons in the modeling region.

Table 15: Synthetic Population Validation Summary

Population Group	HH Obs.	HH Syn	HH Diff	Per Obs.	Per Syn	Per Diff
Permanent	479,250	479,298	0.01%	1,202,855	1,184,800	-1.50%
Seasonal	35,339	35,367	0.08%	63,611	64,185	0.90%
Group Quarters	10,813	10,823	0.10%	10,813	10,823	0.10%
Total	**525,402**	**525,488**	**0.02%**	**1,277,279**	**1,259,808**	**-1.37%**

Figure 8 shows a comparison of the household size distribution from the control data and synthetic population to illustrate the degree of match achieved.

Table 16 and Table 17 show the comparison of household size distributions further disaggregated to the county level. Since the population was synthesized at the TAZ level, the county level match is close as expected. It can be seen that the proportion of seasonal households is quite low.

Figure 8: Jacksonville Regional Household Size Distribution Comparison

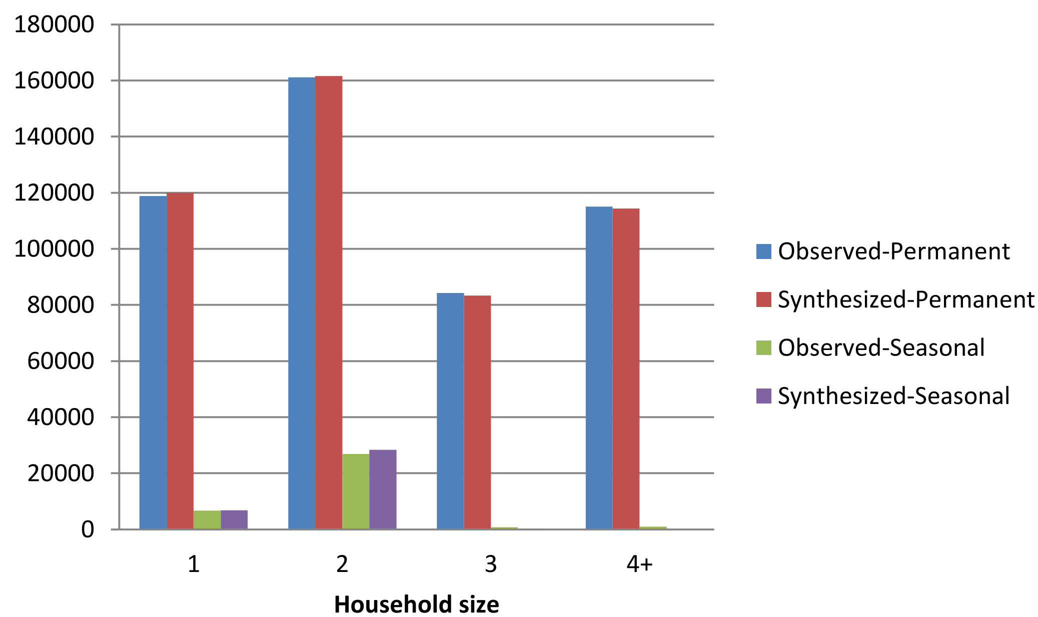


Table 16: Jacksonville Household Size Distributions for Permanent Households by County

Household size		1	2	3	4+	Total
Observed	Clay	10,489	20,296	12,058	18,508	61,352
	Duval	87,640	107,388	57,993	75,330	328,350
	Nassau	4,900	8,940	3,971	5,720	23,531
	St Johns	15,812	24,489	10,212	15,509	66,022
Synthesized	Clay	10,583	20,394	11,965	18,427	61,369
	Duval	88,402	107,661	57,412	74,877	328,352
	Nassau	4,974	8,992	3,914	5,665	23,545
	St Johns	15,942	24,548	10,103	15,439	66,032

Table 17: Jacksonville Household Size Distributions for Seasonal Households by County

Household size		1	2	3	4+	Total
Observed	Clay	583	2,331	66	86	3,066
	Duval	4,840	19,358	550	715	25,463
	Nassau	499	1,996	57	74	2,626
	St Johns	795	3,182	90	118	4,185
Synthesized	Clay	561	2,504	2	0	3,067
	Duval	4,934	20,406	120	23	25,483
	Nassau	493	2,104	21	8	2,626
	St Johns	797	3,370	18	6	4,191

The distributions of household workers for both permanent and seasonal population are shown in Figure 9. At a county level, household workers distributions are given in Table 18 and Table 19 for permanent and seasonal population respectively. The level of match is similarly close to that of the household size distribution.

Figure 9: Jacksonville Regional Household Workers Distribution Comparison

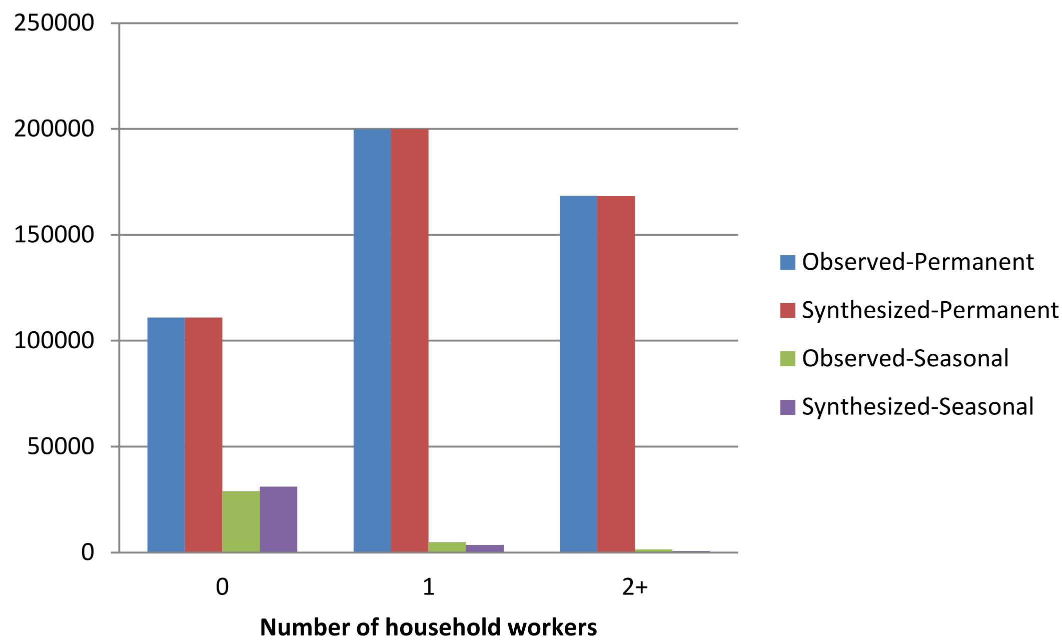


Table 18: Jacksonville Household Workers Distributions for Permanent Households by County

Household workers		0	1	2+	Total
Observed	Clay	12,231	24,171	24,949	61,352
	Duval	74,193	141,661	112,496	328,350
	Nassau	6,783	8,053	8,695	23,531
	St Johns	17,824	25,939	22,259	66,022
Synthesized	Clay	12,233	24,154	24,982	61,369
	Duval	74,116	141,841	112,395	328,352
	Nassau	6,810	8,053	8,682	23,545
	St Johns	17,871	25,963	22,198	66,032

Table 19: Jacksonville Household Workers Distributions for Seasonal Households by County

Household workers		0	1	2+	Total
Observed	Clay	2,510	430	126	3,066
	Duval	20,843	3,575	1,045	25,463
	Nassau	2,149	369	108	2,626
	St Johns	3,426	588	172	4,185
Synthesized	Clay	2,750	268	49	3,067
	Duval	22,335	2,618	530	25,483
	Nassau	2,315	258	53	2,626
	St Johns	3,703	400	88	4,191

The household income distributions for both permanent and seasonal population are shown in Figure 10. Table 20 and Table 21 illustrate that, for both permanent and seasonal households, the income distribution in the synthetic population is close to that of the observed control data at the county level.

Figure 10: Jacksonville Regional Household Income Distribution Comparison

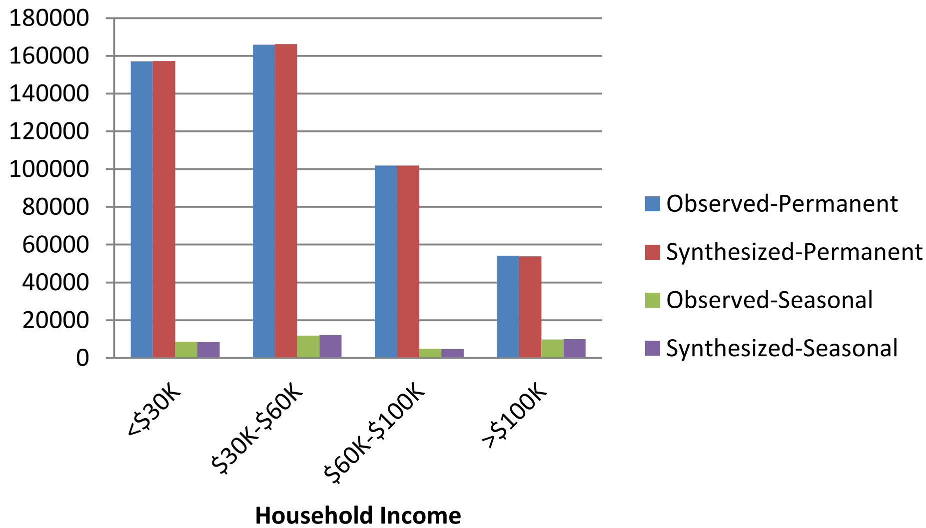


Table 20: Jacksonville Household Income Distribution for Permanent Households by County

Household income		**<$30K**	**$30K-$60K**	**$60K-$100K**	**>$100K**	**Total**
Observed	Clay	16,165	22,164	15,995	7,028	61,352
	Duval	112,892	115,895	66,066	33,497	328,350
	Nassau	7,560	7,917	5,454	2,601	23,531
	St Johns	20,515	19,976	14,448	11,082	66,022
Synthesized	Clay	16,146	22,239	16,005	6,979	61,369
	Duval	113,047	116,117	65,979	33,209	328,352
	Nassau	7,570	7,948	5,460	2,567	23,545
	St Johns	20,556	19,967	14,435	11,074	66,032

Table 21: Jacksonville Household Income Distribution for Seasonal Households by County

Household income		**<$30K**	**$30K-$60K**	**$60K-$100K**	**>$100K**	**Total**
Observed	Clay	748	1,033	430	854	3,066
	Duval	6,214	8,579	3,575	7,094	25,463
	Nassau	641	885	369	732	2,626
	St Johns	1,021	1,410	588	1,166	4,185
Synthesized	Clay	738	1,061	405	863	3,067
	Duval	6,149	8,718	3,412	7,204	25,483
	Nassau	643	896	353	734	2,626
	St Johns	998	1,450	566	1,177	4,191

In addition to evaluating the household validation, it is also important to look at the match among person attributes. The person-level attributes are one of the distinguishing features of the PopGen tool. Figure 11 shows that the person level attribute distribution of age in the synthesized population is a close match to the controls. Table 22 and Table 23 split the distributions of age further by counties. It can be observed that both the distributions and the county level total of number of persons have been fitted well.

Figure 11: Jacksonville Regional Age Distribution Comparison

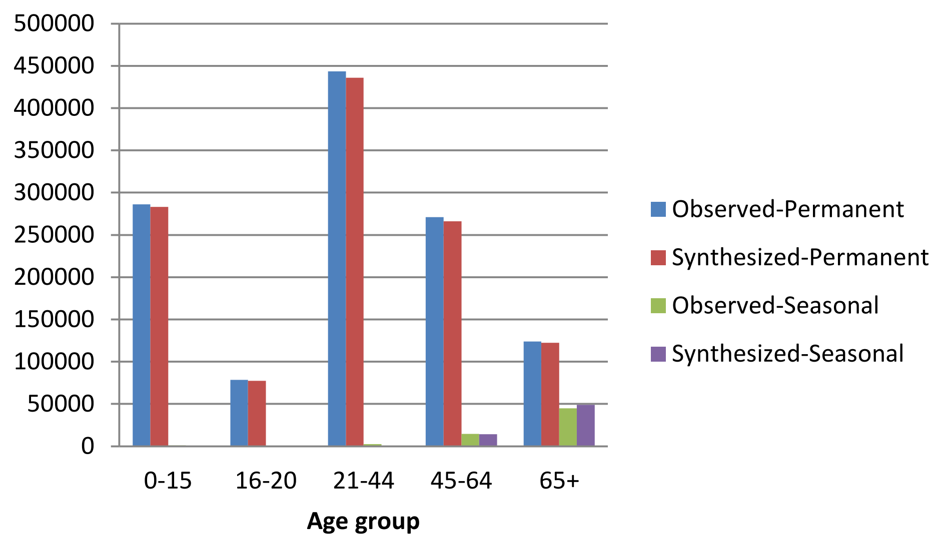


Table 22: Jacksonville Age Distribution Comparison for Permanent Population by County

Age group		0-15	16-20	21-44	45-64	65+	Total
Observed	Clay	43,330	11,951	59,705	39,152	14,317	168,454
	Duval	192,099	53,689	309,132	174,889	80,396	810,204
	Nassau	14,250	3,956	21,236	16,526	7,741	63,709
	St Johns	36,389	9,073	53,278	40,333	21,414	160,487
Synthesized	Clay	42,868	11,733	58,806	38,741	14,179	166,327
	Duval	190,461	52,956	303,452	171,421	79,067	797,357
	Nassau	13,426	3,715	19,795	15,583	7,295	59,814
	St Johns	36,493	9,107	53,681	40,325	21,696	161,302

Table 23: Jacksonville Age Distribution Comparison for Seasonal Population by County

Age group		0-15	16-20	21-44	45-64	65+	Total
Observed	Clay	98	58	202	1,267	3,893	5,518
	Duval	817	480	1,682	10,521	32,333	45,833
	Nassau	84	50	173	1,085	3,334	4,726
	St Johns	134	79	276	1,729	5,315	7,534
Synthesized	Clay	0	1	33	1,200	4,341	5,575
	Duval	23	26	447	10,424	35,278	46,198
	Nassau	10	5	59	1,050	3,672	4,796
	St Johns	8	3	71	1,690	5,844	7,616

Table 24 illustrates that for the GQ population, the age distribution in the synthetic population is representative of the observed marginal numbers in each of the four counties.

Table 24: Jacksonville Non-institutional GQ Population Age Distribution

Age group		18-44	45-64	65+	Total
Observed	Clay	0	121	638	758
	Duval	233	7,841	646	8,720
	Nassau	0	199	33	233
	St Johns	37	943	122	1,102
Synthesized	Clay	0	120	639	759
	Duval	235	7,847	648	8,730
	Nassau	0	199	33	232
	St Johns	37	945	120	1,102

2.2.3 Burlington Synthetic Population

2.2.3.1 Control Data

Estimate the demographic distributions

The first step is to identify specific control variables of interest that are relevant to the travel demand forecasting process. The control variables or attributes are identified for each of the two population segments, permanent households and group quarters population, separately. Since the choice of attributes for Jacksonville was not based on reasons specific to the geographic area, the same attributes were used for Burlington too.

Permanent Households: The household and person control and their categories shown in Table 2 and Table 3 for Jacksonville were also used for PopGen synthesis of the permanent household population in Burlington. The tables also show the specific data sources used to derive distributions for each of the variables. The household level control attributes include:

- The age of the head of household
- Household size
- Household workers
- Household income
- Presence of children

The person level control attributes include:

- Gender
- Age

Distributions for all of the control attributes except those for presence of children were derived from Census Transportation Planning Package (CTPP) tables. The distributions for the presence of children attribute was obtained from the Census Summary File 1 (SF1). Since the CTPP distributions were at the Census TAZ level, these needed to be converted to the CCMPO model TAZ level. The following steps were used create a correspondence between the Census and CCMPO model TAZs:

- A CCMPO model parcel centroid file was created from the parcel boundary shape file. This file also contains the CCMPO model TAZ for each parcel.

- The parcel centroid shape file was intersected with the Census TAZ shape file in ArcGIS and centroids of CCMPO parcels were matched with Census TAZs. This creates a many-to-many correspondence between Census and CCMPO TAZs.
- Using the total number of housing units in all the parcels in a CCMPO TAZ and total number of housing units in all the parcels in a Census TAZ, the proportion of housing units from a Census TAZ that belong to a particular CCMPO TAZ was calculated.

The number of households in various categories of the controls variables were aggregated at the Census TAZ level and distributed to CCMPO TAZs based on the calculated proportions. The data were aggregated again at the CCMPO TAZ level.

SF1 (Census Summary File) data is at the census block group level. A block group-CCMPO TAZ correspondence was similarly created using parcels as the go-between. Parcel centroids were mapped to Census block groups using ArcGIS and the block group level households in each category of the presence of children attribute were distributed in the same proportion to the parcel level. Aggregating the data then to CCMPO TAZ level resulted in the required distributions.

For population synthesis, all dollars are normalized to represent 1999 dollars as closely as possible, since these were used in the 2000 Census that supplies PUMS and control table data.

Non-institutionalized Group Quarters: The control categories used for group quarters population synthesis in Jacksonville were also used for Burlington, and are shown in Table 4. As in Jacksonville, The distribution is extremely simple because of limited census table data for group quarters residents. However, the age distribution helps PopGen to properly locate two important group quarters subpopulations: college students and retirement center residents. There is no IPF required and simple scaling would suffice to match this one-dimensional control. However, PopGen was run in a simplified mode to synthesize the group quarters population.

Estimate the number of households and persons in each TAZ

The total number of permanent households for 2005 at the TAZ level was obtained from the CCMPO model data. Since a separate population for seasonal households was not being synthesized, the additional step for estimating the total number of households at the seasonal level was not required. Permanent population controls were estimated in a rather straightforward manner. An approximate average of 4.5 persons was assumed for the highest household size category (households with 4 or more persons). This average along with the number of households in each size category resulted in an estimate of total number of persons by TAZ.

The total number of non-institutional group quarters residents for the base year was estimated using parcel level data from the CCMPO model. Specific parcels belonging to educational institutions were identified and number of group quarters units on each was aggregated up to the TAZ level.

The demographics distributions described above were rescaled to match the estimated number of households and people residing in each TAZ.

Reformat control data to PopGen specifications

The process for reformatting the control data in Burlington was similar to that performed in Jacksonville. For permanent residents, two PopGen “Marginals Files” are needed with 49 household controls in one file and 10 personal controls in another, as shown in Table 2 and

Table 3. The layout of the marginal input file required by PopGen is shown in Table 5. This file begins with four mandatory fields – state, county, tract, and bg, with bg interpreted as TAZ. After that there is a column for each control category, with entries representing the number of households (or persons) within the category for each TAZ. There are two header rows (the column name in row 1 and the data type of the field in row 2) followed by one of control data items for each TAZ.

For the group quarters population, the three controls in Table 4 need to be included in a household marginals file, but no person marginals file is required.

All of the above steps have been coded in an R-script for Burlington similar to that used for Jacksonville. The inputs required to the R-script are correspondences between Census TAZs, CCMPO TAZs and DaySim TAZs (DaySim TAZs are renumbered CCMPO TAZs since the DaySim software requires external zones to be listed first), CTPP tables in CSV format, SF1 data with CCMPO TAZ mapped (also in CSV format), CCMPO zonal data (dbf format), and group quarters population totals by parcel from CCMPO data.

2.2.3.2 Sample Data

The sample data represents household sample with detailed demographic attributes and is used to draw individual households to form the synthetic population. It is also used to provide a seed matrix of multi-dimensional control attributes for the IPF procedure in PopGen. It is desirable to have a large and socio-demographically diverse sample to draw the synthetic population from. For this reason, as in the case of Jacksonville, PUMS samples from both ACS and Census 2000 were combined to prepare the sample for population synthesis.

There is only one PUMA that covers the model area – PUMA 100. The PUMS sample for permanent households include all occupied housing and person records from this PUMA. The sample for Non-institutionalized GQ residents includes only those housing and person records, from this PUMA, that represent the Non-Institutionalized GQ Population. For 2000 PUMS, these have housing record UNITTYPE=2, and for 2006-2008 ACS they have housing record TYPE=3.

Table 7 lays out the data elements needed for the PopGen Input Sample Files, including the items required by PopGen, the items corresponding to control variable categories, and the items needed by DaySim. The data items are same for the two population segments developed in Burlington

Table 8 shows the exact format of the input sample file as required by PopGen for both Burlington and Jacksonville. An input sample file contains four mandatory fields (State, Pumano, Hhid, Serialno) followed by population attributes. The “Serialno” and “Hhid” are identical IDs for the sample housing unit indexed at 1[1].

2.2.3.3 Population Synthesis

In addition to the control and sample files, PopGen requires a geography correspondence file (shown in Table 9). As in Jacksonville, this file has one row per TAZ, and associates the TAZ with the PUMA (and other larger geographies) to which it belongs. For Burlington, since there is only one PUMA, all the TAZs are just mapped to PUMA 100.

Using all the above mentioned input files, PopGen is run separately for the two population subgroups – permanent and non-institutionalized group quarters. After PopGen is run, output

[1] The duplication is a legacy of an earlier version of PopGen. The current version only requires one unique ID but the code still requires two fields to be present in the input file.

population files are exported. The data elements needed for the PopGen Synthetic Population Output Files are identical in both Burlington and Jacksonville, and are show in Table 10.

2.2.3.4 Synthetic Population DaySim Integration

As described in the earlier Jacksonville population synthesis section, DaySim requires a synthetic population in the form of persons records with the household information attached to each of the person records. Also, DaySim operates at the parcel level which requires population synthesized at the TAZ level by PopGen to be allocated to individual parcels within a particular TAZ. For this purpose, a utility has been created in Delphi which randomly allocates households within a TAZ to individual parcel units. It then outputs a combined synthetic population file (dbf) in a format required by DaySim. The primary inputs to this utility are four PopGen output files (household and person files for two population groups) and a TAZ controls file along with DaySim's regular parcel data input file. Since this utility was created for Jacksonville which had three population segments, the TAZ controls file is an input of permanent households, seasonal households and non-institutionalized GQ residents living in each TAZ. The file format is shown in Table 11. The utility was used by specifying the number of seasonal households in all TAZs as zero.

Table 12 and Table 13 respectively show the formats of PopGen population household and person files that are need by the DaySim utility. A pair of PopGen population household and person files are required for each of the population segments; two in the case of Burlington.

The utility also creates three additional data items (shown in Table 14) required by DaySim during microsimulation. It creates binary variables for each person, indicating whether or not they are a worker and/or a student. It then assigns the household to a specific parcel within the TAZ to which the synthetic household was assigned by PopGen. Permanent households in a TAZ allocated to parcels based on the availability of dwelling units. GQ residents are allocated to parcels containing GQ dwelling units, which are identified in the parcel data input file separately from dwelling units for permanent households.

2.2.3.5 Synthetic Population Validation

The Burlington synthetic population generated using PopGen was validated across the different control dimensions (both household and person) used in order to ensure that the population matched the control variable distributions. The validation was done separately for the two population groups that have been synthesized. Table 25 summarizes the differences in the total number of households and persons in the synthetic population and observed controls. Overall, the synthetic population has just about 1.5% less persons than the observed data.

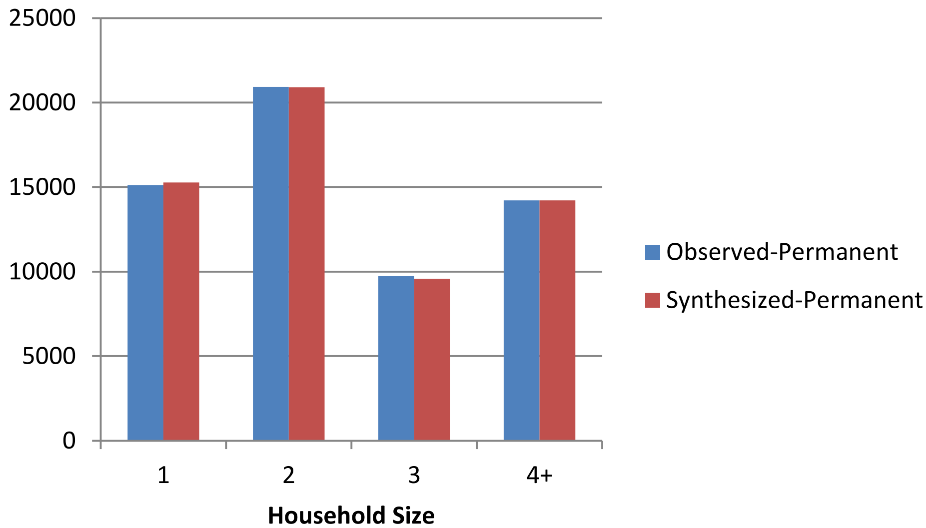
Table 25: Burlington Synthetic Population Validation Summary

Population Group	HH Obs.	HH Syn	HH Diff	Per Obs.	Per Syn	Per Diff
Permanent	59,975	59,975	0.00%	150,263	147,909	-1.57%
Group Quarters	5,474	5,477	0.05%	5,474	5,477	0.05%
Total	65,449	65,452	0.00%	155,737	153,386	-1.51%

Figure 12 shows a comparison of the household size distributions from the control data and synthetic population to illustrate the degree of match achieved. Since there is only one county in the model region (Chittenden County), this can be interpreted as a match at the county level.

Since the population was synthesized at the TAZ level, the county level match is close as expected.

Figure 12: Burlington Household Size Distribution Comparison



Similarly, Figure 13 and Figure 14 show comparisons of distributions for number of household workers and household income between observed and synthetic permanent populations. At the household level, the synthetic population attributes seem to match the observed distributions very closely.

Figure 13: Burlington Household Workers Distribution Comparison

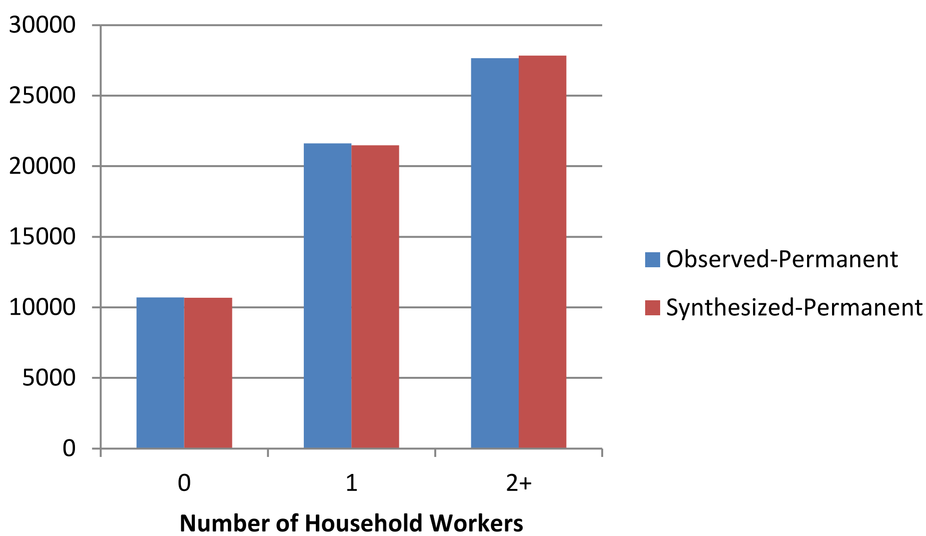
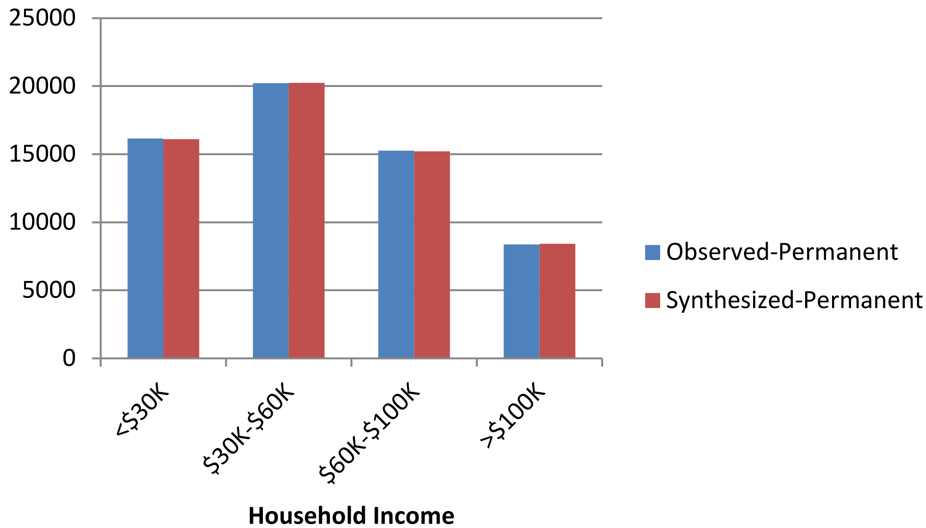
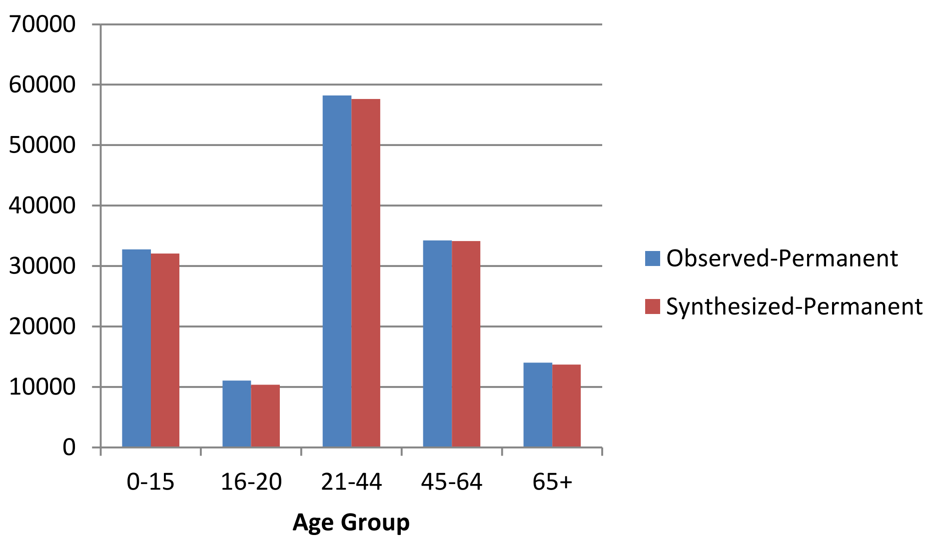


Figure 14: Burlington Household Income Distribution Comparison



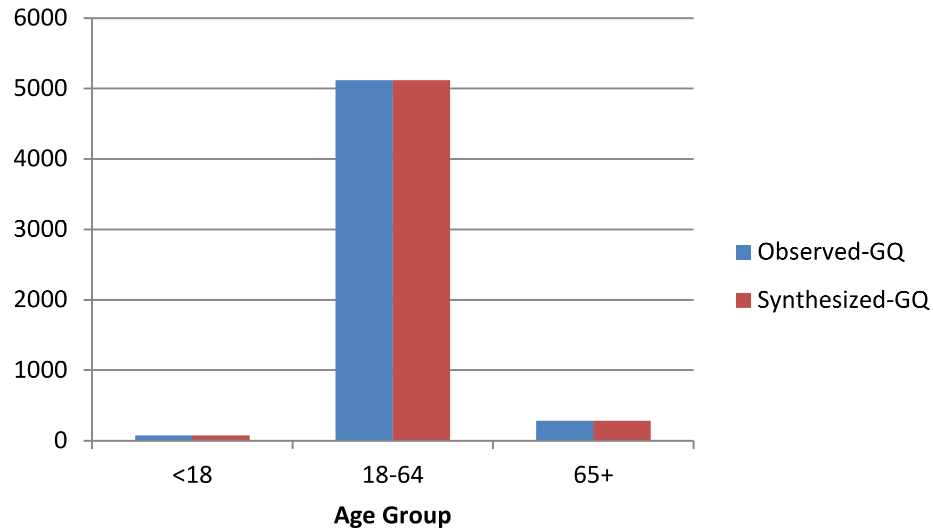
In addition to evaluating the match among household attributes, it is also important to compare the distributions of person attributes in the synthetic populations and observed data. Figure 15 shows that the distribution of person age in the synthetic population matches very well with that observed from Census data.

Figure 15: Burlington Person Age Distribution Comparison



Finally, it can be seen in Figure 16 that the distribution of age of group quarter population is almost the same as that in the observed data.

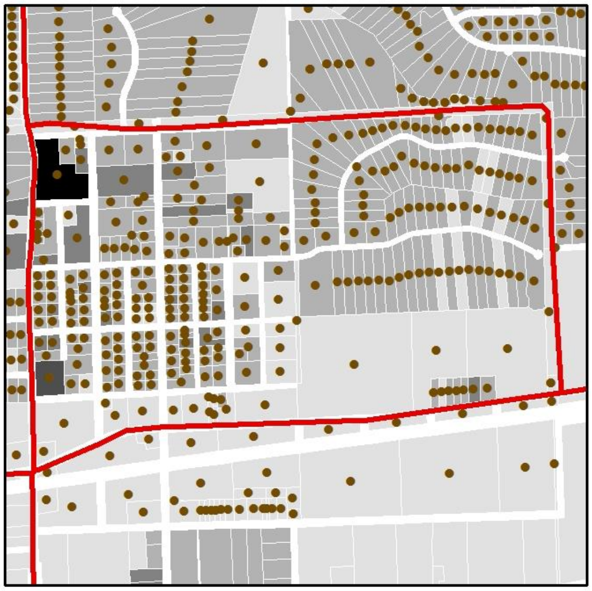
Figure 16: Burlington GQ Population Age Distribution Comparison



2.3 DaySim Parcel Data

A distinguishing feature of DaySim is that it uses parcels as one of the fundamental spatial units. The parcel data input file contains input data at the parcel level of detail, the more detailed of the two spatial levels at which DaySim input data are prepared. The less detailed spatial level uses TAZs; the TAZ input file is described below in Section 2.4. Figure 17 shows the relationship between TAZs (TAZ boundaries are shown in red) and parcels. The parcel polygons, which show the physical extent of the parcel, are shown in gray with the parcel boundaries shown as thin white lines. The area of the parcel is included in the parcel data input file in units of square feet in the "AREA_SQF" field. In this figure, darker gray colors indicate increasing numbers of housing units in the parcel; housing units numbers are recorded in the "HOUSESP" variable in the parcel data input file. The parcel centroids are shown as brown dots. The locations of the parcel centroids are described in the parcel file in the "X_COORD" and "Y_COORD" fields. Omitted from the parcel file (and shown as thick white lines) are highway and other rights of way. Parcels clearly allow a far more detailed, spatially disaggregate description of the land use in a region than TAZ-based models, but consequently result in the need to develop and manage larger quantities of data.

Figure 17: TAZ Boundaries, Parcel Polygons and Parcel Centroids



The parcel data input file is a dBase IV format file (.dbf) with one row of data per parcel. Table 26 shows the fields contained in the parcel data input file. The file begins with several fields that identify the parcel, and describe the physical location and size of the parcel, and then contains fields that describe the quantity of housing, school enrollment, and employment on the parcel and within a quarter mile and a half mile of the parcel. In addition, the parcel file contains information about urban form and the transportation system on and close to the parcel, including the proximity to transit stops and the price and supply of parking. The data sources and the development process of these fields are discussed in the following sections of this report.

Table 26: Parcel data input file format

Label*	Definition
PARCELID	Parcel ID number
X_COORD	X coordinate – state plane feet
Y_COORD	Y coordinate – state plane feet
AREA_SQF	Area – square feet
TAZ	TAZ number
HOUSESP	Housing units – parcel (x 100)
HOUSESQ	Housing units – quarter mile radius (x 100)
HOUSESH	Housing units – half mile radius (x 100)
STUDK12P	Students K-12- parcel (x 100)
STUDK12Q	Students K-12– quarter mile radius (x 100)
STUDK12H	Students K-12– half mile radius (x 100)
STUDUNIP	Students University– parcel (x 100)
STUDUNIQ	Students University – quart. mile radius (x 100)
STUDUNIH	Students University – half mile radius (x 100)
NODES1Q	1 link nodes– quarter mile radius
NODES1H	1 link nodes– half mile radius

NODES3Q	3 link nodes– quarter mile radius
NODES3H	3 link nodes– half mile radius
NODES4Q	4+ link nodes– quarter mile radius
NODES4H	4+ link nodes– half mile radius
DIST_LRT	Distance to nearest LRT stop (miles x 100 -1 if none)
DIST_BUS	Distance to nearest bus stop (miles x 100, -1 if none)
PARKDY_P	Daily paid parking spaces- parcel
PARKDY_Q	Daily paid parking spaces- quarter mile radius
PARKDY_H	Daily paid parking spaces- half mile radius
PPRICDYP	Avg price daily parking- parcel (cts)
PPRICDYQ	Avg.price daily parking- quarter mile (cts)
PPRICDYH	Avg.price daily parking- half mile (cts)
PARKHR_P	Hourly paid parking spaces- parcel
PARKHR_Q	Hourly paid parking spaces- quarter mile radius
PARKHR_H	Hourly paid parking spaces- half mile radius
PPRICHRP	Avg price hourly parking- parcel (cts)
PPRICHRQ	Avg.price hourly parking- quarter mile (cts)
PPRICHRH	Avg.price hourly parking- half mile (cts)
EMPEDU_P	Education jobs – parcel (x 100)
EMPFOODP	Food service jobs – parcel (x 100)
EMPGOV_P	Government jobs – parcel (x 100)
EMPOFC_P	Office jobs – parcel (x 100)
EMPOTH_P	Other jobs – parcel (x 100)
EMPRET_P	Retail jobs – parcel (x 100)
EMPSVC_P	Service jobs – parcel (x 100)
EMPMED_P	Medical jobs – parcel (x 100)
EMPIND_P	Industrial jobs – parcel (x 100)
EMPTOT_P	Total jobs – parcel (x 100)
EMPEDU_Q	Education jobs – quarter mile radius (x 100)
EMPFOODQ	Food service jobs – quarter mile radius (x 100)
EMPGOV_Q	Government jobs – quarter mile radius (x 100)
EMPOFC_Q	Office jobs – quarter mile radius (x 100)
EMPOTH_Q	Other jobs – quarter mile radius (x 100)
EMPRET_Q	Retail jobs – quarter mile radius (x 100)
EMPSVC_Q	Service jobs – quarter mile radius (x 100)
EMPMED_Q	Medical jobs – quarter mile radius (x 100)
EMPIND_Q	Industrial jobs – quarter mile radius (x 100)
EMPTOT_Q	Total jobs – quarter mile radius (x 100)
EMPEDU_H	Education jobs – half mile radius (x 100)
EMPFOODH	Food service jobs – half mile radius (x 100)
EMPGOV_H	Government jobs – half mile radius (x 100)
EMPOFC_H	Office jobs – half mile radius (x 100)
EMPOTH_H	Other jobs – half mile radius (x 100)
EMPRET_H	Retail jobs – half mile radius (x 100)

EMPSVC_H	Service jobs – half mile radius (x 100)
EMPMED_H	Medical jobs – half mile radius (x 100)
EMPIND_H	Industrial jobs – half mile radius (x 100)
EMPTOT_H	Total jobs – half mile radius (x 100)
USED	(unused)
COUNTY	County (used only for usual work validation)
GQUnitsP	Non-institutionalized Group Quarters Units – parcel (x100) (used only with PopGen population)

*DaySim can read these variables in any order, but the variable names must remain the same as shown. All values from the file are read in as integers, with no decimal.

2.3.1 Parcel Files

2.3.1.1 Housing Units - Jacksonville

Parcel-level information on housing units is used to allocate the synthetic population down to individual parcels, and to influence destination choices. These data are available in the model area from parcel-level databases maintained in each county for tax assessment purposes. This section provides an overview of the steps required to take four separate parcel databases, combine them to create a consistent regional database, impute housing units for land use types such as condominium and multi-family housing developments where housing unit numbers were not necessarily present in the parcel database, and verify that the resulting housing unit numbers were reasonable.

The county level databases obtained for each county were each structured differently and contained different data items. The following steps were carried out using each separate database before they were combined:

- Identify and remove non-travel generating parcels such as highway rights of way, bodies of water, or other extraneous parcels. This step did not remove currently undeveloped land that could be developed in the future.
- Merge together into a single parcel any parcels that are the same parcel but are spread across separate GIS shapes with multiple database rows. This step involved summing some fields such as land areas if those were divided across multiple polygons.
- Develop a set of common fields across the four databases such as land use type, effective year, land area, building area, and number of buildings.

Ultimately, a combined database of 618,981 parcels was developed. The parcel database do not specify the number of housing units on a parcel; that must be imputed based on the data that are included in the database. The following steps were taken to impute the number of housing units:

- The databases varied in terms of when they were last updated. Any buildings with an effective year (the year that the structure was built) after 2005 (the model base year) were not included in the analysis.
- Parcels with a Single Family or Mobile Home land use type were assumed to contain one unit.
- Parcels with a Condominium land use type typically contain one unit. Generally, the parcel database show large condominium buildings or developments as a grid of small

parcels, with each parcel representing a single unit. There were some exceptions to this rule that were identified base on a scan of land area, building area, and numbers of buildings. In those cases, the parcels were treated in a similar way to other types of multi-family housing parcels.

- Multi-family housing presented the largest challenge when imputing a number of housing units.
  - The data available were identification as either a development of less than 10 units or 10 or more units, and typically a number of buildings and a total building area.
  - An initial imputation was made to assign a unit for every 1000 square feet of floor area in each building and to assign an extra unit to any remainder in excess of 500 square feet.
  - All parcels in the less than 10 unit category were then constrained to a minimum number of units of 2 and a maximum of 9, while those in the 10 or more unit category were constrained to a minimum number of units of 10.
  - At this point, outliers were identified by looking at the largest developments in terms of imputed units. This identified issues such as large developments that were split into parcels for each building but the total building area for the whole development was assigned to each parcel. Following manual identification and correcting of these issues, the housing unit numbers were recalculated.
  - The number of multi-family units were then compared at a county level with data from the NERPM model to check for aggregate consistency. The 1000 square feet of floor area per unit was increased to 1275 square feet to match the regional multi-family unit total contained in the NERPM model for 2005.

Table 27 shows a summary of the housing unit numbers developed using the parcel databases by unit type and by county. They compare reasonably closely to the county level totals by unit type used in the NERPM model shown in Table 28.

Table 27: Summary of housing unit numbers by type and county (DaySim parcel file)

Description	Clay	Duval	Nassau	St Johns	Grand Total
SINGLE FAMILY	50,990	236,138	19,170	50,340	356,638
MOBILE HOMES	10,183	11,768	5,941	5,594	33,486
CONDOMINIA	11	18,576	2,493	3	21,083
MULTI-FAMILY(>=10)	6,632	97,143	742	8,490	113,007
MULTI-FAMILY (< 10)	851	11,031	1,076	12,676	25,634
Total Single Family	61,173	247,906	25,111	55,934	390,124
Total Multi Family	7,494	126,750	4,311	21,169	159,724
Total Units	**68,667**	**374,656**	**29,422**	**77,103**	**549,848**

Table 28: Summary of housing unit numbers by type and county (NERPM model, 2005)

County	Clay	Duval	Nassau	St Johns	Grand Total
Single Family Units	57,477	251,373	26,190	54,588	389,628
Multi Family Units	9,662	124,499	6,348	19,568	160,077
Total	**67,139**	**375,872**	**32,538**	**74,156**	**549,705**

Parcel-level calculations of proximity to total housing units within ¼ mile and ½ mile buffers are also important urban form measures used in DaySim, and are calculated using a script as described in section 2.3.1.7.

2.3.1.2 Housing Units – Burlington

In the Burlington region, most of the parcel level data on housing, enrollment, and employment were obtained from the Chittenden County Metropolitan Planning Organization (CCMPO) regional travel demand model. The parcel data geodatabase ("year_built_to_share.mdb") contained the table "Parcels_yrbuilt" which was processed to remove non-travel generating parcels such as highway rights of way, bodies of water, or other extraneous parcels. This resulted in polygon data for 50,052 parcels covering the CCMPO model area. The fields in the table included parcel ID, location data, town, and area.

The data on Burlington housing units were obtained from the CCMPO regional travel demand model. The data for the model was were originally collected by the Chittenden County Regional Planning Commission (CCRPC) based on the 2005 municipal Grand List. Comparisons of this these data had been made with data from other sources such as existing parcel data, 2000 Census, and building permits etc., to ensure its accuracy. The data were in the form of a point shape file ("20080612_rpc_2005_housing_points.shp") which contained points for 42,142 residential locations covering the CCMPO model area. The data fields included parcel ID, town, address, dwelling unit type (single family, multi family, group quarters etc.), dwelling unit count, TAZ. Total dwelling units including group quarters is 65,449 .

Some data cleaning was done such as removing non-travel generating parcels such as highway rights of way, bodies of water, or other extraneous parcels. Table 29 shows a summary of the housing unit numbers developed using the parcel databases by unit type for Chittenden County.

Table 29. Summary of Burlington Housing unit numbers by type

Description	Chittenden County
SINGLE FAMILY	36,342
MULTI-FAMILY(2-4)	11,510
MULTI-FAMILY(5+)	12,123
GROUP QUARTERS	5,474
Total Units	**65,449**

2.3.1.3 Employment by Type - Jacksonville

Parcel-level information on the total number of jobs by employment type on each individual parcel is one of the most essential model inputs. In DaySim, the number of workers attracted to each employment site is calibrated to the number of jobs at that site. Detailed information on employment by type was acquired for the Jacksonville area. The employment data for Clay, Duval, and Nassau counties prepared by FDOT, NFTPO, and the consulting firm PBS&J's staff, and were obtained from PBS&J. PBS&J provided a GIS file of employment location points, with records for individual businesses including number of employees and six digit SIC code. Extensive efforts were made to clean and verify the raw employment database that they had obtained and hence it required minimal additional cleaning for use in this model.

The employment data for St. Johns County were obtained in the form of an InfoUSA database that had not been cleaned or processed in any way. These data were similar in format to the data for the other three counties, containing records for individual businesses including number of employees and SIC codes. Prior to processing these data into the format required for the parcel data input file, the data were checked to ensure data quality. These checks include reviewing the largest individual points in terms of number of employees: a common problem is that large employers with many locations in a region allocate all jobs to the “home office” rather than distributed across the firms locations.

Once the employment point databases were obtained and were deemed satisfactory, they were processed using the following steps:

1. **Association of DaySim employment types with each business:** DaySim models employment using nine employment types. Table 30 shows the correspondence between these aggregated employment categories and the more detailed SIC classification. A 2-digit SIC code was derived from the 6-digit SIC codes in the employment databases (this is simply the first two digits of the 6-digit SIC code). Then the correspondence was used to associate the DaySim employment types with each employment location.

Table 30: DaySim Employment Sectors and Correspondence with SIC Categories

ID	Daysim Employment Sectors	SIC Major Categories / Generalized 2-digit categories	SIC 2-digit codes
1	Education	Educational Services	82
2	Food service	Eating And Drinking Places	58
3	Government	Public Administration	91-97
4	Office	Finance and Real Estate, Services (some major categories)	60-67, 73, 81, 86, 87
5	Other	Private Households, Nonclassifiable Establishments	88,99
6	Retail	Retail Trade	52-59
7	Service	Transportation, Services (some major categories)	40-49, 70, 72, 75, 76, 78, 79, 83, 84, 89
8	Medical	Health Services	80
9	Industrial	Agriculture, Mining, Construction, Manufacturing, Wholesale Trade	01-39, 50, 51

2. **Association of employments point with parcels:** the employment points were associated with parcels using GIS to intersect the employment point file and the parcel polygon file. During this process, several checks were carried out to ensure the reasonableness of the association. These included:
    a. Checking that all employment had been assigned to a parcel. Given that the parcel polygon file excludes rights of way and many employment points are geocoded close to the adjacent street, points can fall outside of any parcel. These points were associated with the closest parcel to move then outside of the right of way and into a parcel.
    b. Checking the land use of the parcels with which employment was associated. The parcel polygon file included data on the land use type of the parcel. Employment associated with residential parcels and vacant parcels was checked. In some cases, residences are legitimate business locations, particularly for individual businesses (i.e. those with one employee where the person works from home). Association with vacant parcels can indicate poor geocoding (in which case points can be manually repositioned to the appropriate parcel) or inconsistency between the employment data and the

parcel database; for example recently opened businesses that supersede the latest year built information in the parcel database.

3. **Aggregation of employment at the parcel level**: DaySim requires employment data to be aggregated from individual points within a parcel to totals for each of the nine employment types on a parcel. At this point, parcels with large quantities of employment were spot checked to ensure that a reasonable number of jobs are assigned to individual parcels.
4. **Adjustments to employment in Military Base parcels:** The employment at two military bases was adjusted to reflect data obtained on the number of jobs. For the Naval Air Station Jacksonville, the number of jobs on the parcel was increased by 25,552 to reflect active duty and civilian personnel not accounted for in the employment data, and 10,565 jobs were added to the Naval Station Mayport parcel; together a total of 36,117 jobs were added.

Table 31 summarizes the processed employment data by county and by employment type. Duval County accounts for the majority of the employment in the region, in excess of 80%, with fewer than 50,000 jobs in the each of the other three counties. The distributions of jobs by employment type are similar in each of the four counties.

Table 31: Cross tabulation of employment by employment type and county from DaySim Parcel file

County	Clay	Duval	Nassau	St Johns	Total
Education	3,981	25,632	1,499	2,805	33,917
Food service	5,519	32,233	2,192	4,005	43,949
Government	3,227	65,212	1,980	1,429	71,848
Office	5,210	96,596	2,175	5,118	109,099
Other	95	1,178	9	133	1,415
Retail	10,572	64,647	3,397	5,049	83,665
Service	6,790	78,601	4,614	6,264	96,269
Medical	5,224	57,383	1,681	3,664	67,952
Industrial	6,748	92,779	2,666	5,345	107,538
Total	**47,366**	**514,261**	**20,213**	**33,812**	**615,652**

Table 32shows a comparison of county level employment totals used by the NERPM model for the 2005 model year with those derived from employment databases for use in DaySim. The data received from PBS&J and used in DaySim are relatively similar to the NERPM model data for Clay and Nassau counties. Duval County has 56,000 (12%) more jobs according to the updated PBS&J data, i.e. including 36,000 additional jobs at the military bases. For St. Johns County, the InfoUSA database contains only 34,000 jobs, significantly less than the 53,000 in the NERPM model data. To address this discrepancy, St. John's County employment was subsequently scaled up to match aggregate NERPM-based totals

Table 32: County level employment totals

County	NERPM	DaySim
Clay	41,513	47,366
Duval	458,166	514,261
Nassau	20,579	20,213
St Johns	53,359	33,812
Total	**573,617**	**615,652**

As with housing units, parcel-level calculations of proximity to total employment by sector within ¼ mile and ½ mile buffers are used in DaySim, and are calculated using a script as described in section 2.3.1.7.

2.3.1.4 Employment by Type – Burlington

The parcel level employment data for Chittenden County was derived from the CCMPO model inputs. The CCMPO originally collected employment land use data from two distinct sources: infoUSA (a commercial data provider) and the Vermont Department of Employment and Training (DET). Since the VT DET employment has a privacy agreement and use restrictions, the CCMPO chose to use the infoUSA data and supplemented gaps in the infoUSA data using the VT DET data. The infoUSA data contain information such as the name of the employer, the address of the employer, the general number of employees, and the employer's Standard Industrial Classification (SIC) code.

The CCMPO had geocoded the employers based on the address. As in the case of housing units' data, the parcel employment data were received in the form of a point shapefile. There were points for 7,478 employment locations covering the CCMPO model area. The employment data includes schools, with (for most schools) grades and enrollment ranges. Two fields in the data contained employment numbers. It was found that one of fields that was supposed to contain the number of jobs - "ACTUAL_EMP" – had double counted the number of jobs in at some major employers, such as IBM, that were based at more than one at two sites. Therefore, a second different field "MPO_EMP" was included that these issues resolved and this field was used to derive the number of jobs at each employment location. The total number of jobs in the region is 102,260. Each employment location was associated with a DaySim employment type based on the correspondence with SIC codes shown in Table 30.

In the next step, each employment location was assigned to a parcel by intersecting with the parcel polygon file in GIS. Employment locations that fell outside all parcels were assigned to the nearest parcel. Finally, the number of jobs by employment type were aggregated to the parcel level and appended to the DaySim parcel input data file. Table 33 summarizes the processed employment data by employment type for Chittenden County.

Table 33. Number of Jobs by Employment Type in Burlington

Employment Type	Number of Jobs
Education	9,679
Food service	5,967
Government	4,486
Office	16,123
Other	263
Retail	13,779
Service	19,539
Medical	6,686
Industrial	25,738
Total	**102,260**

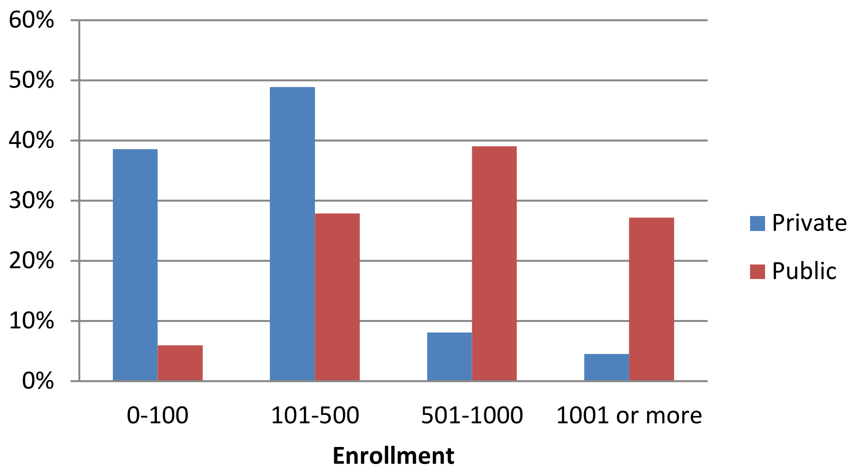
As with housing units, parcel-level calculations of proximity to total employment by sector within ¼ mile and ½ mile buffers are used in DaySim, and are calculated using a GIS script.

2.3.1.5 School Enrollment - Jacksonville

Like workers, the number of students that are attracted to each school location is calibrated to the enrollment by grade-level at that school location. As a result, parcel-level information on school enrollment is necessary. DaySim distinguishes school enrollment into enrollment in grades K-12 and then College/University enrollment.

In the Jacksonville region, the Florida Department of Education (FDOE) has provided school-level information on enrollment by grade for schools enrolling K-12 students. The data was provided for 230 private schools and 357 public schools. The schools were then geocoded based on their addresses to obtain enrollment information at the parcel level. A small number of schools (4 among private and 17 among public) could not be located for various reasons, such as being planned but not yet constructed or open schools, or that the information related to education related administration such as scholarship funds or home schooling programs. Overall, there was non-zero enrollment information identified for 223 private and 287 public schools. Figure 18 shows the distribution of enrollment for both public and private schools. As is expected, enrollment in public schools is skewed more towards the higher ranges than the private schools.

Figure 18: K-12 Enrollment Distribution



Universities, community colleges, and technical schools are identified in the employment database. Enrollment has been identified for some of these institutions by visiting the institution's website, including data on the State University system obtained from the Florida Board of Governors and from FDOE for the community college systems. In the remainder of cases enrollment has been estimated based on the number of employees at the institution. An average ratio of student enrollment to number of employees was calculated for this purpose, which equals 12.43. Table 34 shows the enrollment and employment information for all of the universities/colleges with an indication of whether the enrollment was imputed. Only those institutions with 10 or more employees are shown for brevity in this table. The information from the Jacksonville University campus was not used in the calculation of the average enrollment to employment ratio because of the unusually high value (170).

Table 34: University Level Enrollment and Employment

County	School	Employment	Enrollment	Imputed
Duval	Edward Waters College	20	839	
Duval	F CCJ College Administration	20	189	X
Duval	FCCJ Culinary Inst	28	264	X
Duval	Florida Community College	1970	18598	X
Nassau	Florida Community College	60	566	X
Duval	Florida Coastal School Of Law	95	1539	
Duval	Florida Metropolitan Univ	80	994	X
Duval	Florida Technical College	15	186	X
Duval	ITT Technical Institute	45	559	X
Duval	Jacksonville University Campus	20	3400	
Duval	Jones College	160	650	X
Duval	Logos University	10	124	X
Duval	Remington College	40	497	X
Duval	St Thomas Christian College	20	249	X

Duval	Troy University	20	249	X
Duval	University of Florida College	20	249	X
Duval	University of North Florida	1233	15420	
Duval	University of Phoenix Inc	25	311	X
Duval	Webster University	10	124	X
Duval	Conservative Theological Seminary	14	174	X
St. Johns	University of St Augustine	30	373	X
St. Johns	First Coast Technical Inst	165	734	
St. Johns	Flagler College	250	2716	
St. Johns	Flagler's Legacy	12	149	X
St. Johns	St Johns River Community College	40	1091	
Clay	Florida Metropolitan Univ	10	124	X
Clay	St Johns River Community College	70	941	

Table 35 shows a summary of the final enrollment numbers by type of school and county.

Table 35: Type of Enrollment by County

County	Clay	Duval	Nassau	St Johns	Total
K-12	43,251	195,662	12,188	20,417	271,518
University	1,115	45,349	567	4,764	51,795
Total	**44,366**	**241,011**	**12,755**	**25,181**	**323,313**

Table 36 shows a comparison of derived enrollment with that from the NERPM model at a county level. The enrollment input into DaySim is matches reasonably closely with that used in the NERPM model.

Table 36: Total Enrollment Comparison by County

County	NERPM	DaySim
Clay	39,582	44,366
Duval	227,964	241,011
Nassau	13,740	12,755
St Johns	32,712	25,181
Total	**313,998**	**323,313**

As with housing units and employment, parcel-level calculations of proximity to school enrollment by sector within ¼ mile and ½ mile buffers are used in DaySim. They are calculated using a script as described in section 2.3.1.7.

2.3.1.6 School Enrollment – Burlington

As in Jacksonville, there are two grade levels used in DaySim: K-12 and college\university. For K-12 grade level, the enrollment numbers were derived using the employment data from CCMPO as described in the following steps:

a. All employment locations falling in the primary category of "Schools" (about 100) were considered separately.

b. Some of these were found to be invalid locations such as school district administrative offices, dance\martial arts classes etc. and were filtered out.

c. Each employment location had the number of jobs and a range of student enrollment associated with it. The midpoint of the range of student enrollment was used as the estimate of enrollment at a particular location.

d. Using the total estimated enrollments and the total number of jobs at all school locations, an average value of students per (school) employee was computed, which equals 6.0.

e. This overall average students per employee ratio was then used to estimate the final K-12 student enrollments for each school location.

The total number of students in all the parcels in the region was calculated as 23,706. This value was quite close to 22,403, the value that was found in Vermont Public School Enrollment Report for 2008-09.

For university level enrollment, the numbers were directly obtained from websites of the respective colleges\universities since there were only a few of them. Table 37 shows the colleges included in the parcel level college grade enrollment.

Table 37: University Student Enrollment in Chittenden Co

College\University	Enrollment
St Michael's College	2,700
Champlain College	2,000
Burlington College	200
University of Vermont	11,704
Vermont Hitec Inc	240

As with housing units and employment, parcel-level calculations of proximity to school enrollment by sector within ¼ mile and ½ mile buffers are calculated using a separate script.

2.3.1.7 Parcel Level Buffers of Housing Units, Employment and School Enrollment

Parcel-level calculations of proximity to total housing units, employment by employment type, and school enrollment by school type within ¼ mile and ½ mile buffers are important urban form measures used in DaySim, and are calculated using a GIS script. Figure 19 shows an example of the ¼ and ½ mile buffers around a parcel centroid. The figure shows the parcel centroids of adjacent parcels in brown. All those centroids that fall within a buffer are counted when the various buffer variables are summed.

Figure 19: Example of ¼ mile and ½ mile buffer areas around a parcel centroid

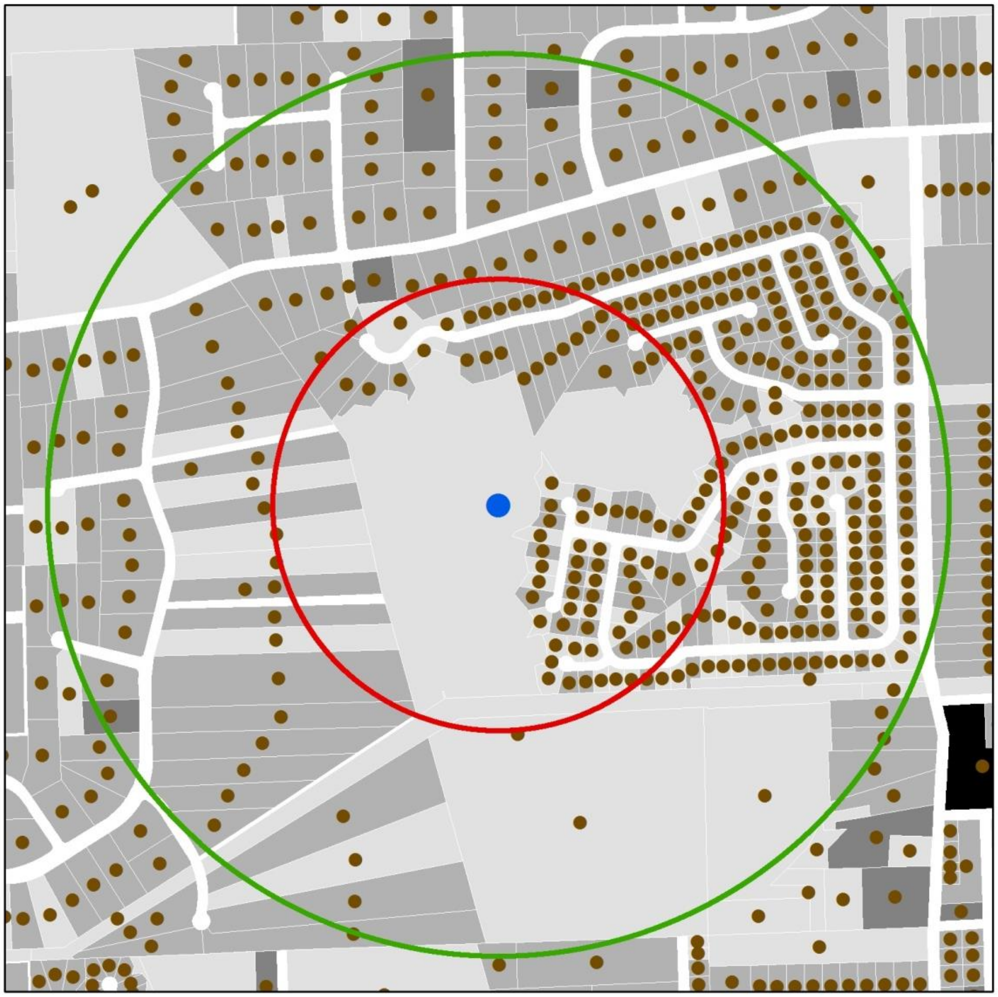


Figure 20 to Figure 23 show maps of housing, employment and School (K-12) enrollment per parcel and within ¼ mile and ½ mile buffers. Figure 20 and Figure 21 compare urban and suburban housing variables.

Figure 20: Housing units per parcel and ¼ mile and ½ mile housing unit buffers, urban area

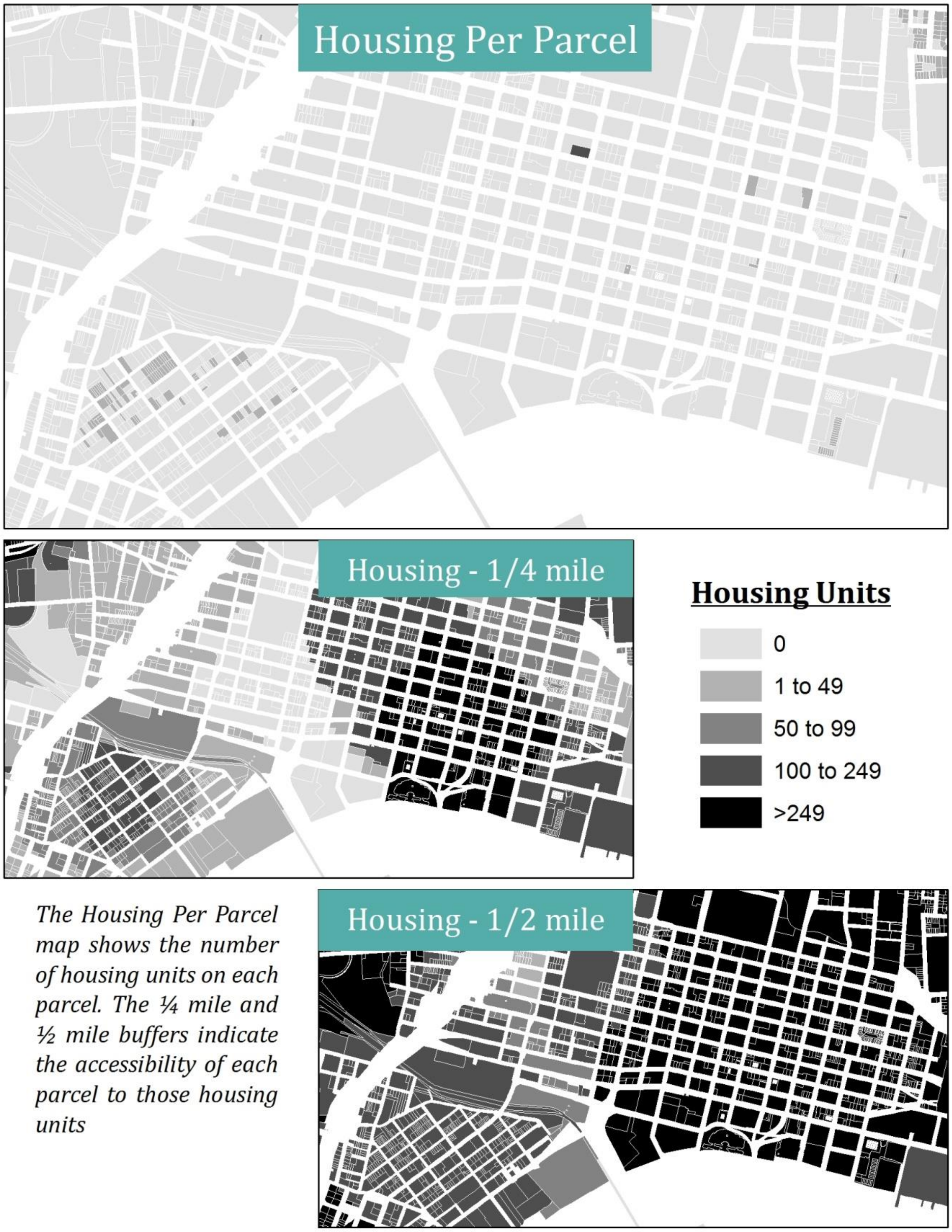
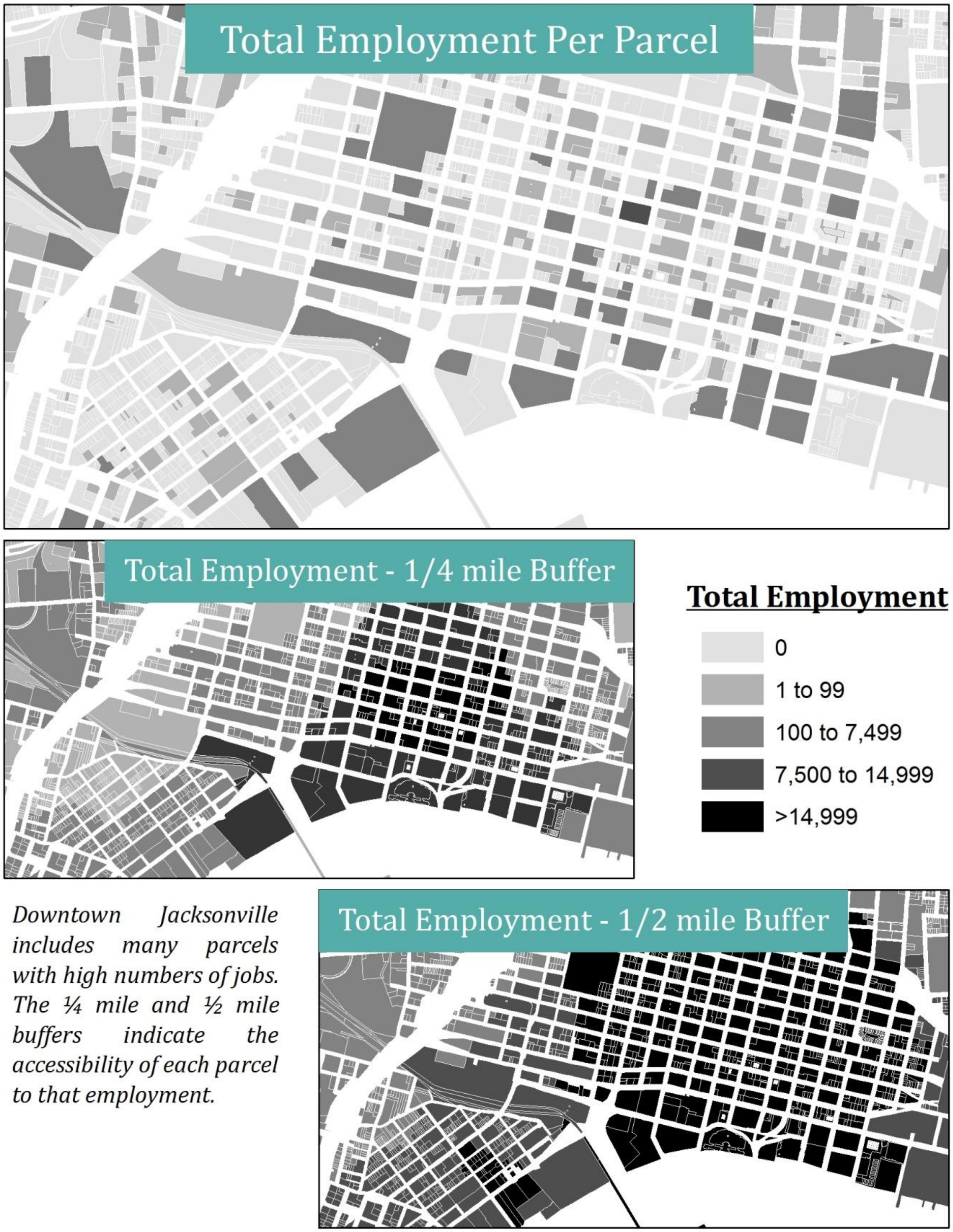


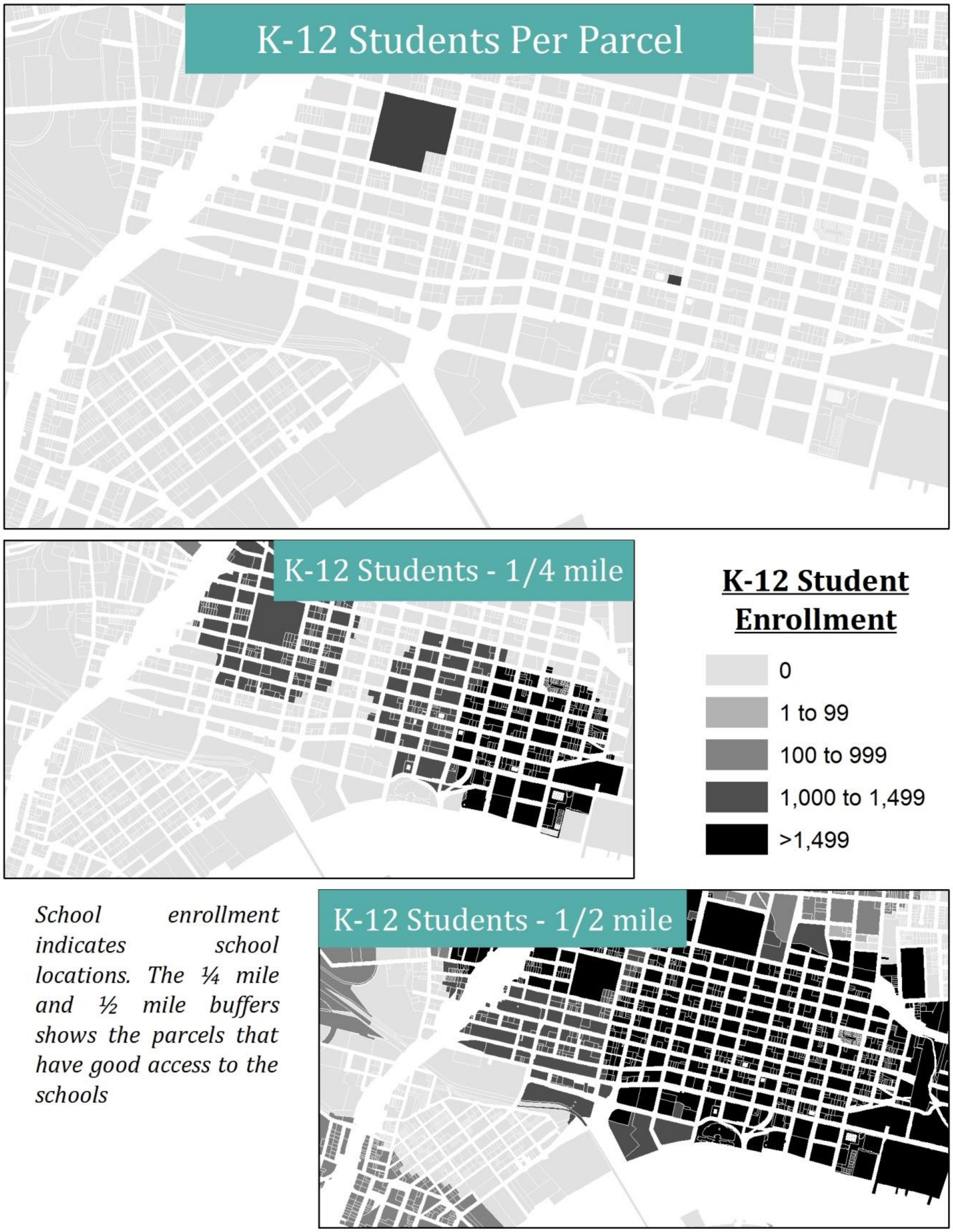
Figure 21: Housing units per parcel and ¼ mile and ½ mile housing unit buffers, suburban area

Figure 22: Total employment per parcel and ¼ mile and ½ mile total employment buffers, urban area



Downtown Jacksonville includes many parcels with high numbers of jobs. The ¼ mile and ½ mile buffers indicate the accessibility of each parcel to that employment.

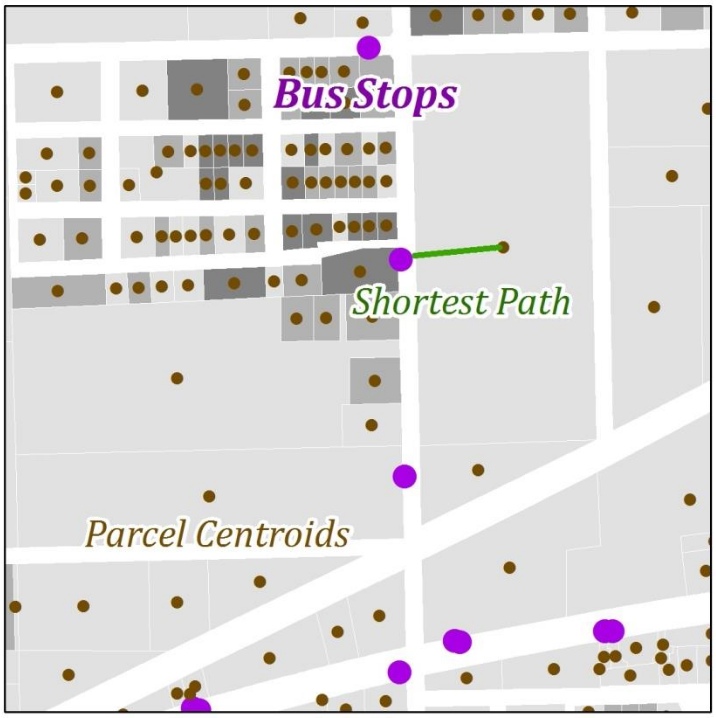
Figure 23: K-12 Student enrollment per parcel and ¼ mile and ½ mile K-12 student enrollment buffers, urban area



2.3.1.8 Transportation Access

In addition to using zone-level information on access times to transit, DaySim also incorporates detailed parcel-level information on the distance to transit by transit sub-mode. In the case of Jacksonville, two transit modes are included: bus and the JTA Skyway. The Jacksonville Transportation Authority provided GIS data on transit stop locations, and a GIS-based script has been developed to calculate distances to the closest transit stop for every parcel in the region.

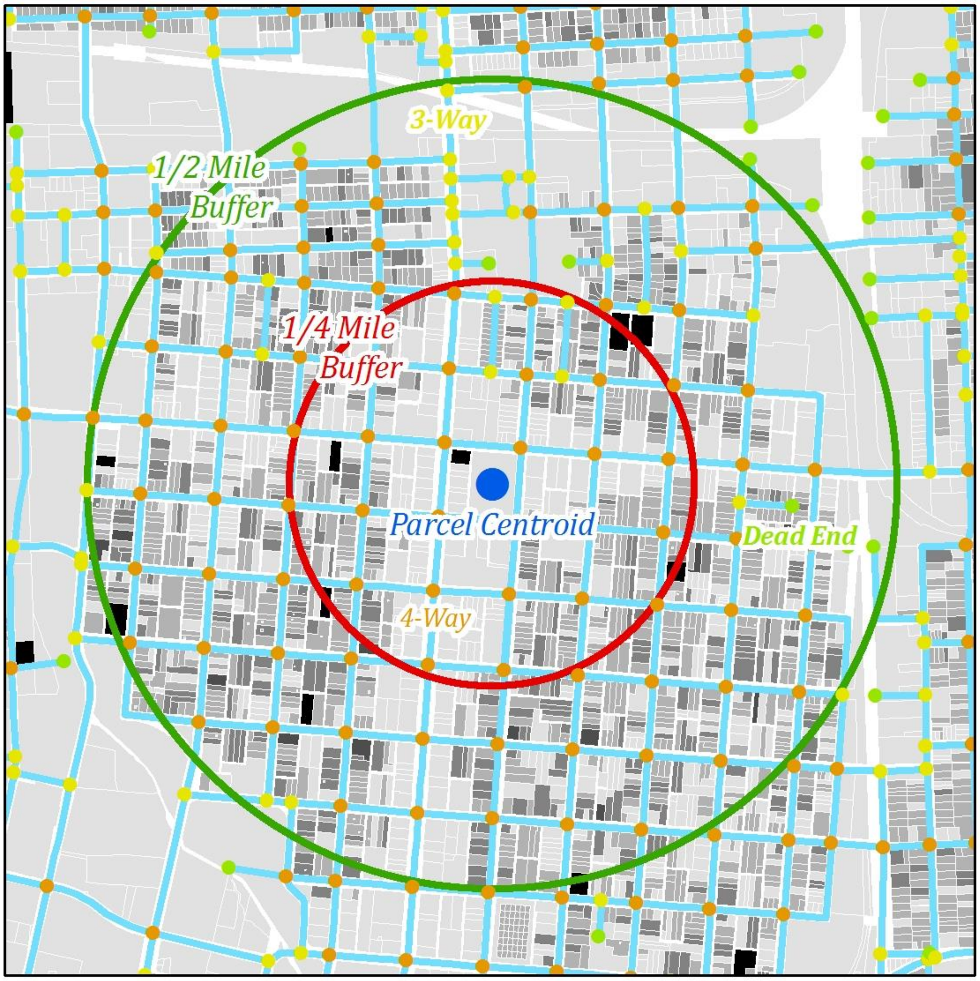
Figure 24: Example of closest bus stop to parcel centroid



2.3.1.9 Urban Form

A unique parcel-level measure of urban form that DaySim incorporates is the number of intersections or nodes of different types within ¼ mile and ½ mile buffers. These intersection types include, dead-ends (1 link), T-intersections (3-links), and tradition intersections (4+ links), and help characterize the pattern of urban development. An automated process has been developed to calculate these urban form measures for Jacksonville based on detailed GIS street centerline files. In Burlington, an updated Tiger Lines road map for Vermont was downloaded and used for this purpose. Both of these networks are more detailed than the modeled network, which does not include all streets. This GIS process first analyses the GIS street centerline file to locate nodes and assign an intersection type code to them based on the number of links joined to the node. The process then creates buffer areas around each parcel and then counts the number of intersections of each type that fall within the buffers.

Figure 25: Example of ¼ mile and ½ mile node buffers



2.3.1.10 Parking

DaySim uses information on the number and prices of both daily and hourly parking spaces on the parcel and within ¼ mile and ½ mile buffers of each parcel. The project team has inventoried the location and pricing information for paid off-street locations, but has not yet obtained or developed accurate information on capacities.

The point based parking locations are assigned to individual parcels to develop the parcel data (e.g. "PARKDY_P", Daily paid parking spaces on the parcel). The buffer variables are calculated in similar way to the urban form network buffers described above, using a GIS process to create buffer areas around each parcel and then sum the parking capacity within the buffers and calculate the weighted average of the parking price at the parking lots in the buffer.

2.4 DaySim TAZ Data

Parcels are the primary spatial units used in DaySim. However, current implementations have used a limited set of TAZ level data, including PUMA and summary district correspondence codes, and physical attributes such as the land area and coordinate locations. The Jacksonville model uses the TAZ system from the NERPM model, except that the zones are renumbered so that external zones are first (1-23) and internal zones are numbered consecutively (24-1335). Similarly, the Burlington model uses the TAZ system from the CCMPO model.

Table 38 shows the file layout for the Jacksonville implementation of DaySim. The “XCORD”,”YCORD”, and “SQFT_Z” fields were developed in ArcGIS using the TAZ shape file supplied with the NERPM model. PUMA’s were assigned to each TAZ by intersecting the TAZ shape file with a PUMA boundary shape file obtained from the Census Bureau.

Table 38: TAZ data input file format

Label**	Definition
TAZ	Zone number
AUTACC	Auto access time (min x 100) *
AUTEGR	Auto egress time (min x 100) *
PRKCOST	Parking cost in zone (cents/hour) *
DAVIS	Davis dummy (0/1)
PEDENV	Pedestrian environment score *
PUMA	PUMA code for zone
RAD	RAD code for zone (aggregation of zones) *
XCORD	X coordinate of zone centroid (state plane ft)
YCORD	Y coordinate of zone centroid (state plane ft)
PKNRCOST	Park and ride lot cost in zone (cents)
SQFT_Z	Area of zone (square feet)

* not used in models

** DaySim can read these variables in any order, but the variable names must remain the same as shown: All values from the file are read in as integers, with no decimal.

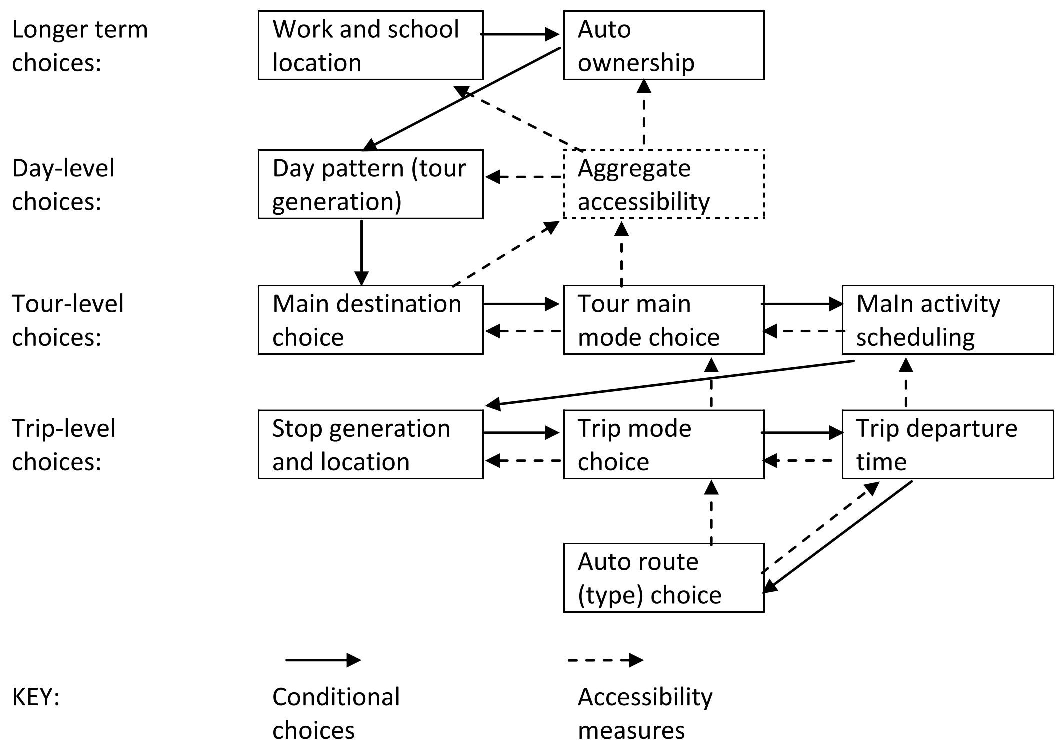
2.5 DaySim Pricing Enhancements

Key goals of the SHRP2 C10A model system development effort include providing enhanced representation of travelers’ sensitivities to price, as well as to incorporating findings from other SHRP 2 Capacity projects. This section describes how DaySim and TRANSIMS have been refined and configured to provide a more robust capabilities with respect to tolls and other types of road user charges are modeled in the integrated DaySim/TRANSIMS model framework. This section starts with a more theoretical discussion of an ideal model representation, and then addresses the more practical aspects of our current implementation. In several aspects of this work, the project team has made use of the findings from the SHRP 2 C04 project “Improving Our Understanding of How Highway Congestion and Pricing Affect Travel Demand”. The objective has been to implement the key findings from that study in a manner that retains as much behavioral detail as possible while also remaining practical for model application.

2.5.1 Choice model context

DaySim is an activity-based (AB) model structure that includes several different levels of travel choices, as shown in Figure 26. The solid arrows in the figure depict the “downward” flow of conditionality of the simulated choices—auto ownership is conditioned by the simulated longer-term work and school locations, the day activity pattern (tour generation, plus some aspects of intermediate stop generation) is conditional on all longer-term choices, the tour-level choices are conditional on the longer-term and day-level choices, etc. The dashed arrows represent the "upward" flow of accessibility in the models, where travel times and costs have the most immediate effects on the trip-level models, but ultimately affect all of the DaySim models.

Figure 26. DaySim Model Structure



The DaySim structure has been carefully designed to include accessibility effects at all choice levels, as consistently and comprehensively as possible. (In the literature, this is termed "vertical integrity", with consistent information flows both upwards and downwards.) This is done through the use of accessibility "logsums", where a logsum is a statistical construct used in discrete choice modeling that captures the expected utility across all available choice alternatives. As much as possible, DaySim uses fully disaggregate logsums, which essentially combine two models into a single joint, simultaneous model. An example is the tour-level main destination choice model, which uses logsums representing the composite attractiveness of traveling to a given destination across all available modes. For the longer-term and day-level choices, however, the use of fully detailed disaggregate logsums is not practical, as there are too many possible combinations of tour- and trip-level alternatives that correspond to each upper level choice alternative. Instead, these models use "aggregate accessibility logsums", which are logsums from a simplified tour-level model across all possible destinations, modes, and time periods, segmented by (a) residence zone, (b) travel purpose, (c) car availability level, and (d) distance from the nearest transit stop.

2.5.2 Incorporating auto route (type) choice into DaySim

For modeling highway pricing and congestion effects, the relationship between auto route choice (at the bottom of Figure 26) and the other travel choices is critical. As the available auto routes change in terms of their travel time and/or cost at different times of day, the most direct effects are on the travel mode and departure time chosen for a particular trip. (However, it is also important to represent effects on all of the other DaySim choices—a subject we return to below.)

In previous implementations of DaySim, auto route choice has been handled outside of DaySim itself. Instead, for a given auto sub-mode, such as SOV , and a given time period, such as AM peak, the best route has been pre-determined in a network package such as CUBE, TransCAD or EMME, and DaySim has simply used matrices of travel time, distance, and cost along the best path between each O-D pair. Our previous work of integrating DaySim with TRANSIMS has also followed that same general approach, but with two key improvements: (a) a more realistic, dynamic representation of traffic congestion than is typically possible using a static equilibrium assignment approach, and (b) incorporation of much more detail in terms of how congestion levels vary across the day. Whereas DaySim typically uses highway skims from only 4 or 5 different time periods across the day, our current DaySim/TRANSIMS implementation uses skims from 22 time periods, with durations as short as 30 minutes long during the peak periods.

If finding the "best" path through a network simply involves finding the shortest time path, then it may be satisfactory to do all path-building and assignment outside of DaySim in a more aggregate environment. This may not be the case, however, when toll cost and/or operating cost become key considerations in choosing a route. Each traveler may have different trade-offs between travel time and cost—their so-called value of time (VOT) or willingness to pay (WTP) for time savings. The same is true if additional variables such as travel time variability/reliability are added to the route choice model, in which case different travelers may also have different values of reliability (VOR) relative to travel time and cost.

Ideally, the network path choice would be fully integrated into an AB model such as DaySim. When applying a time of day choice for a given tour or trip, for example, DaySim would evaluate all available paths through the network at each time of day for that given traveler on that given tour or trip. In other words, network path choice would be done "on the fly" in a fully disaggregate manner depending on each traveler's tradeoffs between travel time, toll, distance and any other important route characteristics that are known in the network.

Although the project team studied possibilities for setting up such an "on the fly" integration of Daysim with TRANSIMS, it is not yet practical from the standpoint of computation and run times. Another possible solution is to set up different user classes depending on VOT, and to use different network skim files with pre-determined best path for each class. For example, if the two best paths for a given O-D are a travel time of 40 minutes with no toll, or a travel time of 30 minutes with $2.50 toll (and no difference in distance or operating cost), then any traveler willing to pay $2.50 for a 10 minute savings (VOT higher than $15/hour) would chose the tolled path and any traveler with VOT lower than $15/hour would choose the free path. In the same example, suppose we define 3 user classes: (1) VOT = 0 to 10 $/hr, (2) VOT = 10 to 20 $/hr, and (3) VOT > 20 $/hr. The best path for class (1) would clearly be the free route, and the best path for class (3) would clearly be the tolled route. For class (2) however, it is not clear which path is best, because the indifference point of VOT = $15/hr falls in the middle of the class (2) range of 10-20 $/hr.

To deal with this type of inaccuracy, we have adopted an approach which is commonly used in practice. This involves:

- Providing network skims of the best paths for two different route types: (1) routes from the full network, including tolled links, and (b) routes the non-tolled network only, excluding tolled links.
- Incorporating a binary route type (tolled vs. non-tolled) choice model into DaySim.
- To further increase the accuracy of this approach, provided different sets of best tolled and non-tolled paths for different ranges of VOT. (Even the non-tolled links include operating cost, so the best non-tolled path may also vary by VOT class.)
- Within each VOT class, use the VOT at the high end of that range to select the best path in TRANSIMS to use as input to DaySim.

So, in the example above with 3 VOT classes, there would be 6 different routes selected in TRANSIMS and input to Daysim:

1. VOT= 0-10/non-tolled network: The lowest generalized-cost route with VOT set at $10/hour, excluding tolled links.
2. VOT= 10-20/non-tolled network: The lowest generalized-cost route with VOT set at $20/hour, excluding tolled links.
3. VOT= 20+/non-tolled network: The lowest generalized-cost route with VOT set at $50/hour (an arbitrary, high value), excluding tolled links.
4. VOT=0-10/full network: The lowest generalized-cost route with VOT set at $10/hour, including tolled links.
5. VOT=10-20/full network: The lowest generalized-cost route with VOT set at $20/hour, including tolled links.
6. VOT=20+/full network: The lowest generalized-cost route with VOT set at $50/hour, (an arbitrary, high value) including tolled links.

Using the same example, if the best two paths are 40 minutes with no toll and 30 minutes with a toll of $2.50, then the best non-tolled route for all VOT classes (1-3 above) would be the 40 minute free route, as the tolled route is excluded. The best path from the full network for VOT = 0-10 (4 above) would also be the free route, but the best path from the full network for the other two VOT classes (5 and 6 above) would be the tolled path, since both of those classes use a VOT set higher than $15/hour to pick the best path.

Inside of DaySim, travelers with VOT in the range 0-10 would face a binary choice between two identical non-tolled paths, so there is essentially no choice. Travelers with VOT higher than 10 would all have a choice between the tolled and non-tolled paths. The probability that any traveler will pick the tolled path increases with VOT, so moving the route type choice inside of DaySim makes it more sensitive to small variations in willingness to pay, and ultimately more accurate. The binary route choice model is a probabilistic model, however, so even for very high VOT there may be a small probability of selecting the free route.

The following sections provide more details about how the binary model is implemented in DaySim. First, however, we provide more perspective on the need to include a route type choice model within DaySim, and to provide different skims to DaySim for different ranges of VOT. The second feature of using different VOT user classes is not always used in practice, and may seem unnecessary, since the binary route type choice model already accounts for differences in willingness to pay. In cases with simple pricing scenarios, such as those that include one or two isolated HOT lanes or Express lanes, the best tolled route is usually the facility in question, and that is not likely to vary across VOT classes. For a more detailed pricing scenario, however, such as mileage-based pricing on the regional freeway network, there can be a large set of different tolled paths to choose from, and the best tolled path may vary according to VOT. In such a case, providing different best paths for different ranges of VOT helps to compensate for the fact that we are only choosing a single best tolled path to input to DaySim rather than providing a larger set of possible tolled paths. The implication here is that the more complex the regional pricing scenario, the larger the number of VOT-specific user classes which should be used.

2.5.3 Binary route type choice model

The route type choice model implemented in DaySim works as follows:

*V(n,i) = s . b(i) * Time(n,i) + s . c(i) * Distance(n,i) * opcost*

*V(t,i) = s. a(i) + s. b(i) * Time(t,i) +s. c(i) * (Toll(t,i) + Distance(t,i) * opcost)*

P(t,i) = 1 – P(n,i) = exp[V(t,i)] / (exp[V(t,i)] + exp[V(n,i)]]

Where:

V(n,i) and *V(t,i)* are the systematic logit utilities for the non-tolled and tolled routes, respectively, for individual traveler I, and *P(t,i)* and *P(n,i)* are the corresponding binary logit probabilities.

Time(n,i), Time(t,i), Distance(n,i), Distance(t,i) are the travel time and distance along the best non-tolled and tolled routes, respectively, for traveler I, depending on the traveler/trip's origin, destination, time of day, and value of time (VOT) class.

Toll(n,i) is the toll along the best tolled route for traveler I, depending on the traveler/trip's origin, destination, time of day, and value of time (VOT) class.

a(i) is an alternative-specific constant for the tolled route for traveler i

b(i) is the travel time coefficient for traveler i

c(i) is the travel cost coefficient for traveler i

s is a scale factor applied to all coefficients, denoting the scale of this model relative to mode choice

opcost is the auto operating cost per mile

The strategy for providing "best path" skim values for time, distance and toll from Transims to DaySim was explained in the previous section.

The assumptions and methods used for setting coefficients a, b ,c, and s are given in the next section.

Note that if the two paths are identical in terms of time, distance and toll (=0), then the non-tolled path is selected as the chosen route type without applying the model.

Also note that we treat operating cost per mile as a constant in DaySim (which can be varied by the user to represent future fuel cost assumptions). If DaySim is enhanced in the future to include a model of vehicle type choice (economy, SUV, hybrid, etc.) , then operating cost can be treated as traveler-specific. Also, it is possible that network simulation software such as Transims could be enhanced to provide an O-D/time of day-specific estimate of average fuel usage based on speeds and traffic conditions (stop and go, etc.) along the route. In that case, average fuel usage could be another skim variable used as input to DaySim.

2.5.4 Traveler-specific coefficients

In setting traveler-specific coefficients for the model, we have used the findings from the SHRP 2 C04 study to the greatest extent possible, both for the functional forms and the magnitudes. The values are set as follows:

Work tours

c(i) = -0.15/$ / [((income(i) / 30,000) ^ 0.6) * (occupancy(i) ^ 0.8)]

b(i) = -0.030/min * draw from a log-normal distribution, with mean 1.0 and std. deviation 0.8

a(i) = -1.00

s = 1.5

Non-work tours

c(i) = -0.15/$ / [((income(i) / 30,000) ^ 0.5) * (occupancy(i) ^ 0.7)]

b(i) = -0.015/min * draw from a log-normal distribution, with mean 1.0 and std. deviation 1.0

a(i) = -1.00

s = 1.5

The cost coefficient c is set at -0.15/$ for both work and non-work tours. It is adjusted according to the household income of the traveler, using a power function with a somewhat higher exponent for work tours (0.6) than for non-work tours (0.5). When applied to specific car occupancy levels, the cost coefficient is also adjusted downwards for cost-sharing, again using a power function with a somewhat higher coefficient for work tours (0.8) than for non-work tours (0.7)

The base travel time coefficient is set at -0.030/min for work tours and -0.015/min for non-work tours. For an SOV trip for a traveler with income = 30,000, this corresponds to a VOT ratio of 60* -0.030 / -0.15, or $12 $/hr for work tours, and 60 * -0.015 / -0.15, or $6/hr for non-work tours.

The C04 study also found significant random taste variation around the base travel time coefficient, with the best results assuming a log-normal shape to the distribution, which is typical for value of time analysis. Although the results are not conclusive with regard to the amount of random variation, the C04 study and past analyses of this type generally support a coefficient of variation (std. deviation/mean) in the range of 0.7 to 1.0. Here we assume a somewhat higher coefficient of variation for non-work trips, since that covers a wider variety of trip types. (Note: The code for performing random draws from approximate normal (Gaussian) and log-normal distributions uses the ratio of uniforms method of A.J. Kinderman and J.F.

Monahan augmented with quadratic bounding curves. The original algorithm was published in TRANSACTIONS ON MATHEMATICAL SOFTWARE, VOL. 18, NO. 4, DECEMBER, 1992, PP. 434-435.)

The alternative-specific constant for the tolled route is set at -1.0 for both work and non-work tours, as evidence shows some aversion to paying tolls, all else equal. Note that it would be possible to also simulate normal or log-normal taste variation around this coefficient for each individual. However, we do not have any empirical evidence to go by, and, statistically, it would be very difficult to estimate taste variation parameters on both the toll constant and the time or cost coefficient at the same time.

The final parameter is s, the model scale relative to mode choice. If we think of the binary toll/non-toll model as a nest of mode alternatives under each of the auto alternatives in a mode choice model, then the unscaled time and cost coefficients b and c are those used in the mode choice model, while the scaled coefficients s.b and s.c are those used in the lower level route type choice nest when the logsum parameter for the nest is 1/s. Empirically, the C04 report contains nest logsum parameters on the route type choice nest ranging from 0.9 (constrained) for the New York RP data to 0.5 for various SP data sets. Here, we assumed a logsum of roughly 0.67, and the scale s is the inverse of that, at 1.5.

Conceptually, the larger the value of s, the more sensitive and deterministic the route type choice model probabilities will be, and the less sensitive the logsum from the model will be to the attributes of the unchosen/inferior alternative. The logsum from the model is important, because it is that value which is fed "upwards" from the auto route type choice model to all of the other DaySim models, as described in the following section.

2.5.5 Use of the route type choice model within DaySim

Conceptually, the route type choice model can be thought of as a binary nest beneath each of the auto alternatives (SOV, HOV2, HOV3+) in the DaySim mode choice model. So, whenever auto time and cost for one of those auto sub-modes is referenced in the DaySim models, they need to be replaced by the composite utility from both the tolled and non-tolled paths under each of those modes, just as they would be in a fully nested model. In DaySim, this is done by using the route type choice model to return a "generalized auto time" logsum whenever it is applied. The generalized time is calculated as follows:

$$V(n,i) = s \cdot b(i) * Time(n,i) + s \cdot c(i) * Distance(n,i) * opcost$$

$$V(t,i) = s \cdot a(i) + s \cdot b(i) * Time(t,i) + s \cdot c(i) * (Toll(t,i) + Distance(t,i) * opcost)$$

$$GT(i) = LN\ [(exp[V(t,i)] + exp[V(n,i)]] / (s \cdot b(i))$$

The two utility equations are the same as presented earlier, and are used to set the logit probability of choosing either route. The generalized time *GT(i)* is simply the logsum across those two alternatives, divided by the scaled travel time coefficient *(s . b(i))* to obtain units of minutes. (Because a, b and c are always negative and there are only two alternatives, the logsum will virtually always be negative as well, so the generalized time will be positive. However, a check has been placed in the code to avoid cases of negative generalized time.)

When there is no tolled alternative, then the *V(t,i)* term is not used, so the generalized time simplifies to *V(n,i) / (s . b(i)) = Time (n,i) + c(i)/b(i) * Distance(n,i) * opcost* , which is simply travel time plus operating cost divided by VOT. In upper level models, this generalized time is typically multiplied once again by a time coefficient, *b(i)*, so it becomes *b(i) * Time(n,i) + c(i) * Distance(n,i) * opcost* , which is the unscaled version of *V(n,i)* above. Note that the generalized time now includes the effects of toll and operating cost as well as travel time, so all explicit utility terms related to time, toll and operating cost were replaced in the code for those models by the single generalized time term (times a relevant time coefficient, where appropriate).

Table 39 provides a summary of how the route type choice model is used within the various component models within DaySim. Note that only one of the lowest level models, trip mode choice, actually simulates a route type choice (toll or non-toll) as a prediction, but nearly all of the DaySim models use the route type choice model in the form of the generalized auto time logsum. This feature ensures that the effects of pricing at various times of day are represented consistently at all levels of the model system. Also, most of the models take account of the effects of pricing and congestion separately for the SOV, HOV2 and HOV3+ modes, which allows the effects of HOT lanes and other occupancy-specific types of pricing and facilities to be accurately represented. The effects of pricing are also treated consistently for each of the 22 different skim periods in the tour and trip time of day choice models, so time-of-day variations in prices and congestion can have nuanced effects on demand. The tour-level and upper level models also react to pricing for both legs of a tour round trip.

As indicated in the table of the upper level models use disaggregate logsums from the tour mode choice model and/or aggregate mode/destination choice accessibility logsums to “carry up” the effects of pricing and congestion in a way that is as consistent as possible with discrete choice theory.

Table 39. How Route Type Choice Model is used in DaySim

DaySim model	Predicts route type choice?	Uses logsum as generalized auto time?	Used for modes…	Used for periods…	One way or round trip?
Work location	No	Yes	SOV, HOV2, HOV3+***	Assumed***	Round trip***
School location	No	Yes	SOV, HOV2, HOV3+***	Assumed***	Round trip***
Auto ownership	No	Yes	SOV, HOV2, HOV3+***	Assumed***	Round trip***
Day pattern choice	No	Yes	SOV, HOV2 **	Assumed**	Round trip**
Tour destination choice	No	Yes	SOV, HOV2, HOV3+*	Simulated*	Round trip*
Tour mode choice	No	Yes	SOV, HOV2, HOV3+	Simulated	Round trip
Tour time of day choice	No	Yes	Predicted tour mode	All possible	Round trip
Stop generation and location choice	No	Yes	Predicted tour mode	Predicted tour periods	One-way via stop detour
Trip mode choice	Yes	Yes	SOV, HOV2, HOV3+	All possible	One way trip
Trip time of day choice	No	Yes	Predicted trip mode	All possible	One way trip

* via disaggregate tour mode choice logsum, ** via aggregate accessibility logsums, *** via both

2.5.6 Feedback of DaySim results to Transims

The DaySim model system produces a list of person-trips for a single day for the entire regional population. With the incorporation of the route type choice model for tolling, three new variables have been added to the DaySim trip level output file:

- The toll paid for the trip
- The trip-specific (unscaled) time coefficient
- The trip-specific (unscaled) cost coefficient

This extra information can be used by TRANSIMS to (a) know whether or not to exclude tolled links from possible paths when assigning the trip to the network, and (b) use the ratio of time and cost coefficients to determining the best value-of-time specific path of each type. Both of these types of information will help to ensure that the choice behavior being predicted by DaySim is consistent with the route choices and traffic flows being predicted by TRANSIMS.

2.5.7 Treatment of Travel Time Variability / Reliability

The project team has not yet included travel time variability in the route type choice model or other choice models within DaySim. Although the C04 report provides a good deal of useful evidence regarding tradeoffs between cost, usual travel time and travel time variability/reliability, we have not yet determined a feasible way to simulate spatial, OD-specific levels of travel time variability in TRANSIMS as input to the DaySim demand models.

As discussed in the C04 report, most proxy-type variables that can be generated from a single run of the network model (e.g. congested time minus free flow time, etc.) tend to be so highly correlated with the main travel time variable that they provide very little new information. Also, given the long TRANSIMS run times, any procedures that would require multiple network simulation runs to produce "day to day" distributions of O-D travel times are not practical at this point. The project team reviewed the latest available versions of the SHRP 2 C04 and L04 reports and considered methods suggested therein which might be both useful (in terms of adding real network spatial and temporal information) and feasible (in terms of computation and run times).

2.6 TRANSIMS Network

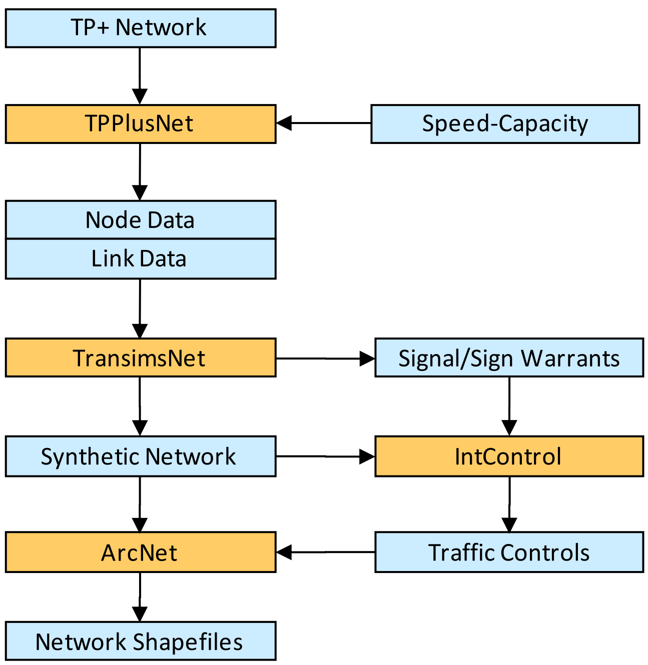
The supply side models developed for the Jacksonville and Burlington model implementations are based on the TRANSIMS network and travel assignment process. This process assigns a sequence of trips or tours for individual household persons between specific activity locations to roadways, walkways, and transit modes on a second-by-second basis for a full travel day. The network includes detailed information regarding the operational characteristics of the transportation facilities that may vary by time of day and by vehicle or traveler type. This includes the number of lanes, the lane use restrictions, the traffic controls and signal timing and phasing plans, turning restrictions, tolls and parking fees.

Most of the detailed network coding can be synthetically generated from traditional transportation modeling networks or GIS files. Traffic engineering warrants and coding rules can be customized for local conditions. The resulting data for a regional network can be edited to more accurately reflect actual conditions in the field. Since, however, TRANSIMS and SHRP2-C10 are designed to address transportation planning needs and future operational and policy scenarios, it is often helpful to design and develop the network models to dynamically adjust to future conditions rather than be fixed or limited to existing traffic controls.

2.6.1 Network Conversion Process

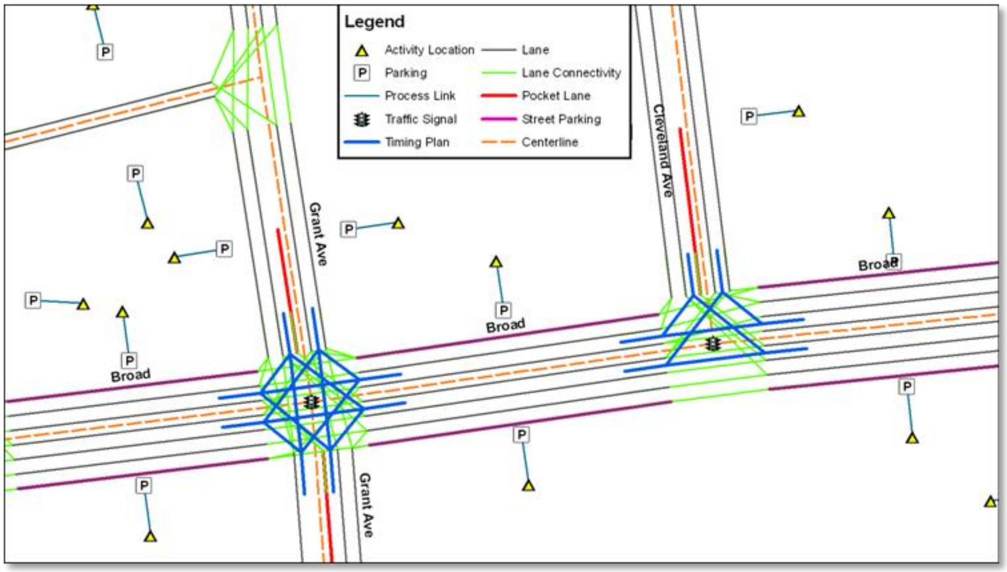
The TRANSIMS suite includes a number of tools to synthesize a TRANSIMS network from a traditional MPO networks. These tools provide a quick method of developing a detailed TRANSIMS network without a lot of extra data collection and arduous network coding. This gets the model up and running quickly and uses the trip assignment process to identify locations where the synthetic process requires refinement. The generic process for converting a TP+ network is depicted in Figure 27.

Figure 27. Network Conversion Process



The TPPlusNet program reads the one-way link records and the speed-capacity lookup table used in the regional model network (in Jacksonville this is aTP+ network while in Burlington this is a TransCAD network), and re-formats, re-groups and re-configures that data into standard TRANSIMS input link and node data files. The modified link and node files are then read by TransimsNet to synthesize the additional information needed for a network simulation. This includes pocket lanes, lane connectivity, parking lots, activity locations and signal and sign warrants. The signal and sign warrants are typically reviewed and edited prior to the execution of IntControl. IntControl synthesizes traffic signal timing and phasing plans, detectors, and signs. The ArcNet program is then used to create ArcGIS shapefiles to display and edit the network. Figure 28 shows a typical TRANSIMS network following this conversion.

Figure 28. Typical TRANSIMS Network



2.6.2 Jacksonville Network Development

In consideration of the flexibility requirement and the project goal to develop a fine-grained network, the team started the model development by creating three network resolutions from the NERPM regional modeling datasets:

1. PLANNING Network

This network is equivalent to the NERPM regional modeling network.

2. ALLSTREETS Network

This network is equivalent to the NERPM regional modeling network plus all other existing minor streets such as neighborhood streets and alleys.

3. FINEGRAINED Network

This network adds local thru streets to the PLANNING network to provide greater distribution of travel to, from, and within traffic analysis zones.

In addition to the information included in the NERPM network files, the TRANSIMS conversion process synthesizes the operational details required for network simulation. These include traffic controls, pocket lanes, lane connectivity, and lane-use or vehicle-use restrictions. The following sections describe the generic as well as network-specific conversion and enhancement processes including network data inputs, their treatment and application. These networks were prepared using TRANSIMS Version 4 tools and formats. Version 5 of TRANSIMS is expected to be available for use soon, and new tools for converting TP+ to TRANSIMS networks will support the implementation of the C10A model system in TRANSIMS v5.

2.6.2.1 NERPM Network

Florida Department of Transportation (FDOT) District 2's Northeast Florida Regional Planning Model (NERPM) maintains multiple networks in a single master network file. The master files are those files that are universally applicable to all scenarios in NERPM. These files are not altered from scenario to scenario. Instead, these files contain source data from which scenario specific information is extracted.

1. MERGED-GIS – This is a set of files that collectively form an ESRI shape file corresponding to the master network.
2. MERGED.NET – This is the master network file from which all scenarios and alternatives are derived.
3. TCARDS.PEN – This is the turn penalty file that contains all turning movement penalties and prohibitions for all scenarios. Penalty sets are used to distinguish scenarios.

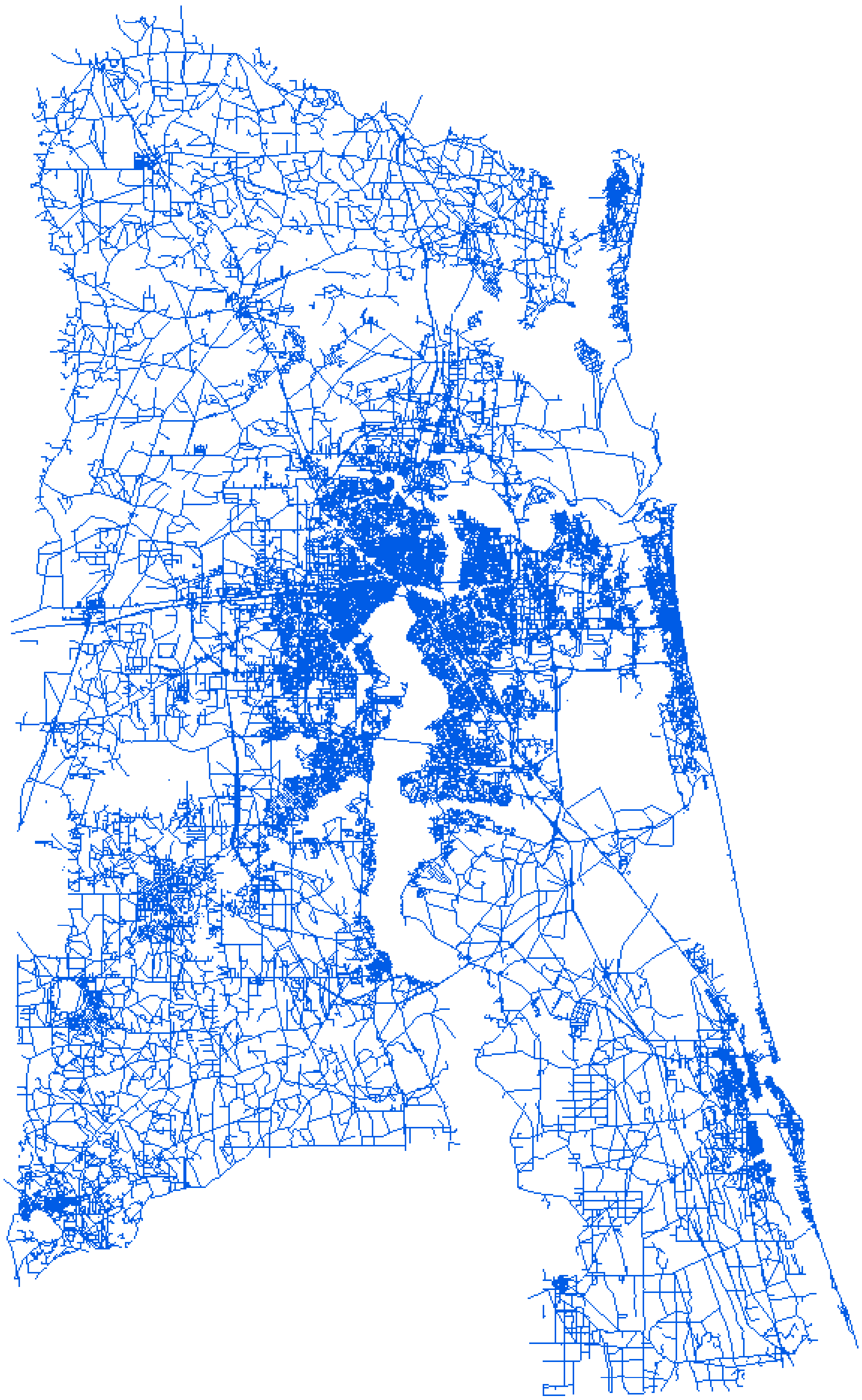
Thus various modeling year networks are coded as scenarios with specific attributes. As such, their values change from scenario to scenario. Scenario-specific attributes are identified by the presence of catalog keys designating the year and alternative of the scenario. For example, the attribute for facility type for the 2005 base year scenario is called FTYPE_05A, whereas the facility type attribute for the 2030 network is called FTYPE_30A. A break-down of the default scenario-specific networks is as follows:

- 00A – 2000 Base Year;
- 05A – 2005 Interim Year;
- 10A – 2010 Existing-plus-Committed (no socioeconomic data);
- 15A – 2015 Interim Year;
- 25A – 2025 Interim Year;
- 30A – 2030 Cost Feasible Plan Horizon Year; and
- 30N – 2030 Needs Plan (no socioeconomic data).

This study is focused on using the 2005 scenario for the base year network and the financially constrained 2030 scenario for the forecast year network.

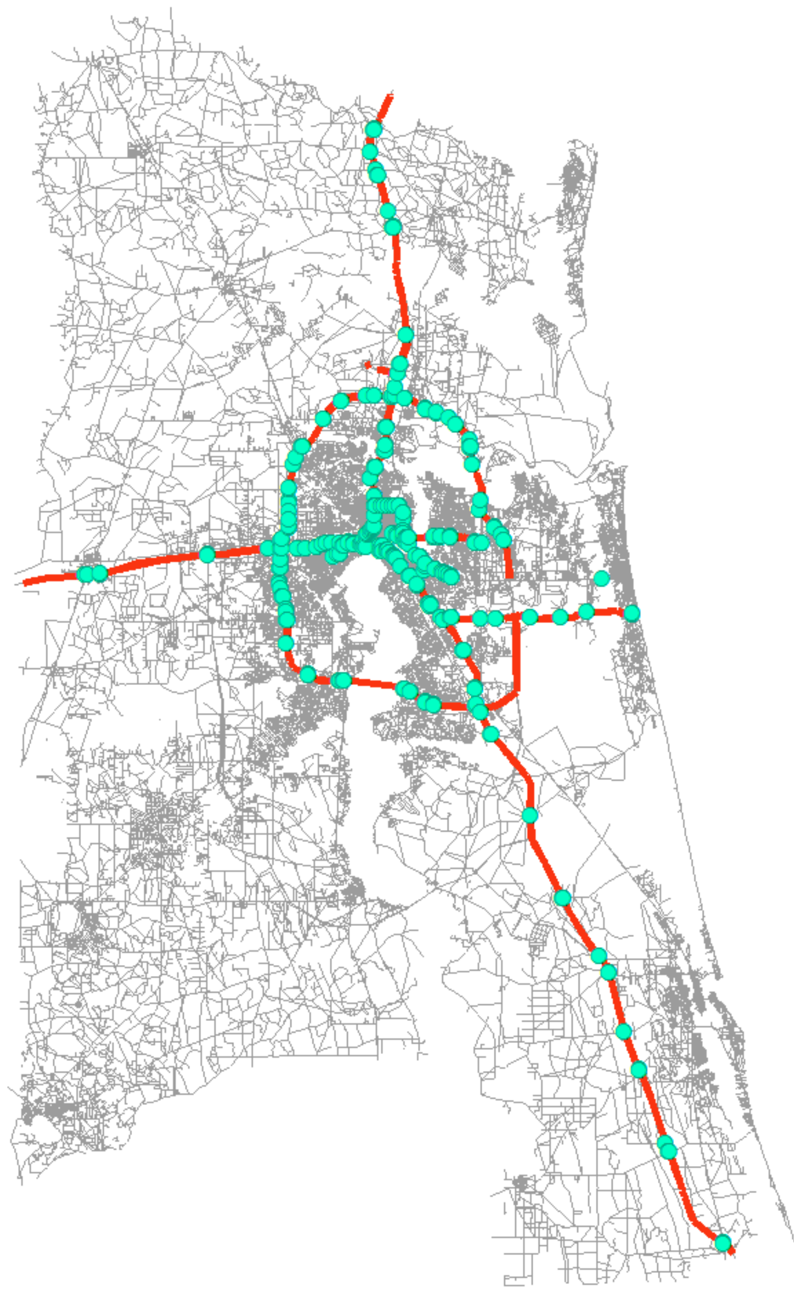
The MERGED.NET master network is conflated with the MERGED-GIS ESRI shape file and allows viewing the true shapes of the links in CUBE. The shape information is also exported to the output shape files for input to TransimsNet program.

Figure 29. NERPM Master Network



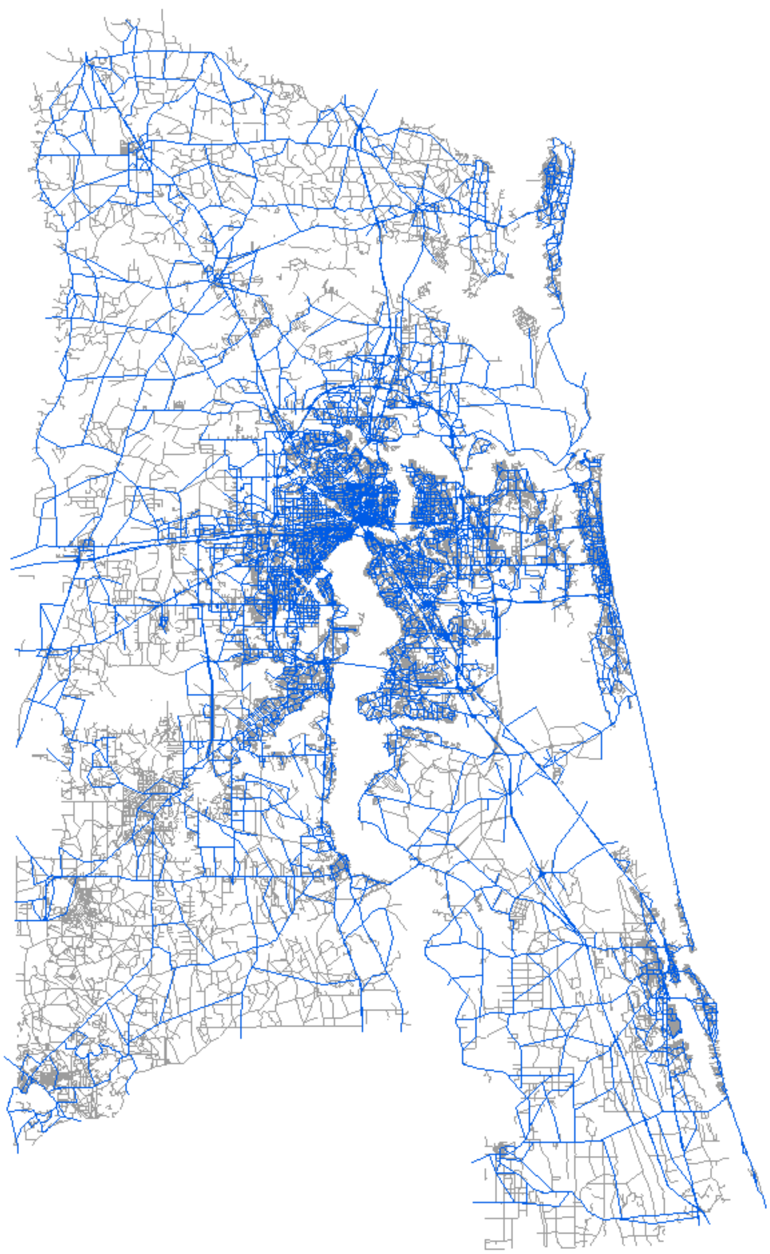
The NERPM model network topology is accomplished by having distinct nodes that are super-imposed. This distinction is not visible at normal zoom levels. However, grade-separate roadways have unique Anode-Bnode pairs relative to the underpass/overpass roadways. Figure 30 highlights nodes where links are seemingly intersecting with their cross streets, but in reality need to be represented as grade separations.

Figure 30. Locations of NERPM Super-imposed Nodes



The 2005 regional planning network has approximately 9,800 directional links, 6,500 nodes and 1,642 zones. Shape points or curvature information is available for approximately 6,000 directional links. The project team converted the 2005 TP+ network with time-of-day variations to a time-dependent TRANSIMS network. This network was subsequently used to route and simulate the TP+ trip tables.

Figure 31. NERPM Year 2005 Network



Note: the 2005 NERPM network is shown in light blue

The NERPM master network attributes are described in Table 40.

Table 40. NERPM Network Attributes

FTYPE_{year}{alt}	This is the facility type attribute. It distinguishes such facility types as freeway, arterials, collectors, etc. Values range from 11 to 99 and generally follow standard FSUTMS highway network coding practices. A facility type value of zero is used to indicate that a particular link is not present in a given scenario. For any given scenario, the facility type for many of the links in the master network is zero. This is one of three attributes used to calculate link speeds and capacities, the other two being area type and number of lanes.
ATYPE_{year}{alt}	This is the area type attribute. It distinguishes such land uses as CBD, Residential, Rural, etc. Values range from 11 to 55 and generally follow standard FSUTMS highway network coding practices. This is one of three attributes used to calculate link speeds and capacities, the other two being facility type and number of lanes.
LANES_{year}{alt}	This is the attribute that designates the number of directional lanes on any given link. Values range from 1 to 9 and follow standard FSUTMS highway network coding practices. This is one of three attributes used to calculate link speeds and capacities, the other two being facility type and area type.
IMPROV_{year}{alt}	This attribute indicates whether a particular link is a roadway improvement project that first becomes active in this scenario. Values are Y(es)/N(o).
AGENCY_{year}{alt}	This attribute identifies the agency responsible for making the roadway improvement as indicated above if that agency is known. Values are the names of the funding agencies.
DESC_{year}{alt}	This attribute describes in plain text the nature of the roadway improvement as indicated above.

2.6.2.2 NERPM Speed Capacity

The Jacksonville regional model stores speed and capacity information for links in TP+ formats. An example file for the base year 2005 is shown below in Figure 32. For a given area-type range, facility-type range and lane range combination, this file provides a corresponding speed and capacity value. The speed is considered to be free-flow speed and is translated as such during conversion to TRANSIMS formats.

Figure 32. NERPM Speed Capacity File

SPDCAP.05A

```
10101010 1 2  1810 45.0
10101010 3 3  1863 45.0
10101010 4 9  1891 45.0
10101111 1 2  1778 40.0
10101111 3 3  1863 40.0
10101111 4 9  1905 40.0
10101212 1 2  1810 45.0
10101212 3 3  1863 45.0
10101212 4 9  1891 45.0
10101515 1 2  1810 37.5
```

2.6.2.3 TAZ-Area Type Definition

Area-Type information is typically stored as an attribute of the TAZs in most regional models. TransimsNet was therefore designed to read this TAZ-Area-Type equivalence to allow users to control the generation of synthetic TRANSIMS network elements such as pocket-lanes by area-type in addition to facility-type. Figure 33 shows an example of the TransimsNet control parameters.

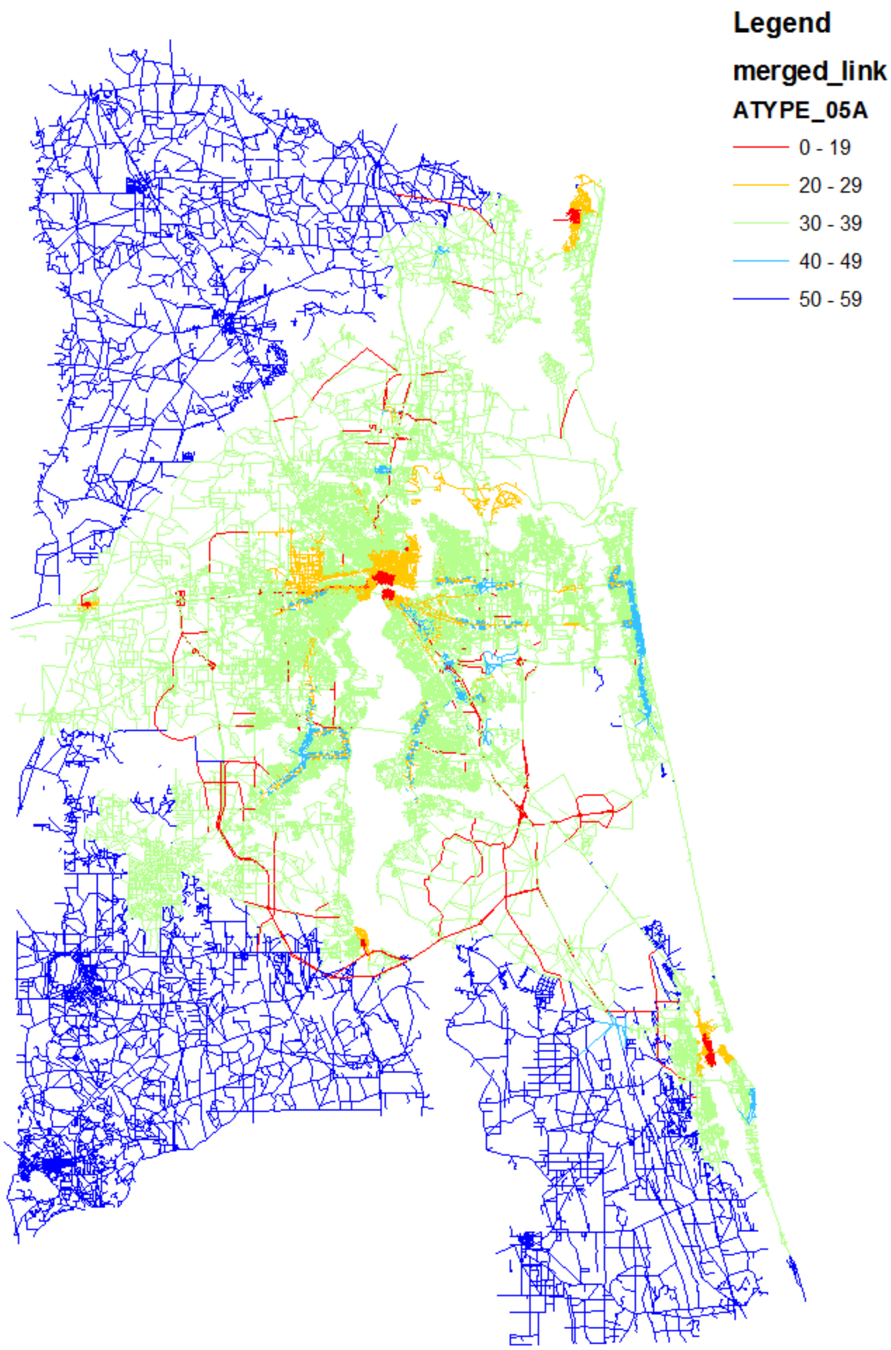
Figure 33. TransimsNet Controls by Area Type

```
FACILITY_TYPE_ACCESS_FLAGS         0,0,1                          #-- freeway,expressw
ACTIVITY_LOCATION_SIDE_OFFSET      25                             #-- meters

POCKET_LENGTHS_FOR_FACILITY_1      100, 200, 350, 400, 500        #-- freeway=1
POCKET_LENGTHS_FOR_FACILITY_2       60, 150, 200, 250, 300        #-- expressway=2
POCKET_LENGTHS_FOR_FACILITY_3       40,  80, 100, 125, 150        #-- principal=3
POCKET_LENGTHS_FOR_FACILITY_4       30,  70,  80,  90, 100        #-- major=4
POCKET_LENGTHS_FOR_FACILITY_5       30,  70,  80,  90, 100        #-- minor=5
POCKET_LENGTHS_FOR_FACILITY_6       30,  70,  80,  90, 100        #-- collector=6
POCKET_LENGTHS_FOR_FACILITY_7       30,  70,  80,  90, 100        #-- local=7,
POCKET_LENGTHS_FOR_FACILITY_9       60, 150, 200, 250, 300        #-- ramp=9
POCKET_TO_FACILITY_BY_AREA_TYPE      7,   7,   7,   7,   7        #-- allow pocket tur
```

However, the NERPM area-types are only available on links and not on zones. The area-types are also subdivided to range from 11 to 99. Figure 34 shows the distribution of area-types by link.

Figure 34. NERPM Link Based Area Types



Since the area-types were link-based, a representative area-type for a given TAZ could not be easily established due to the presence of multiple area-types within a TAZ. This problem was overcome by weighting with link-based area-types with their lane-feet in order to create a TAZ-AreaType equivalence.

2.6.2.4 Network Corrections

As mentioned earlier in this chapter, the super-imposed nodes in NERPM network were collapsed during the TransimsNet application to synthesize TRANSIMS network elements. Only the nodes that connect what would otherwise be a single continuous link are considered during the collapsing process. Attention is paid to link attributes such as the facility type, number of lanes and speeds. The nodes are not collapsed if any of these attributes differ or, in other words, only homogenous links are merged.

During the network conversion and simulation processes a number of issues with the network were revealed that required the study team to implement a series of network refinements. These issues can be divided into two basic types. The first type is refinements pertaining to large abrupt changes in roadway attributes such as facility type, thru lanes and speeds. These refinements were included in the TPPlusNet conversion scripts to maintain automated procedures to retain their applicability to future year networks and network alternatives. These abrupt changes can be classified into the following:

1. Discontinuities in facility types
2. Discontinuities in thru lanes
3. Discontinuities in speeds

Figure 35 shows the locations where such discontinuities were observed.

Figure 35. All Locations where TPPlusNet issued warnings

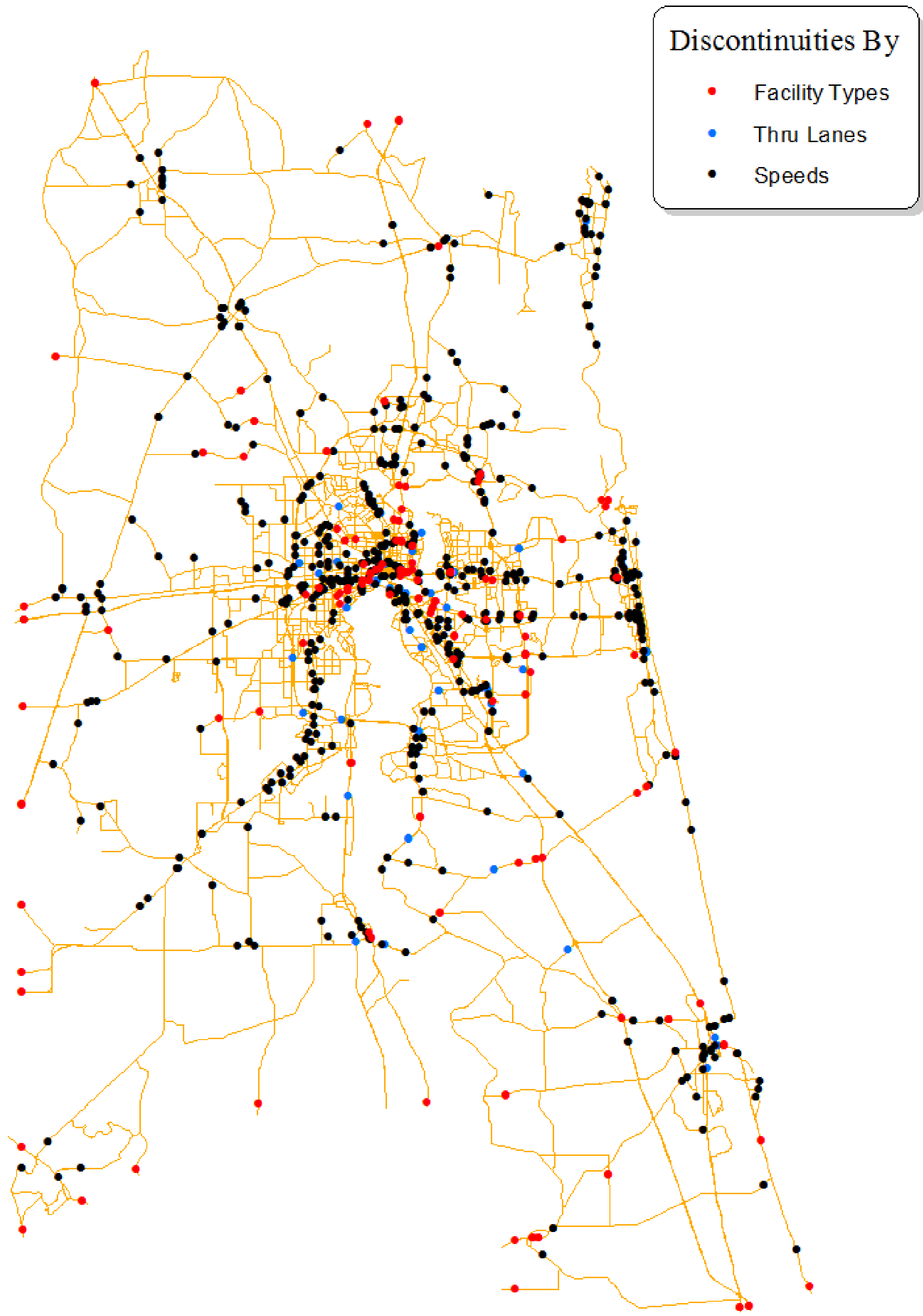
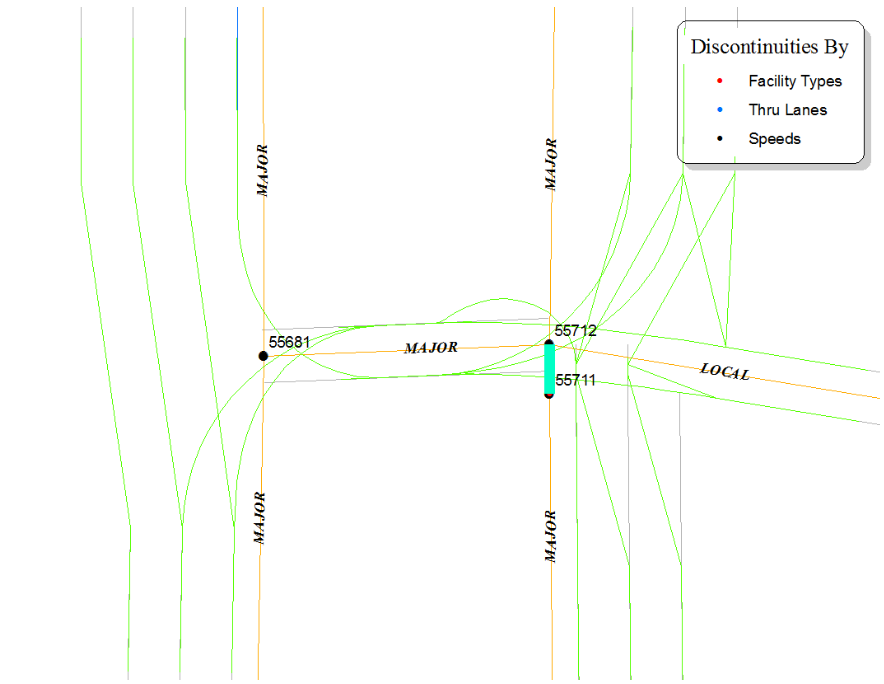


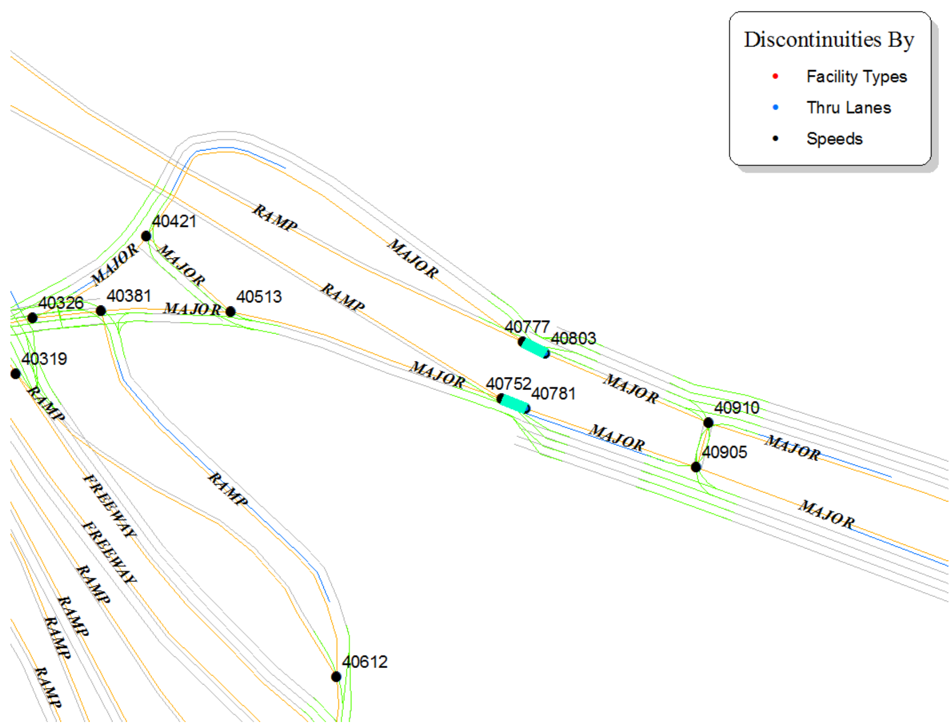
Figure 36 illustrates an example location where the functional class changed to local (FTYPE_05 = 46) for a short distance in between a continuous facility type of major arterial (FTYPE_05 = 23) as the link approached an intersection. This resulted in an extremely short link that complicated intersection operations.

Figure 36. Example location showing short link due to facility type change



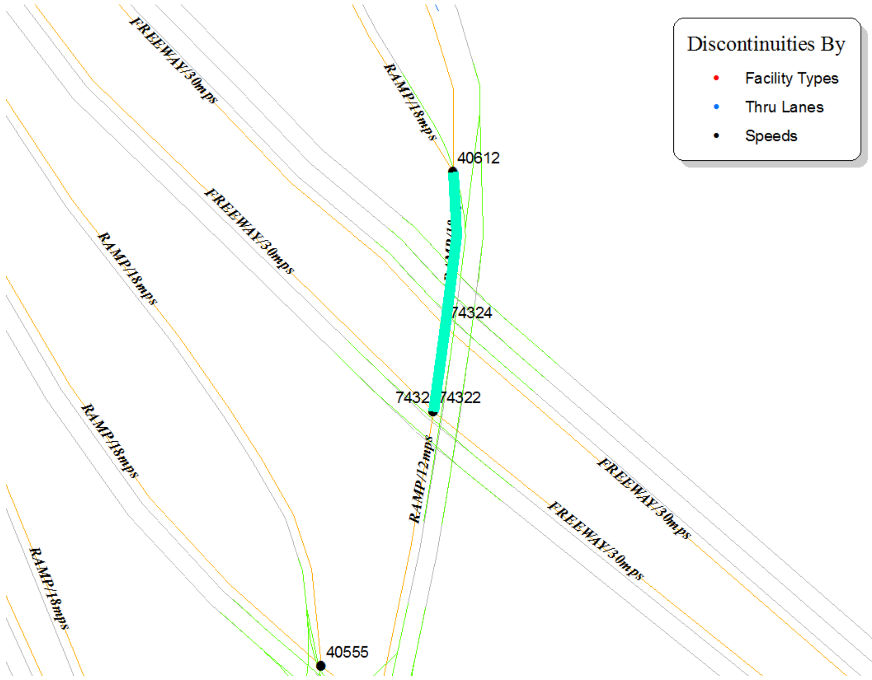
The next type of discontinuity was the unrealistic change of two or more thru lanes as seen in Figure 37. Since lanes are the primary source of capacity within the simulation and lane changing is one of the primary reasons for congestion and lost vehicles, these types of errors are extremely problematic.

Figure 37. Example location where the roadway cross-section changed abruptly



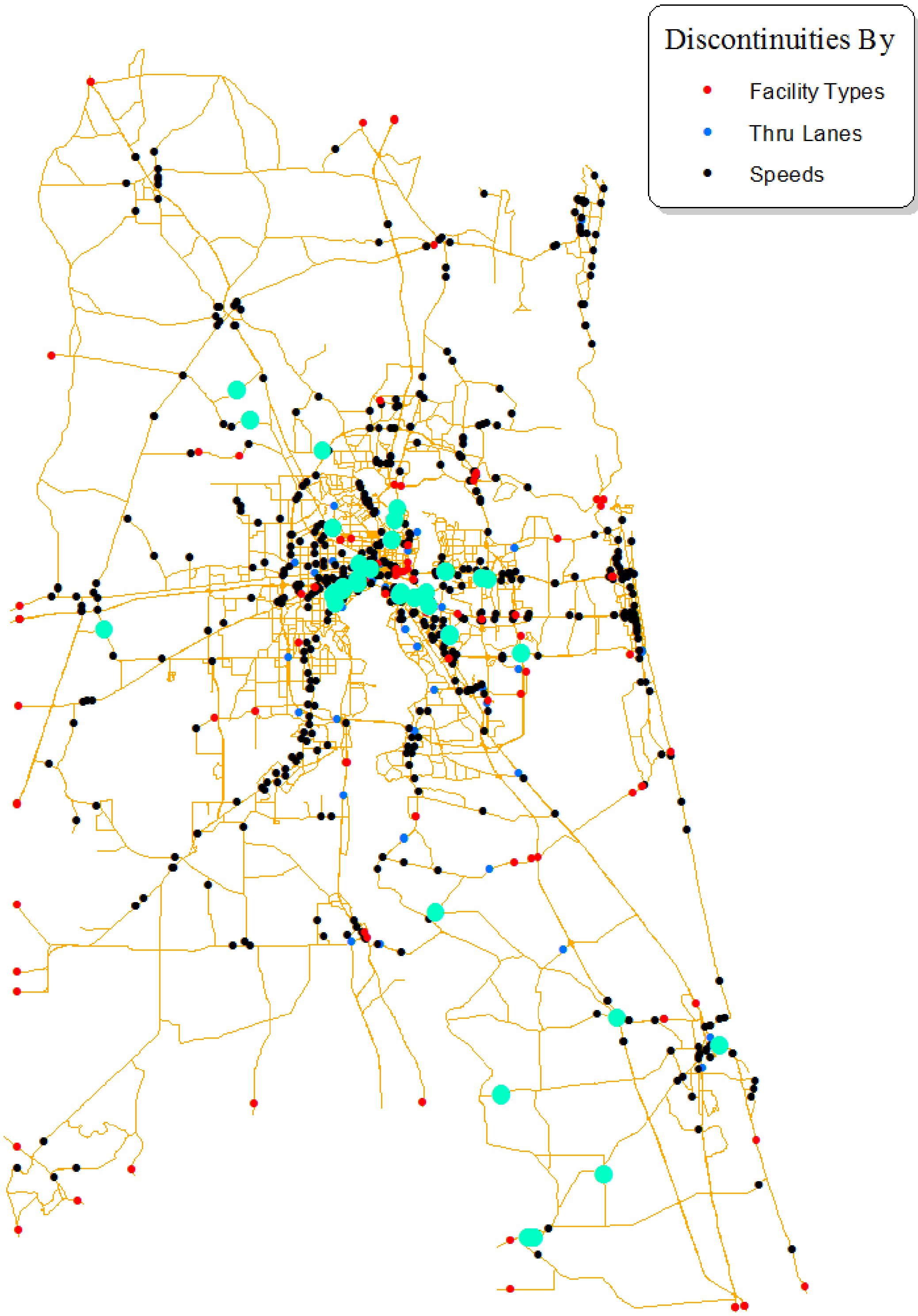
The last type of discontinuity was a large change in the speeds as seen below in Figure 38. The speed discontinuity was the result of a change in the coded area type for the link.

Figure 38. Example location with sudden drop in speeds



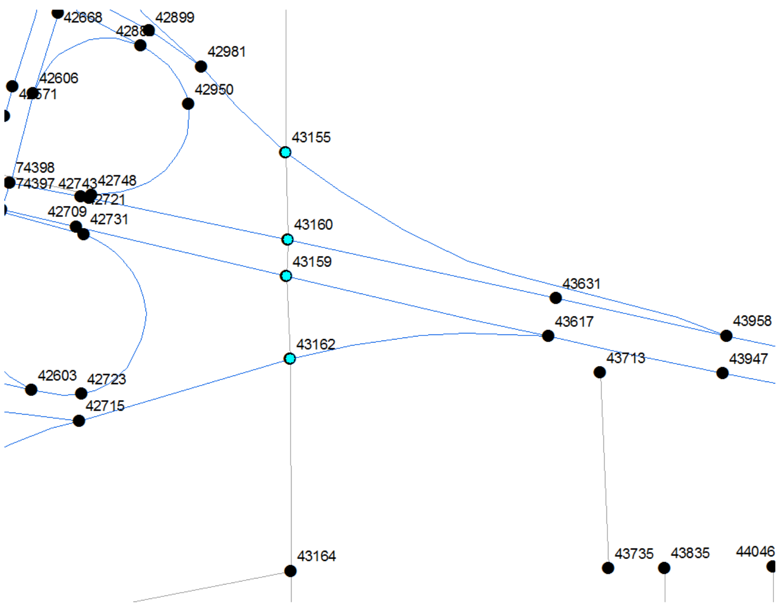
The above three coding errors result in a significant number of unnecessary nodes and short links which cause congestion problems in the simulation. Most of these issues were addressed by modifying the TPPlusNet conversion script and collapsing the nodes. Figure 39 below shows the distribution of locations were these nodes were dropped. Since the edits were performed at the input level, they can easily be applied to all network resolutions and analysis years.

Figure 39. Highlighted nodes were collapsed as a result of removing discontinuity in link attributes



The second group of issues relate to more systematic problems raised by the assignment visualization. The visualization results showed significant backups on freeways at several locations far away from their real access points. Investigation showed several locations where freeways intersect with arterial roadways due to network coding errors in the input NERPM master network. These were locations where links should have been coded using grade separated nodes, but the link was assigned to the wrong node number. An example is shown in Figure 40. Again, to address this issue, network edits were performed to the NERPM TP+ network to keep the conversion procedure intact.

Figure 40. NERPM Freeways Intersection Arterials Due to Coding Errors



2.6.2.5 Intersection Controls

Traffic Signals

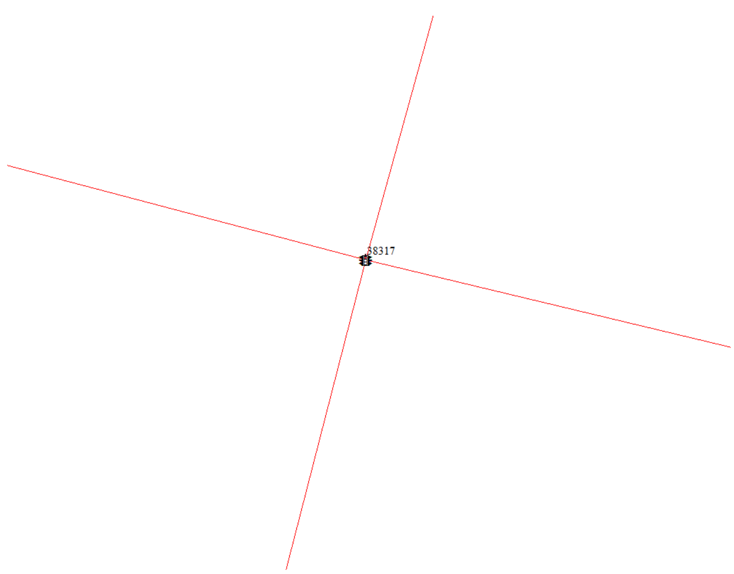
TRANSIMS includes a number of ways of changing the configuration of the network by time of day. In addition to the roadway configurations, traffic controls can also vary by time of day. The signal timing and phasing plans can be adjusted to optimize time-of-day flow conditions. Signal progression tools are also available to coordinate fixed time signals along specified corridors or throughout a grid system. Demand-actuated signals can include multiple detectors and simulate ramp metering behavior. The signal formats also allow changing signal types by time of day.

A rich data set containing the location of all signals in ESRI shape file format was available from FDOT for the entire state of Florida. As discussed in earlier sections, this signal location information was used to replace rule-based signal warrants from TransimsNet. However, some processing was required by the project team to interface this data with TRANSIMS programs. There was no equivalence between these signals and the NERPM node numbers. The signals had to be spatially or manually tagged to the 2005 scenario of the NERPM master network nodes. As an outcome of this process, one or more nodes were identified as part of a single signal in the Jacksonville region. This is due to the fact that in a number of cases a single real-world intersection is represented by two or more nodes in the NERPM network. For example, when a divided arterial is represented as a pair of two one-way links, their intersections are represented by two nodes.

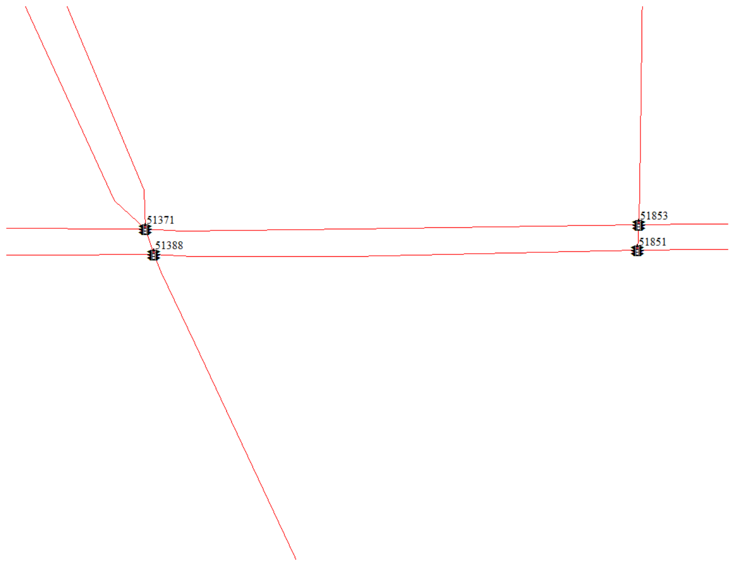
The tagging process was useful in two respects; (1) it served as a direct resource for the replacement of rule-based synthetic signal warrants in TRANSIMS Version 4 network conversion across all the three network resolutions, and (2) it provided for leveraging the group control feature of TRANSIMS Version 5 signal format wherein a single signal could be defined over more than one node.

The following steps describe how this process of matching FDOT signal locations with the NERPM master network (MERGED.NET) was carried out:

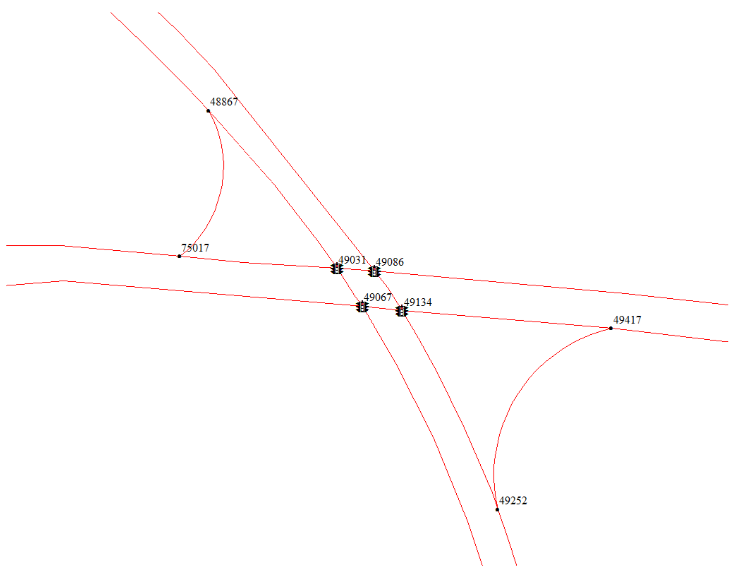
1. The merged_node and merged_link shape files were projected onto the coordinate system of the traffic_signal_locations shape file which was NAD_1983_UTM_Zone_17N.
2. Using a spatial join, all of the nodes were first joined to the nearest traffic signal. A filter was used so that only nodes whose joining distance was less than or equal to 50 feet were considered.
3. When deciding which nodes should be attached to a specific traffic signal there were a few general cases that came up regularly that were used to make this decision. These cases are explained below with examples. The RED nodes indicate the nodes that were attached to the traffic signal.
    a. If two single-link roadways intersected at a single node intersection, then only that node was attached to that traffic signal.



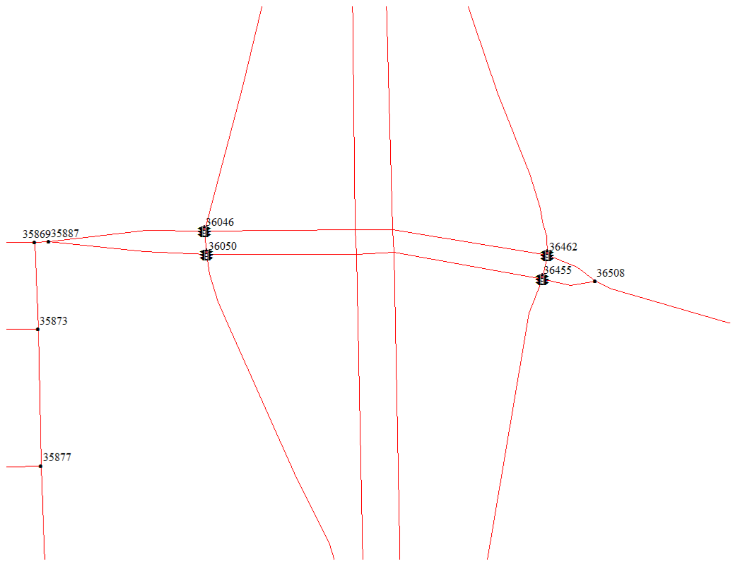
    b. If a single-link roadway intersected a double-link roadway, then the nodes at the two intersection points were attached to the traffic signal.



c. If two double-link roadways intersected, then all four nodes of the intersection were attached to the traffic signal.



d. At freeway on and off ramps there were usually two traffic signals; one on either side of the freeway. Usually only two nodes were attached to each traffic signal as shown.



Traffic signal intersections that varied from these examples were evaluated on a case-by-case basis. The intersections were examined using Google Earth and a determination was made as to which nodes belonged to each signal.

4. Only traffic signals with a VALUE_ field of "02" were used in this mapping. Others such as flashing beacons and school signals were omitted.
5. The output shapefile with mappings was formatted for use with TransimsNet.

Unsignalized Intersection

The sign warrant creation logic in TransimsNet was used to synthesize traffic controls for the un-signalized intersections. This logic considers the facility class levels and their differences in addition to the area-type information to determine whether and what type of sign control is required for each of the approaching roadways at an intersection. The TransimsNet program allows the user to define rules for generating signal warrants and places sign warrants for the remaining intersections. Since signal location information was used from FDOT dataset, no rules were specified in TransimsNet, resulting in the creation of sign warrants for all intersections in the region. The IntControl program was later provided with the FDOT signal location-based signal warrants and TransimsNet generated signal timing and phasing plans and sign controls for the intersections. Naturally, there were many conflicting sign and signal warrants, where signal warrants were preserved and sign warrants were discarded.

2.6.2.6 Network Conversion Process

The network conversion process starts with applying TPPlusNet to convert TP+ network formats to generic TRANSIMS link and node format. During this application, the coordinate system is also converted from Florida StatePlane 0901 in Feet to UTM 17N in meters. The resulting TRANSIMS network is maintained in metric units. This process is depicted in Figure 41.

The TPPlusNet conversion script is critical to the conversion process. It maps the NERPM functional codes to TRANSIMS facility type strings and populates the hourly capacity, number of lanes, maximum speed and free flow speed fields.

Table 41 shows how the TP+ functional class codes were mapped to TRANSIMS facility types.

The number of lanes coded in the NERPM network represents "all-day" travel lanes and excludes parking and turn lanes. In TRANSIMS, all lanes are coded along with link names and distances, which are converted from feet into meters.

The generic TRANSIMS link, node and shape files are then provided to TransimsNet along with TAZ area-type equivalence file to generate synthetic TRANSIMS network elements as shown in Figure 41.

To overcome the problem of visualizing network topology at grade separated crossings, the collapse-nodes feature in TransimsNet was used to merge homogenous links and thereby reduce the number of input nodes that are kept in the output network.

Following this step, synthetic intersection controls are created using the process shown in Figure 42 and the activity location fields are updated and expanded to include zonal attributes.

Figure 41. Network Conversion Process

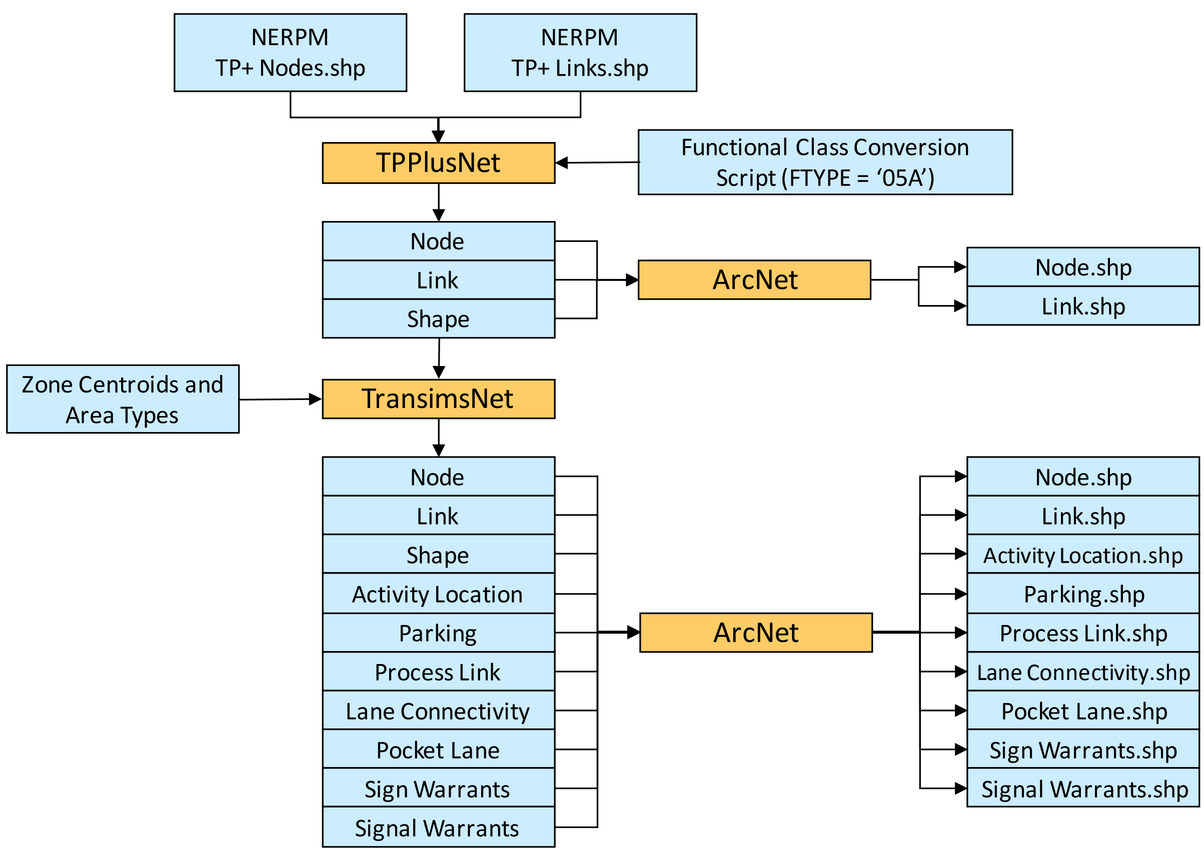
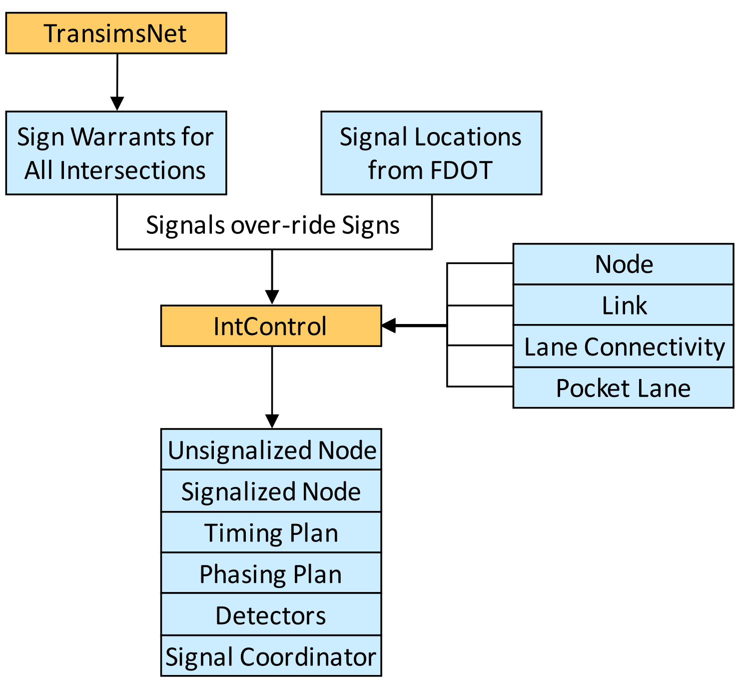


Table 41. TRANSIMS Functional Class Mapping

FUNCLASS Codes	TRANSIMS Facility Type
11-12, 79-84, 90-91	FREEWAY
16-17, 61, 85, 93	EXPRESSWAY
70-78, 86-89, 97-98	RAMP
20-22, 60, 94-95	PRINCIPAL
23-25, 62-63	MAJOR
30-38, 64	MINOR
15, 40-43	COLLECTOR
65-68	FRONTAGE
44-49	LOCAL
29	FERRY
50	ZONE CONNECTOR
52	EXTERNAL
92	OTHER

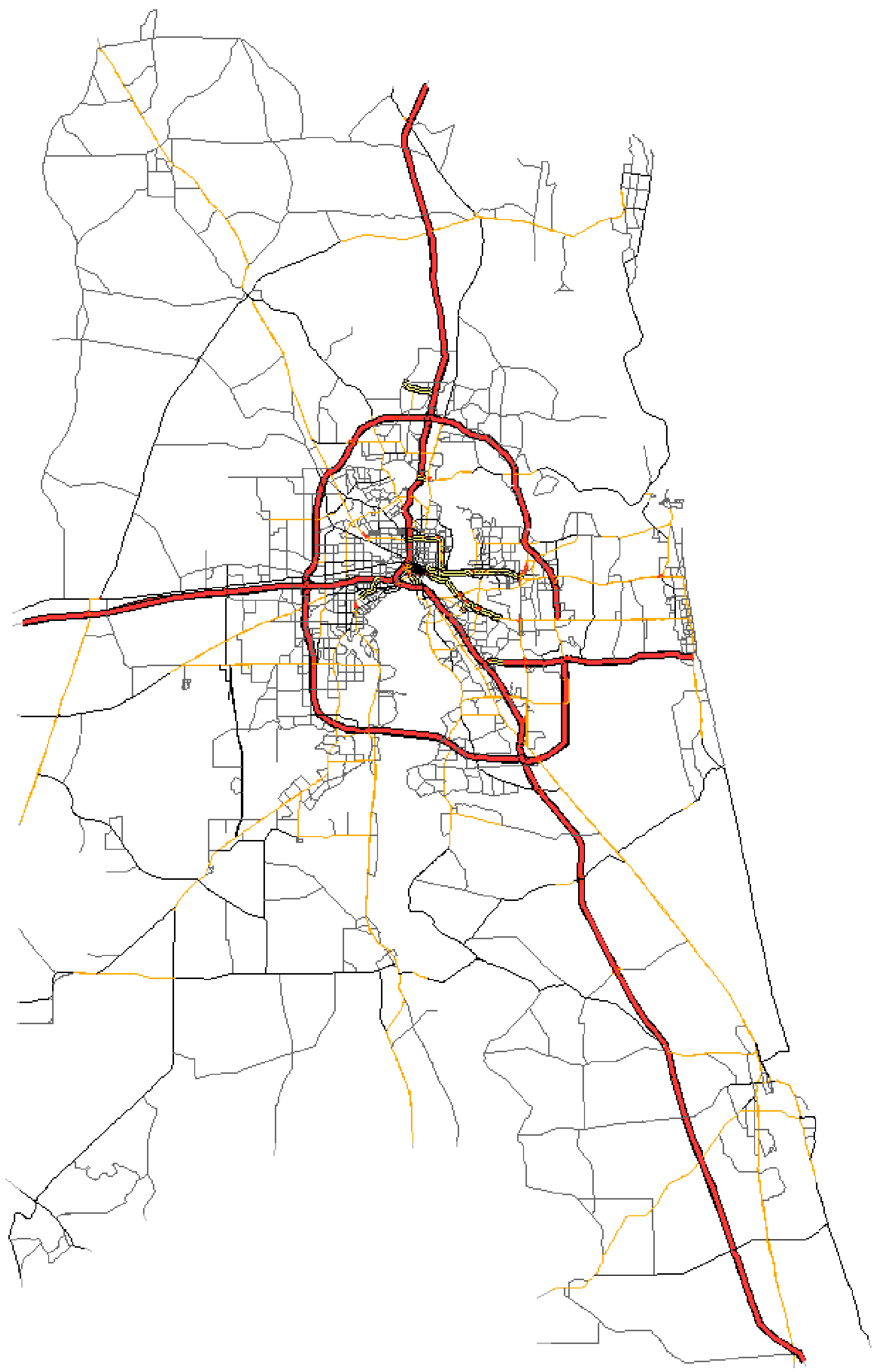
Figure 42. Signalized and Un-signalized Intersections



2.6.2.7 PLANNING Network

The PLANNING network is a network resolution including only the NERPM regional modeling links. The corresponding scenario in NERPM model is year 2005 or '05A'. Only the records defined in the '05A' scenario, i.e., records containing non-zero values for field "FTYPE_05A" were included. The resulting network is shown in Figure 43. This network includes 6,525 nodes and 9,864 links.

Figure 43. Resulting PLANNING Network

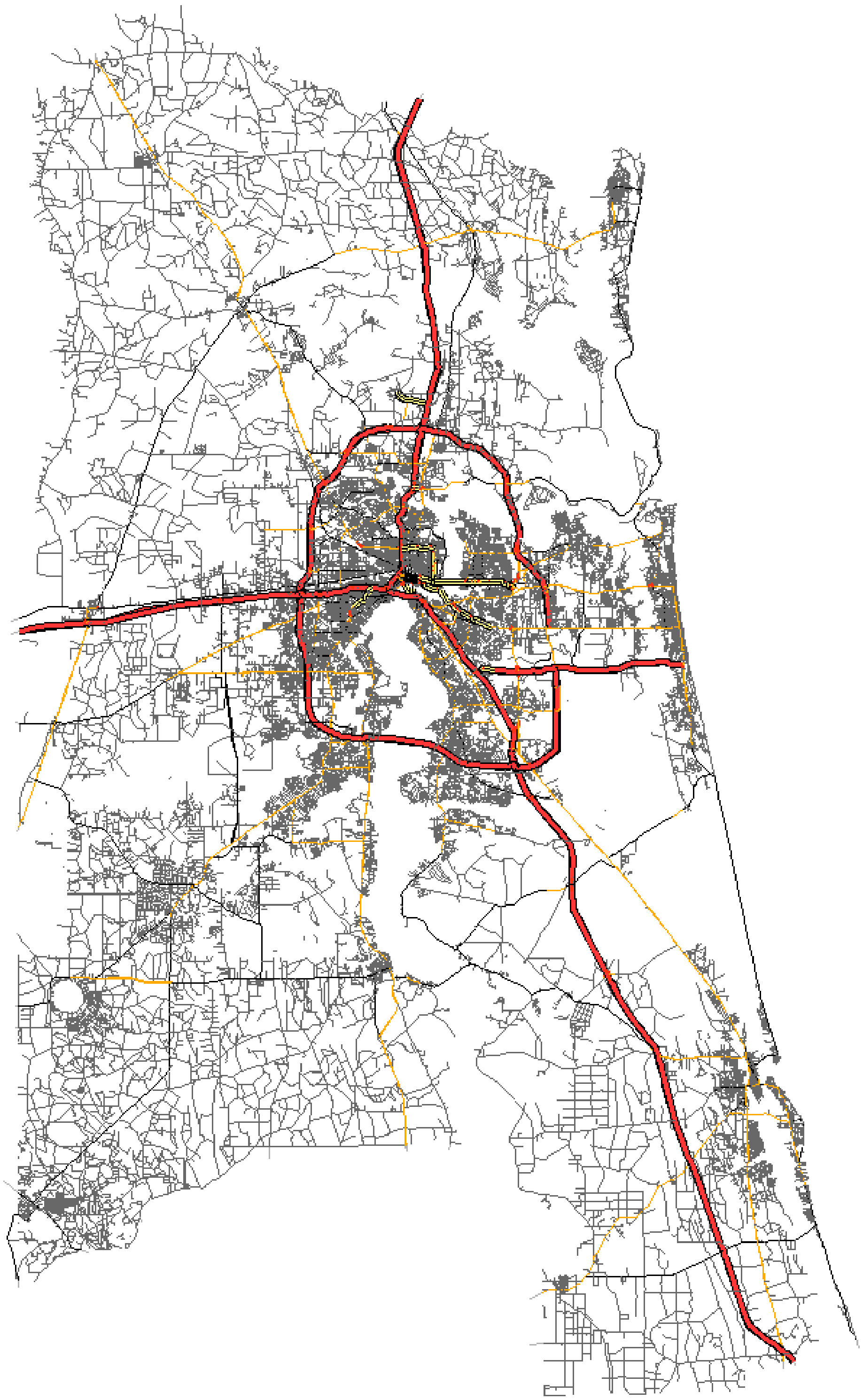


2.6.2.8 ALLSTREETS Network

The PLANNING network and the ALLSTREETS networks were envisioned as the two ends of network resolutions for this project. While the PLANNING network was limited to the NERPM modeling links for the 2005 scenario, the ALLSTREETS network utilized links that were not defined in any of the NERPM scenarios. Such links in the NERPM master network are assumed to be existing streets of very low facility class for instance, neighborhood streets, and do not carry any significant levels of traffic. These additional links were categorized as locals in TRANSIMS, equivalent to NERPM facility number '42' and marked with a common name for identification purposes ("ADDON_DETAIL").

Modification of the TPPlusNet conversion script to include these additional streets was the only difference in the ALLSTREETS network conversion process compared to that of the PLANNING network. The resulting TRANSIMS network is shown in Figure 44. It contains 51,420 nodes and 69,361 links.

Figure 44. Resulting ALLSTREETS Network



2.6.2.9 FINEGRAINED Network

In accordance with one of the primary goals of this project, a fine-grained network was developed as an intermediate resolution network for greater policy sensitivity and increased fidelity without the huge computational overhead associated with the ALLSTREETS network. Since additional local through streets are the fundamental difference between the ALLSTREETS and the PLANNING networks, different filtering methodologies can produce several different intermediate resolution networks.

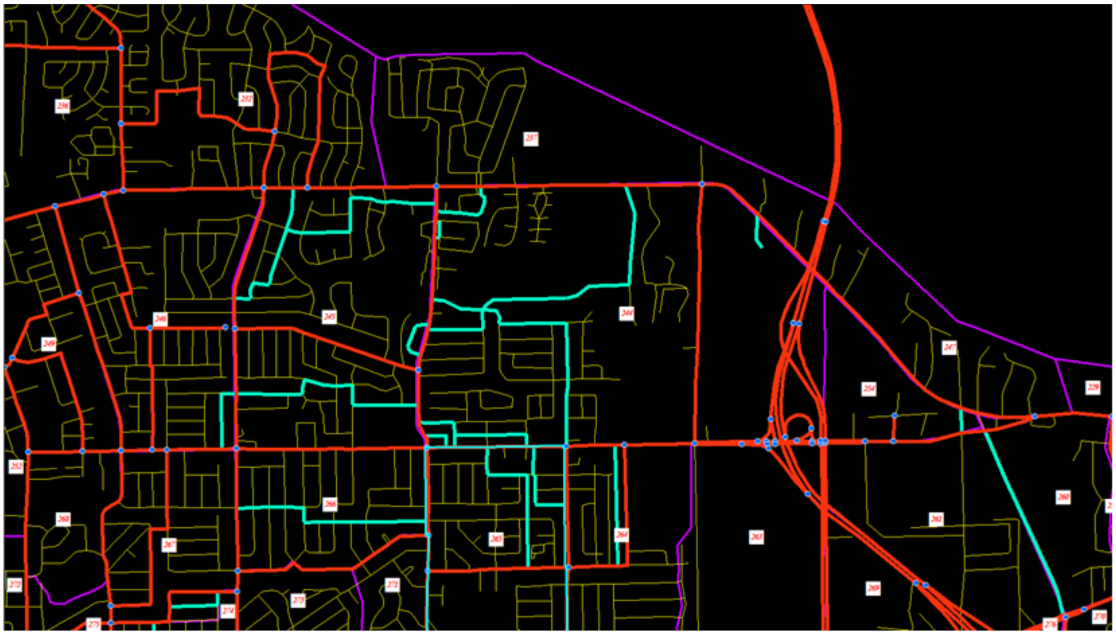
The selection process, however, needs to address the following considerations to have any reasonable or meaningful impact on the TRANSIMS simulation process:

1. While all streets bring more realism to the modeling process, they come with heavy computational costs, potentially resulting in very large and unreasonable processing times.
2. Additionally, not all streets bring value to the TRANSIMS simulation modeling. For example, long dead-end links may provide better estimates for disaggregate travel times, but do not impact nearby links or path alternatives; hence the simulation results would more or less be un-changed.
3. The chosen network resolution would need to strike a balance between reasonable representation of roadway accessibility and its additional burden on model processing times.

Based on these considerations, an approach balancing the mobility and accessibility factors within the project scope was developed. The presence of additional network detail provides the opportunities to consider a wider range of path options or diversions to, from and within traffic analysis zones. Links were selected for inclusion in the FINEGRAINED network based on the frequency in which they were used on paths between zone origins and destinations.

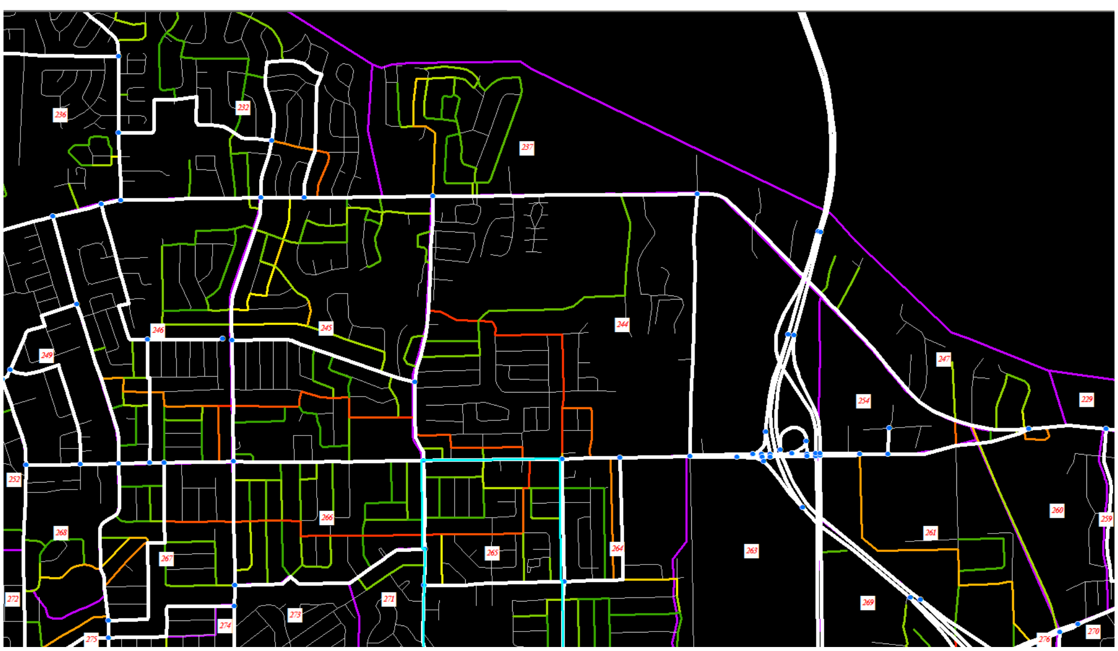
Figure 45 demonstrates this conceptual approach showing planning links in red, zone boundary in purple, additional links in brown and potential links for selected zones in blue based on distance-based paths built from a single zone.

Figure 45. Example of Qualified Links



Since the analysis of mobility options was the primary motivation for this approach, the PLANNING network is used as a benchmark for keeping a subset of the additional ALLSTREETS links that provided new path-diversions. Multiple levels of network resolutions can thus be created from this subset by filtered based on an accessibility "score". This score could alternatively be considered a parameter for choosing the intermediate network resolution. The score would compute the percent of the regional employment, population and households within a given distance from roadways. Figure 46 shows an approach for further filtering the additional streets by examining volume levels.

Figure 46. Filtering Qualified Links by Volume Levels

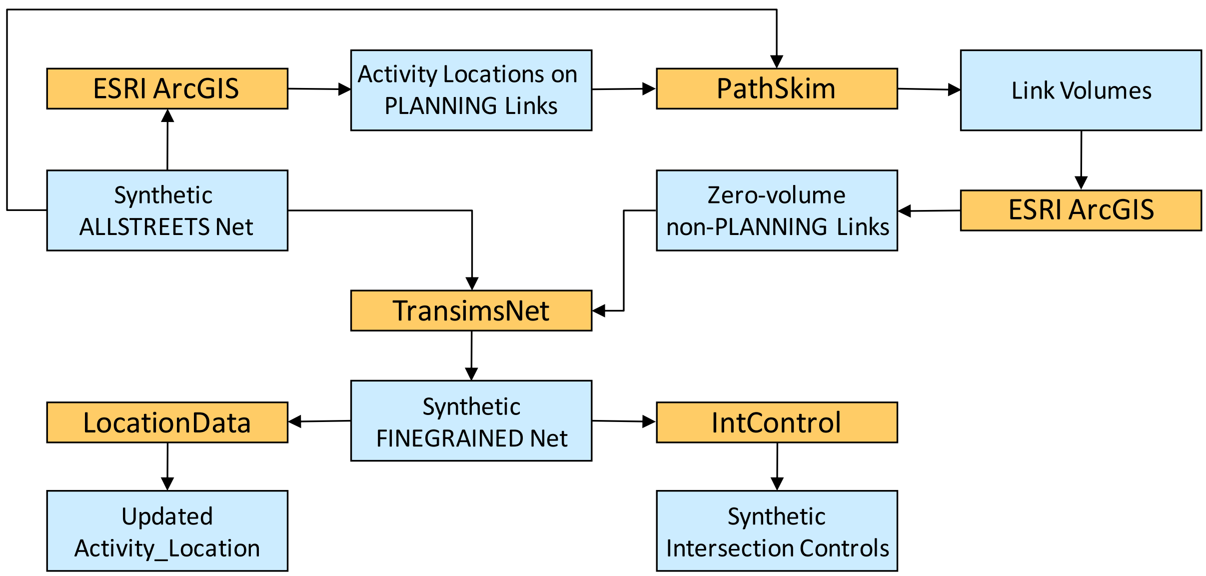


This process of selecting local thru streets requires building numerous paths between targeted activity locations. The TRANSIMS Version 5 program, PathSkim, with enhanced capabilities was used for this purpose. PathSkim is backwards compatible with Version 4 network formats which makes its application straight forward.

Additionally, it was found that deleting undesirable streets from the ALLSTREETS network was an easier way to create the FINEGRAINED network than adding selected streets to the PLANNING network.

Figure 47 shows the process of creating a FINEGRAINED network given the PLANNING and ALLSTREETS networks using PathSkim.

Figure 47. FINEGRAINED Network Conversion Overview



The accessibility scores for the resulting FINEGRAINED network were computed by measuring the Euclidean distance of each parcel in the region from its spatially nearest roadway. The distribution of employment, population and household in the region is summarized by this distance in Table 43.

Table 42. Accessibility Computation

S. No.	Distance (m)	Distance (miles)	PLANNING							FINEGRAINED						
			Parcels	Emp	Pop	HH	% Emp	% Pop	% HH	Parcels	Emp	Pop	HH	% Emp	% Pop	% HH
1	16.1	0.01	10,385	12,333	15,294	6,264	2%	1%	1%	19,495	17,370	31,512	12,822	3%	3%	3%
2	32.2	0.02	47,365	61,108	70,707	28,651	11%	6%	6%	81,706	82,396	138,684	56,141	14%	12%	12%
3	48.3	0.03	80,569	151,652	124,839	50,505	26%	11%	11%	125,866	176,814	216,632	87,690	31%	18%	18%
4	64.4	0.04	111,272	223,312	175,343	71,028	39%	15%	15%	161,358	248,365	279,140	113,074	43%	24%	24%
5	80.5	0.05	138,861	275,838	232,602	94,367	48%	20%	20%	190,431	299,207	341,108	138,304	52%	29%	29%
6	96.6	0.06	163,117	327,106	286,118	115,977	56%	24%	24%	214,491	350,131	394,641	159,915	60%	33%	33%
7	112.7	0.07	187,394	362,727	344,530	139,612	63%	29%	29%	237,538	385,291	450,734	182,627	66%	38%	38%
8	128.7	0.08	210,519	400,430	399,719	161,872	69%	34%	34%	258,688	419,907	501,369	203,119	72%	42%	42%
9	144.8	0.09	231,568	426,750	452,307	183,124	74%	38%	38%	277,403	443,605	550,639	222,992	77%	47%	47%
10	160.9	0.10	249,946	445,855	499,480	202,156	77%	42%	42%	293,863	461,575	594,959	240,874	80%	50%	50%
11	241.4	0.15	327,751	506,998	686,590	278,188	87%	58%	58%	361,953	515,980	758,228	307,196	89%	64%	64%
12	321.9	0.20	385,118	530,225	818,972	332,153	91%	69%	69%	412,523	537,705	870,301	352,815	93%	74%	74%
13	402.3	0.25	426,539	545,337	902,178	365,841	94%	76%	77%	450,015	550,012	943,052	382,199	95%	80%	80%
14	482.8	0.30	458,811	555,324	957,838	388,304	96%	81%	81%	479,454	559,577	992,539	402,117	97%	84%	84%
15	563.3	0.35	484,129	562,376	1,005,355	407,632	97%	85%	85%	502,681	565,587	1,035,266	419,450	98%	88%	88%
16	8046.7	5.00	618,981	579,535	1,182,431	478,072	100%	100%	100%	618,981	579,535	1,182,431	478,072	100%	100%	100%

The FINEGRAINED network chosen for this project had the following accessibility score:

- 80% of the regional employment is within 1/10th of a mile of a modeled roadway.
- 50% of the regional population is within 1/10th of a mile of a modeled roadway.
- 50% of the regional households are within 1/10th of a mile of a modeled roadway.

Figure 48 through Figure 50 show the employment, population and household distributions for the PLANNING and FINEGRAINED networks.

Figure 48. Employment Accessibility

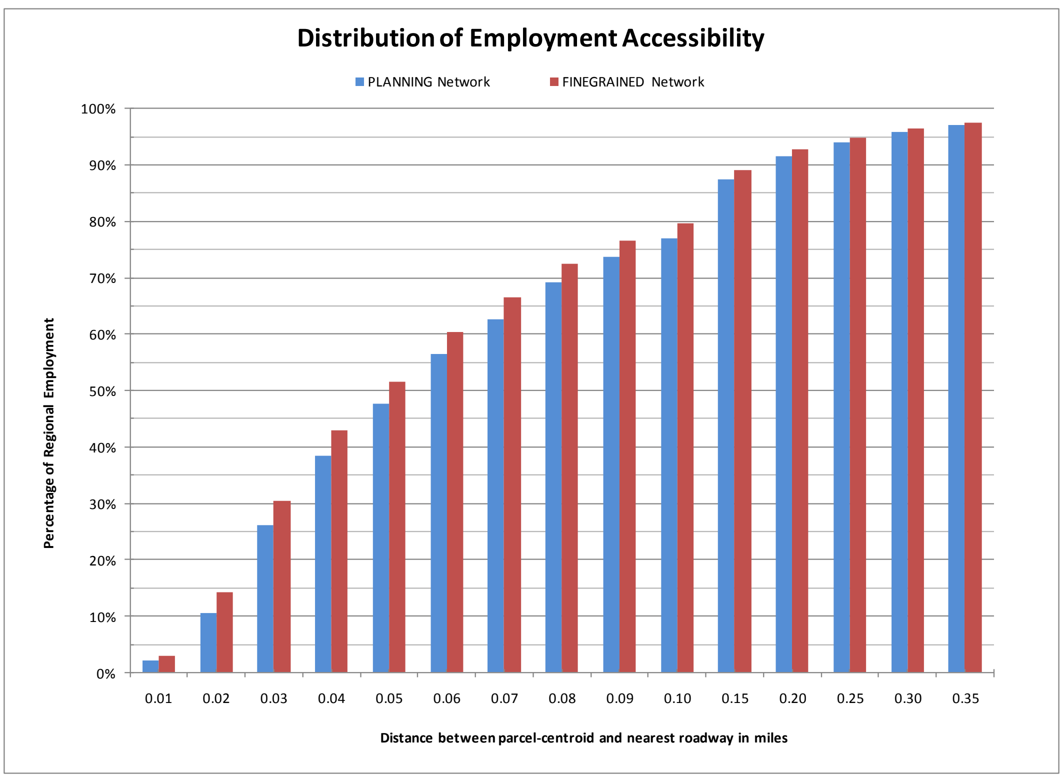


Figure 49. Population Accessibility

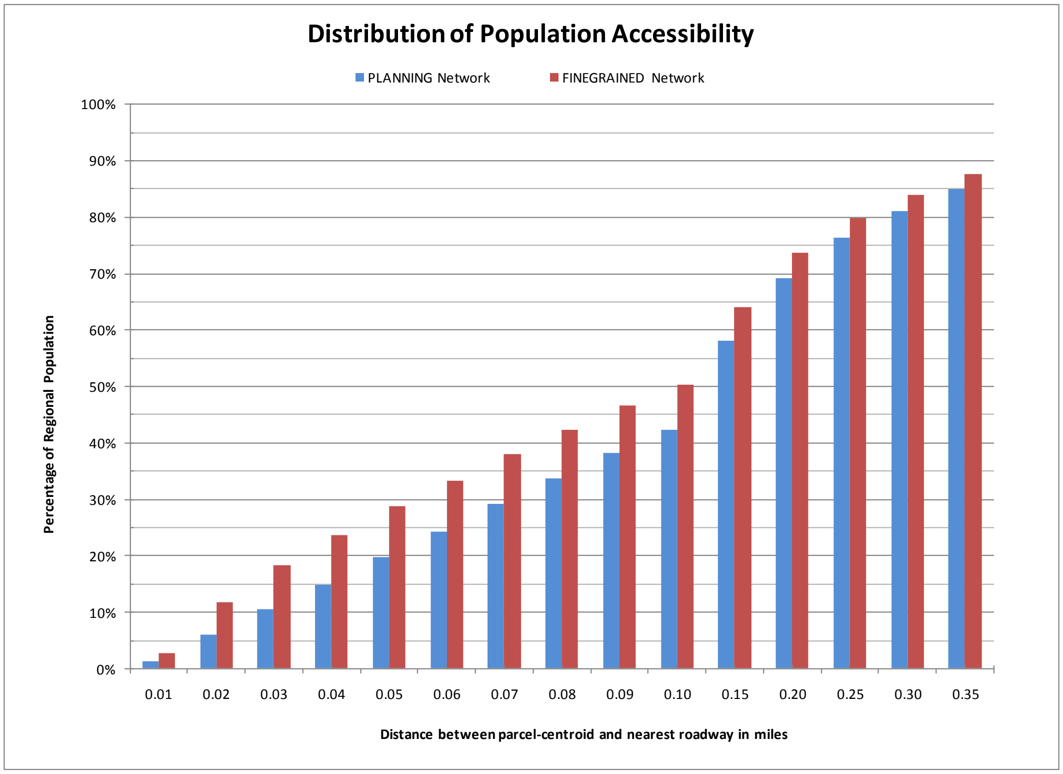
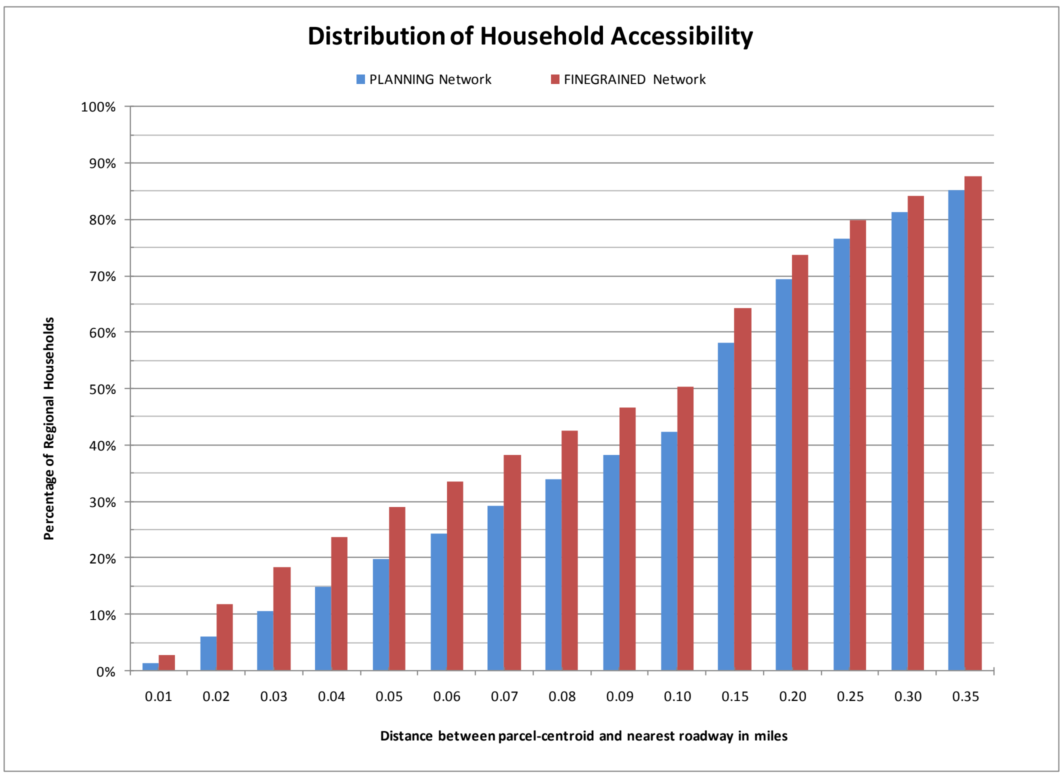


Figure 50. Household Accessibility



The resulting FINEGRAINED network is shown in Figure 51. This network includes 10,577 nodes and 16,910 links. Figure 52shows the distribution of the additional network detail in the FINEGRAINED network in contrast to the PLANNING network. A comparison of the network attributes included in the three network resolutions is provided in Table 46.

Figure 51. Resulting FINEGRAINED Network

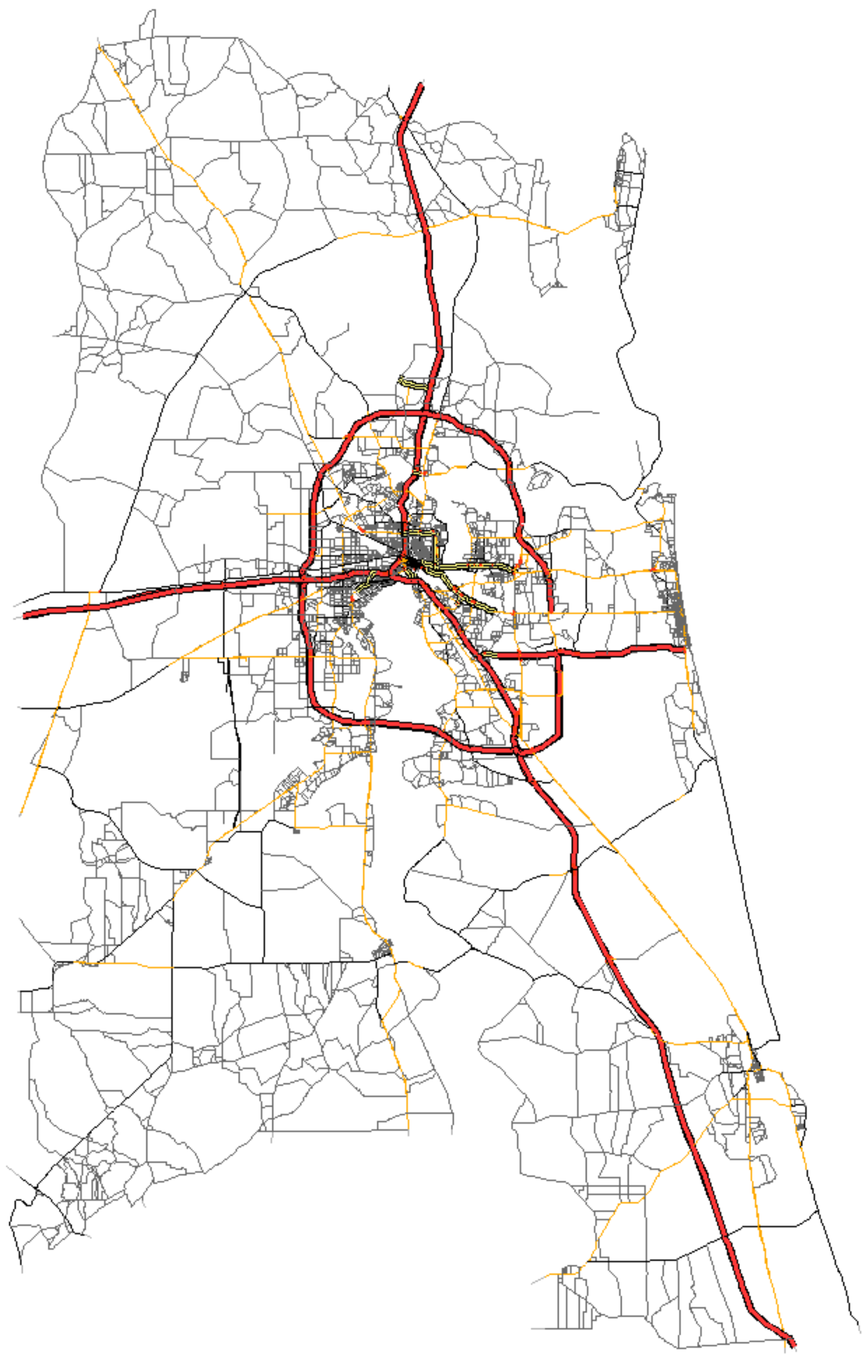


Figure 52. Distribution of Network Detail in FINEGRAINED Network compared to PLANNING Network

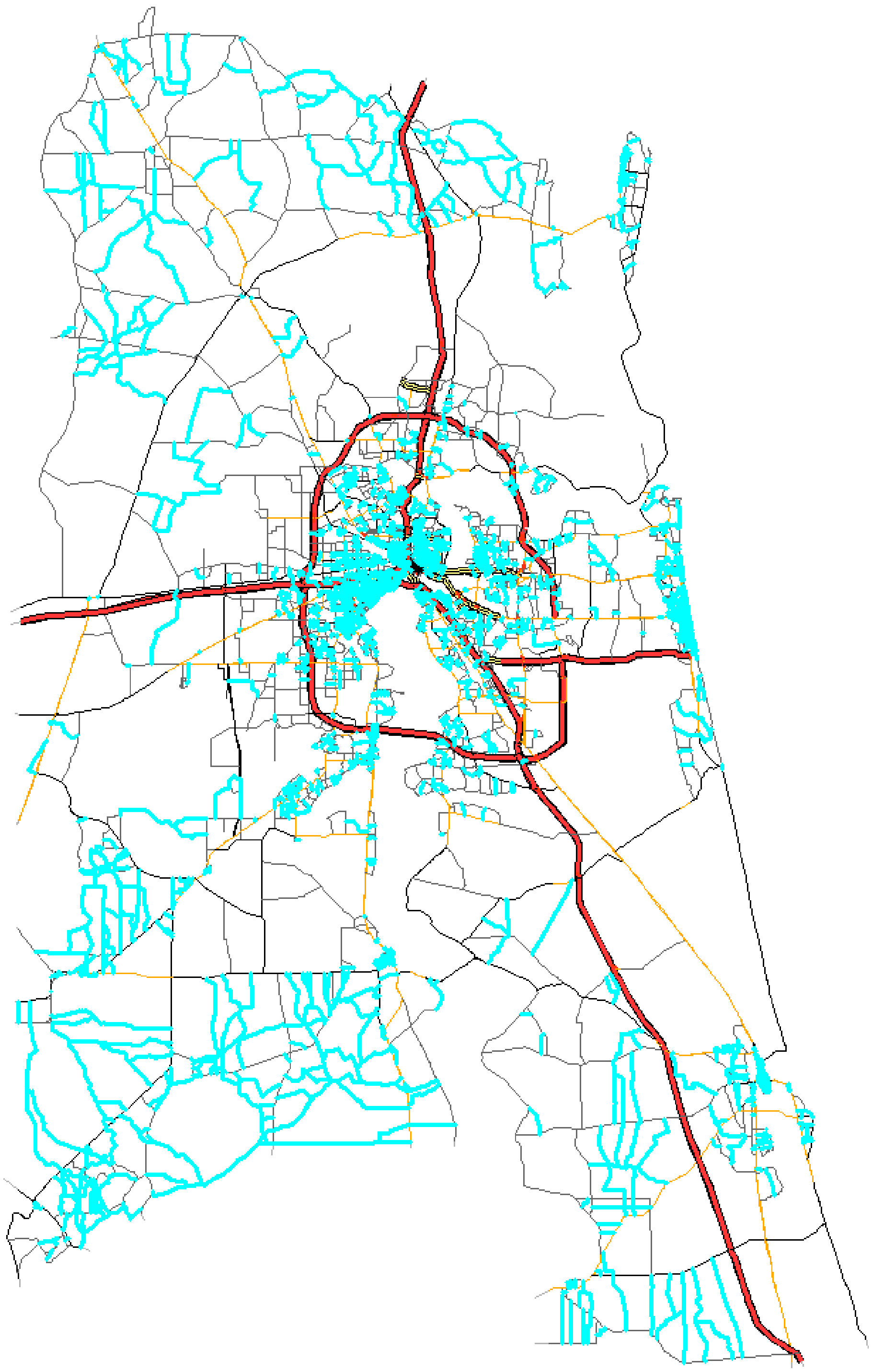


Table 43. Network Attributes Comparison

Attribute	PLANNING	ALLSTREETS	FINEGRAINED
Nodes	6,519	51,305	10,577
Links	9,850	69,106	16,910
Links with Shape Points	5,960	25,250	9,345
Activity Locations	29,272	216,246	58,846
Parking Lots	29,272	216,246	58,846
Process Links	58,544	432,492	109,692
Pocket Lanes	6,461	76,152	16,574
Lane Connections	37,758	303,106	74,302
Unsignalized Nodes	3,820	36,215	8,458
Signalized Nodes	934	1,140	1,058

2.6.2.10 Network Conversion Summary

The overall network conversion process is summarized in the following steps:

1. Enable “True Shape Display” in CUBE utilizing Merged-GIS ESRI shape-file and TP+ network files.
2. Export the TP+ network data files to link and node ESRI shape-file format. Note that it is not required to perform any attribute-based filtering to create specific scenarios in this step. The entire master network is exported. The filtering process is performed in subsequent steps.
3. Run TPPlusNet to convert the TP+ link and node files to generic TRANSIMS input files. During this process a conversion script is used to translate NERPM facility type codes, number of lanes, speeds and capacities to TRANSIMS coding rules. Zone centroids and connectors are excluded. For the PLANNING network, records containing non-zero values for field “FTYPE_05A” are selected. For the ALLSTREETS network, links that are not defined in any other scenarios are included as locals and marked with a separate name. Warnings are raised if area-type or lanes are not defined for such links, since those are required to obtain speed and capacity information from the speed-capacity lookup table. These warnings are examined and appropriate values for missing attributes are coded based on neighboring links. This step is repeated until all the warnings are addressed.
4. Run ArcNet to review basic link information and network continuity related to facility types, speeds, capacities and the number of lanes.
5. Run TransimsNet with the shapes file to create the synthetic TRANSIMS network files such as pocket lanes, lane-use, lane connectivity, parking lots, activity locations, process links and signal and sign warrants. Note that rules for creating signal-warrants are not provided in TransimsNet hence the resulting signal warrants file is empty. In the absence of these signal-rules, all of the intersections in the region are “filled” with sign-warrants, as it were. The FDOT signal location information takes the place of signal-warrants and replaces the pre-filled sign-warrants wherever applicable.
6. Run ArcNet to visualize and review the resulting network. The focus of this review is pocket lanes and intersection connectivity. The location of signals and signs are typically reviewed and edited as well.

7. Run IntControl using the signal and sign warrants to generate the signal timing and phasing plans, demand actuated detectors, and sign controlled intersections.
8. Run ArcNet again to review the traffic control data.
9. Run LocationData to post the zonal attributes to activity locations and update their TAZ numbers based on the supplied TAZ boundary layer. This method uses a point-in-polygon approach to update the initial TAZ number assignment to activity locations in TransimsNet based on Euclidean distances from zone-centroids.
10. Run ArcNet to visualize the resulting network.

The process used to create the FINEGRAINED network requires the output of steps 1 through 5 for both the PLANNING and ALLSTREETS networks. The process then implements the following additional steps:

1. In ArcGIS: from the ALLSTREETS network, flag all links (street name equal to "ADDON_DETAIL") not part of the PLANNING network.
2. In ArcGIS: select all activity locations in the ALLSTREETS network that are located on links included in the PLANNING network and save this list to an output file. This file will be supplied to PathSkim to create paths between activity locations.
3. Run PathSkim to build paths between these selected activity locations. This can potentially be a very time consuming process (~55,000 x ~55,000 = ~2.8 billion paths) taking over 46 hours on an eight core workstation. Save the output Link Delay file from PathSkim in text format without any turning movement data. Note that it is not necessary to build paths for all activity location combinations. A sufficient number of samples, approximately 10 per zone, requires much less processing time and is quite representative.
4. In ArcGIS: join the PathSkim output Link Delay file to the ALLSTREETS link shape-file to flag all links that have non-zero volumes. Be sure to include all PLANNING network equivalent links. Flip this selection and save the list of non-PLANNING zero-volume links to a file.
5. Run TransimsNet in Update mode with the ALLSTREETS network and a list of deleted links as input. This process will delete the requested links and their corresponding parking lots, process links and activity locations. In addition, all nodes, links, and lane connections are also refreshed including the evaluation of sign and signal warrants.
6. Continue the process outlined in steps 6 through 10 above.

2.6.3 Burlington Network Development

The Burlington TRANSIMS network was developed following a similar process to that described for Jacksonville, although only a single network was built. The TransCAD-based highway network maintained by the Chittenden County Metropolitan Planning Organization (CCMPO) for use in their daily regional travel demand model was used as the starting point for developing a detailed microsimulation TRANSIMS network for the region.

The 2005 base year model highway network has approximately 1,700 links and 1,300 nodes which represent the major roadway facilities in the county. Interstate I-89 is the only interstate highway in the county which serves the 18 cities and towns in Chittenden County, most notably Burlington, the largest city in the State of Vermont. The TransCAD base year highway network is a typical "planning" level network though the links do reflect true shapes while zonal access to the street grid is modeled with centroid connectors. Table 44 presents the extent of roadway network coverage by facility type. Figure 53 presents the 2005 base year CCMPO TransCAD 4-step planning model network.

Table 44. CCMPO TransCAD Base Year Network Lane Miles by Facility

Facility	Links	Lane Miles	% Share
Interstate	66	166	11%
Limited Access Hwy	18	18	1%
Principal Arterial	285	159	10%
Minor Arterial	178	163	11%
Major Collector	163	174	11%
Urban Local	54	165	11%
Rural Major Collector	313	305	20%
Ramps	80	18	1%
Internal Centroid	530	325	21%
External Centroid	17	32	2%
Total	1,704	1,523	100%

Figure 53. CCMPO TransCAD Base Year Network

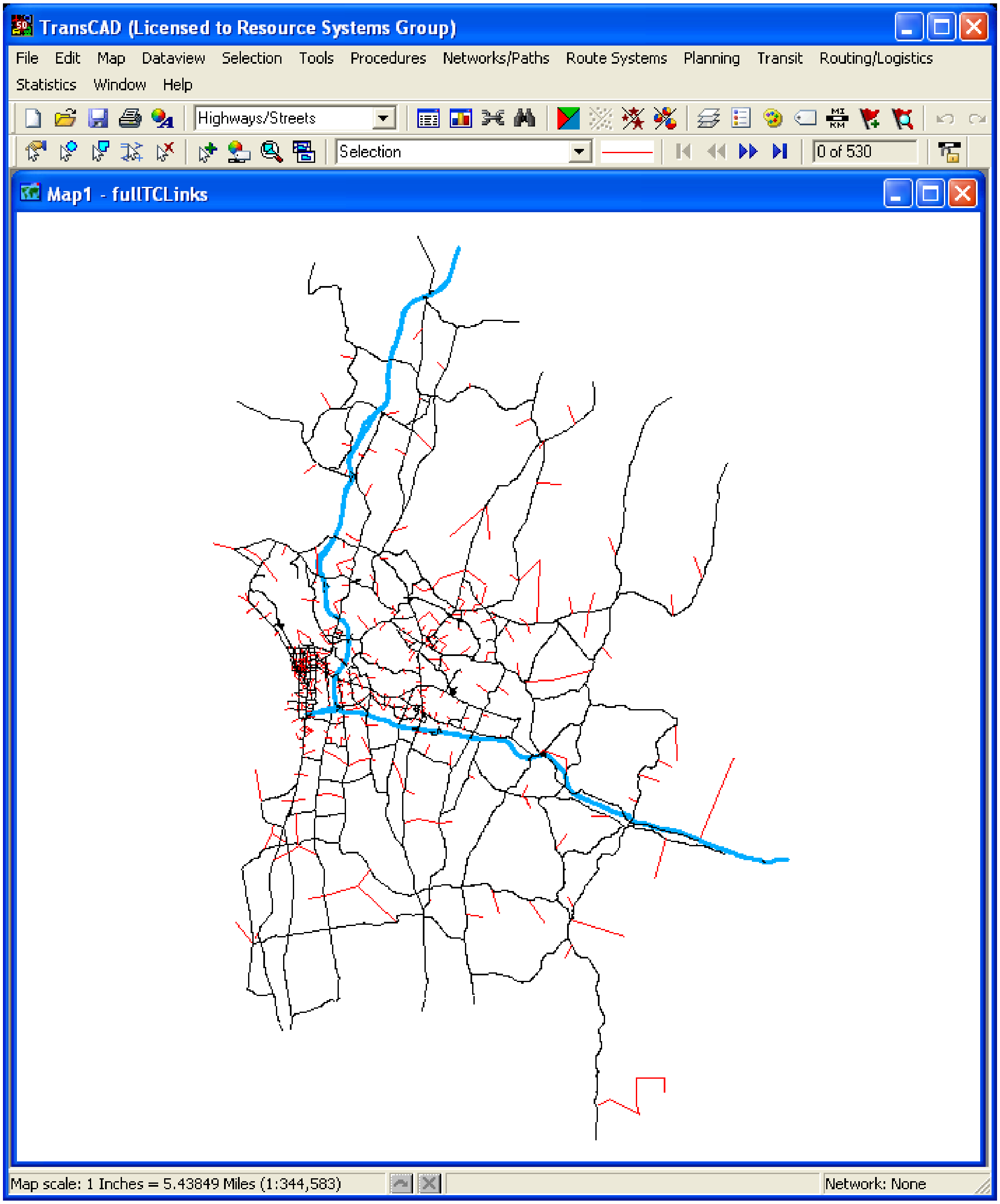


Table 45 below provides the number of records in each of the TRANSIMS network files generated by the network preparation described above. The Burlington TRANSIMS network for the 2005 analysis year has 524 nodes, 779 links and 2,608 activity locations. To synthesize TRANSIMS simulation network data, the number of links into and out of a given node was used along with intersection logic to construct turn pockets, lane connectivity and traffic controls, both signs and signals. Unique logic was applied depending on the facility type of the link. For instance, arterial intersections examine the relative orientation of each movement and the functional class of each link to determine when and where to include turn pockets and signals or signs. In general, if an approach has opposing traffic, a turn pocket was added to accommodate the movement. The signal warrants are determined based on the number of legs and the user-specified functional class by area type signal warrant parameters.

Table 45. Burlington TRANSIMS Network Size

Network File	Records
Nodes	524
Links	779
Activity Locations	2,608
Parking Lots	2,608
Process Links	5,216
Pocket Lanes	310
Lane Connectivity	3,100
Unsignalized Nodes	328
Signalized Nodes	114

Activity locations were automatically synthesized using the TransimsNet utility. The program creates activity locations (loading points for TRANSIMS) along every block face separated by a user-specified location spacing variable (eg. 100 meters). A minimum block length of 30 meters and no more than 3 activity locations per block face are two additional criteria which dictate the placement of activity locations in the simulation network. The automated procedure works very well but has two problems which typically needed to be corrected subsequent to running TransimsNet. The first was that in the more rural areas of Chittenden County, there were traffic analysis zones in the 4-step model which were not associated with at least a single activity location. This typically occurs in cases where the traffic analysis zone represents open land with very little road frontage and/or where the roadway network is sparse. An ArcMap overlay of the traffic analysis zones on top of the activity locations was utilized to manually associate TAZs with activity locations on the nearest appropriate roadway for the cases not automatically allocated correctly by TransimsNet.

2.6.3.1 Enhancing the Synthesized Network Integrity

The automatic synthesis of activity locations using TransimsNet can also produce loading points where no loading should occur in reality, for example in the middle of highway interchanges. Rather than manually remove these activity locations, a polygon layer representing the areas where this occurred was built. Geographic rules were then applied to systematically remove all locations within these undesired polygon locations. This provided an automated means of importing new and future year 4-step planning networks and automatically correcting/updating the activity locations synthesized by TransimsNet.

2.6.3.2 Regional Signal Retiming

The original Burlington TRANSIMS network was developed as part of an earlier FHWA TRANSIMS Demonstration project, and during this project effort the project team elected to review and update the fixed traffic signal timing and phasing plans developed as part of the original network development. A regional signal re-timing and re-phasing of the traffic signals in the simulation network was performed using the TRANSIMS utility IntControl. An automated and iterative re-timing and re-phasing of the traffic signal data was conducted using 90-second cycle lengths and link volumes resulting from the simulation of the increased regional demand. The signal timings and phasing were iteratively updated until link flows reached an acceptable level of calibration against observed ground counts.

2.7 Auxiliary Demand

2.7.1 Jacksonville

DaySim provides detailed estimates of the long-term and short-term travel choices of Jacksonville residents when traveling within the region, but this travel demand doesn't fully represent all trips that use the regional transportation networks. Commercial and truck traffic comprise a significant share of all roadway volumes, typically up to 20% or more. In addition, non-residents enter the region through key external gateways to access jobs, shopping or other opportunities, or may simply pass through the region. Similarly, residents may leave the region to satisfy other needs. Special generators may also create demand not explicitly represented person travel demand models.

"Auxiliary demand" refers to the regional demand that is not forecast by the DaySim model system, but that must be represented in the Jacksonville DaySim-TRANSIMS-MOVES integrated model system in order to reasonably assess network performance and the impacts of different policies or improvements. This auxiliary demand is derived from the existing NERPM model system currently used in Jacksonville, with spatial and temporal detail added to support integration with the detailed demand and supply simulation models. Because this demand is exogenous to the DaySim-TRANSIMS model system, the total demand, and the spatial distribution, mode, and timing of these trips is fixed within a given forecast or horizon year, but will of course vary across model run years. However, network times and costs influence the routes used, so the network assignment of this auxiliary demand is not fixed

The auxiliary demand in Jacksonville can be generally grouped into four main classes: internal-internal commercial vehicle trips, internal-external personal and commercial vehicle trips, external-external personal and commercial vehicle trips, and internal-internal special generators. Table 46 lists the demand components of each of these four main classes, the total trips associated with each component, and the relative share of total regional demand that this component represents. This table shows that approximately 19% of the total regional demand in the DaySim-TRANSIMS model system is derived from auxiliary demand, which seems generally consistent with reported practice. However, a closer inspection reveals that about 12% of total regional demand is comprised of commercial vehicles.

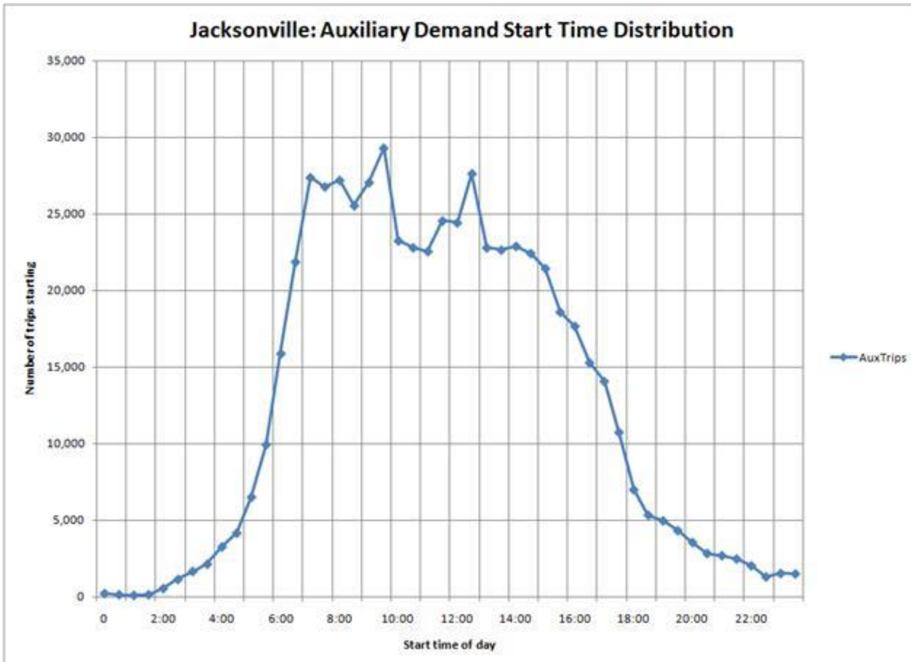
Table 46. Jacksonville Auxiliary Demand Summary

Auxiliary Demand Segment	Total Trips	Regional share
II Four-wheeled truck	269,695	8%
II Single-unit truck	75,957	2%
IICombo Truck-trailer	30,190	1%
IE SOV	94,640	3%
IE HOV	48,158	1%
IE Light-duty	3,318	0%
IE Heavy Duty	14,188	0%
EE SOV	22,709	1%
EE HOV	17,864	1%
EE Light-duty	1,686	0%
EE Heavy Duty	8,867	0%
Airport	41,080	1%
Total Auxiliary Vehicle Demand	628,352	19%
Total DaySim Vehicle Demand	2,708,077	81%
Total Vehicle Demand	3,336,429	100%

Regarding special generators, it should be noted that the original NERPM model contains more special generators than were ultimately included in the DaySim-TRANSIMS model system, including state parks, military bases, and malls. However, only the airport special generator was ultimately maintained in the integrated model system. In some cases, it was not necessary to treat the NERPM locations as special generators because the employment and population assumptions used in DaySim generate sufficient demand. The primary examples of this are with military bases and university group quarters. In other cases the data was not used due to counterintuitive patterns.

The auxiliary demand was temporally disaggregated from daily numbers using the same household survey information that was used in the original NERPM trip matrix conversion process. Figure 54 shows the diurnal distribution of trip start times. Future refinements to this temporal disaggregation process may include using vehicle class-specific or external station-specific diurnal distributions derived from traffic counts, or using scheduled airport takeoffs and landings to impute the temporal distribution of travelers coming from or going to the airport. It is also necessary to disaggregate the auxiliary demand spatially from TAZs down to TRANSIMS activity locations, which are the fundamental spatial units used in the TRANSIMS network assignment process. The subzone distribution of the trips is based on simple activity-location weights, though additional refinements such as the use of size variables reflecting employment and population can be easily implemented.

Figure 54. Auxiliary Demand Time-of-Day Distribution



2.8 Component Integration

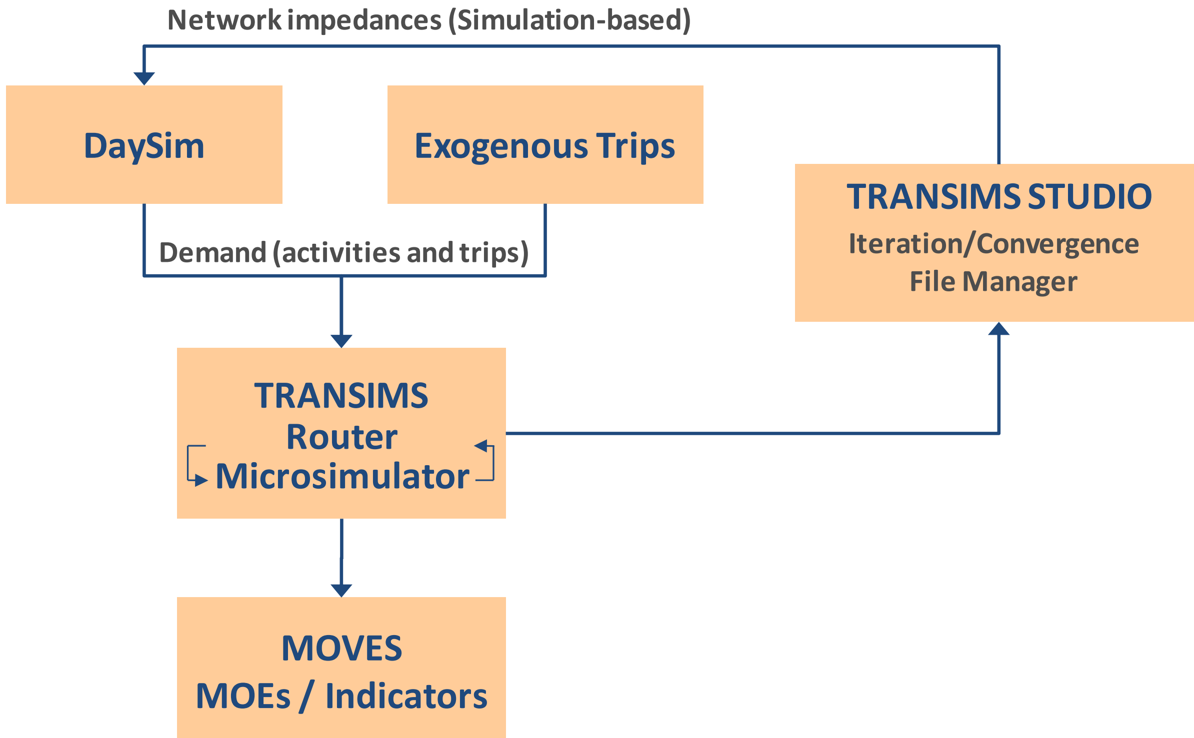
A key goal of the SHRP2-C10A project is that integrated demand-supply model is implemented in a dynamic modeling framework that is easily transferable to the local jurisdictions for policy analysis. In support of this goal of transferability, the model system incorporates a system manager that controls the execution of the three primary model system components: DaySim, TRANSIMS, and MOVES.

This model system manager, TRANSIMS Studio, is a Python programming language based Integrated Development Environment (IDE) built specifically to run TRANSIMS Version 4 applications. It has two basic components: (1) a Python-based library called Run Time Environment (RTE) and (2) a full featured Python GUI. The RTE is at the core of TRANSIMS Studio and is responsible for executing a series of TRANSIMS programs and external programs such as DaySim and MOVES in an iterative modeling framework. The GUI is fully-featured allowing users to manage and view input and output files; develop program controls and processing scripts, and track model execution status.

The TRANSIMS Studio model manager is configured to:

- Run on Windows or Linux,
- Run on stand-alone or clustered computers (e.g., TRACC),
- Run Jacksonville or Burlington models,
- Run Tour-based (DaySim) or Trip-based models (converted static demand), and
- Start the assignment process from free-flow speeds ("cold" start) or from the loaded speeds of a previous assignment ("warm" start)

Figure 55. Integrated Model System Components



2.8.1 Studio Components

The integrated model implemented in Studio is comprised includes four software components:

1. The TRANSIMS Studio user interface and application management software,
2. The TRANSIMS and DaySim modeling software,
3. Python scripts that define the modeling process, and
4. A folder structure housing network and other input data

The modeling scripts and input datasets for Jacksonville, Florida and Burlington, Vermont are a deliverable of this project and distributed through SHRP2. All of the software components are available free of charge from the following open-source websites:

The TRANSIMS Studio software can be obtained from:

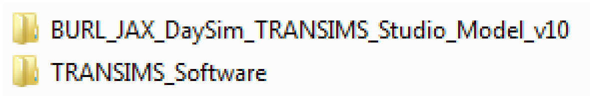
http://sourceforge.net/projects/transimsstudio/

The latest TRANSIMS software can be obtained from:

http://sourceforge.net/projects/transims/

The modeling package consists of two primary folders: (1) a folder containing the TRANSIMS and DaySim software and (2) a folder containing all the model data. Figure 56 shows the folder names created for this project. The model uses the concept of relative-paths which makes it easy to move the modeling package folders without having to set the full system paths in every control file. File paths also use the Linux directory convention to enable the programs to run on both Windows and Linux operating systems without modification. This enables the user to easily setup the model and also easily re-package or move the model folders.

Figure 56. Model package folders



The model data folder is further subdivided into a folder named “RTE” and a folder containing the actual model data. Figure 57 and Figure 58 show the condensed and expanded subfolders inside the model folder.

Figure 57. Model folder structure

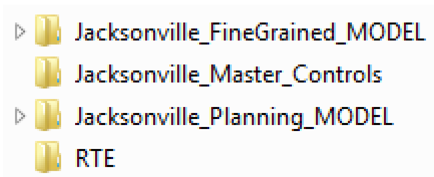
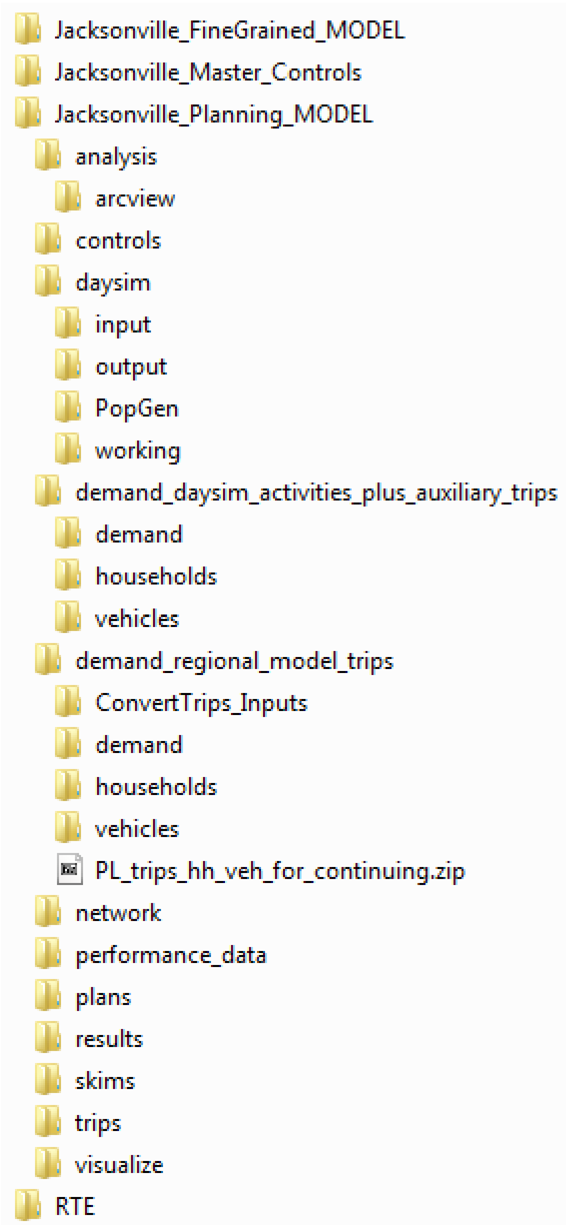


Figure 58. Expanded Model folder structure



All of the Python scripts used in the model are placed under the "RTE" folder. The following is a brief description of the purpose and functionality of each script:

1. **SysDef.py**
   This script is the central location where configuration keys that impact the entire model are defined. Therefore changes to this script are usually only performed when beginning a model run. This script inherits relevant standard Python libraries and is inherited by each of the other scripts.

2. **Main.py**

This script defines the various procedures and software steps used to perform global and assignment iterations. This is the only script that is actually executed as part of a model run. This script inherits all model definitions and global variables from Sysdef.py and functions from other scripts.

3. **AssembleSkims.py**
   This script defines the process for creating free-flow or loaded speed zone-to-zone skims for twenty-two time periods as input to DaySim.

4. **DaySim.py**
   This script runs DaySim and prepares the DaySim outputs for input to TRANSIMS.

5. **MsimIterations.py**
   This script defines the various steps involved in a TRANSIMS assignment iteration.

6. **MiscUtilities.py**
   As the name indicates, this script contains several miscellaneous procedures. These include the startup procedures and various data processing functions.

7. **VisualizerPrep.py**
   This script prepares the Microsimulator outputs for visualization using TRANSIMS-VIS.

In addition to the Python scripts which are identified by the ".py" extension, a number of other file types are created during the course of model execution. These include:

1. ".pyc" or compiled Python script,
2. ".gui" for display in the navigator pane of TRANSIMS Studio,
3. ".log" for display in the execution log window of TRANSIMS Studio,
4. ".pfm" for TRANSIMS Studio management,
5. ".pid" recording the process id,
6. ".job" containing model execution commands, and
7. ".res" which contains the information for resuming a model run.

A batch file is provided in the same folder to help remove temporary files created during model execution. However, this batch file is intended to be run only when a model needs to be started afresh because it clears all model logs and tracking information.

The TRANSIMS Studio settings are saved in a file with a ".prj" extension which is also saved in the "RTE" folder. This file is also called the TRANSIMS Studio "project" file.

2.8.2 Application Options

The model scripts were designed to store key model parameters and application options in a central location – SysDef.py – and reference these parameters across all other scripts using the high level processing steps defined within a single application script – Main.py – that controls the model execution and data flow. In addition, the procedures included in all Python scripts except Main.py are encapsulated within one or more functions to facilitate the "resume" feature of TRANSIMS Studio. Functions that perform small limited tasks help break down the complex model flow into simpler sub-steps. Thereby making it easier to track and resume from a sub-step following an abnormal execution termination.

A typical application of the modeling process involves decompressing the model package at a certain location on the local or network drive, opening the TRANSIMS Studio software, and loading the TRANSIMS Studio project file by navigating to the "RTE" folder inside the model package. After this the user opens SysDef.py in the navigator pane to confirm or edit the folder paths and model variables. Next, the user opens Main.py from the navigator pane to check the number of global and assignment iterations. After everything is set, the "play" button for the Main.py script is pressed to launch a model run.

Information is shared between the scripts by means of variables that are declared within the scope of RTE. These variables are prefixed with the term 'var.' and are globally accessible across all scripts that import the script where these are defined. For instance, the link delay resolution is defined inside SysDef.py script as "var.LINK_DELAY_RESOLUTION" and is used within MsimIterations.py for generating controls for TRANSIMS programs. This variable is coded inside a master control file by replacing the prefix "var." with "@" and adding a suffix "@" to the variable name. For instance, "var.LINK_DELAY_RESOLUTION" is referenced in a master control file as "@LINK_DELAY_RESOLUTION@". This functionality helps dynamically adjust control key values, if required, during TRANSIMS applications.

For the assignment iterations, the user can choose between the default simulation-based iterations or optionally skip simulation and perform only Router-based iterations by relying on volume-delay functions (VDF) in lieu of simulation-based delays.

The following two model parameters are used to switch between Router-based iterations or Microsimulator iterations:

2.8.2.1 For Router-Based Iterations (Planning Mode)

var.MODEL_ASSIGNMENT_OPTION = 'No_Simulation'
var.SKIMS_LINK_DELAY_SOURCE = 'Router_Based'

2.8.2.2 For Microsimulator-Based Iterations (Planning+Operations Mode)

var.MODEL_ASSIGNMENT_OPTION = 'Default'
var.SKIMS_LINK_DELAY_SOURCE = 'Msim_Based'

Other model parameters defined in SysDef.py that impact the TRANSIMS supply side model are shown below along with their default values and a description of their impact on the modeling process:

var.NUM_PARTITIONS = 8
This variable defines the number of partitions (.t*) to be used in the model. Partitioning helps utilizing multiple-cores across one or more machines depending upon the resources dedicated for the job. Partitioning primarily helps in the application of the Router and PlanPrep programs by enabling tasks to be performed in parallel. When running the model on a single machine, this value is set equal to the number of cores

available on a machine. A default value of 8 is provided corresponding to a modern workstation with 8 cores.

var.NUM_PATHSKIM_THREADS = 8
This variable applies only to the TRANSIMS Version 5 software PathSkim, which is used to create skims. It specifies the number of threads to use in its application. A default value of 8 is provided corresponding to a modern workstation with 8 cores.

var.MAXIMUM_PERCENT_SELECTED = 10
This parameter applies to the PlanCompare program and defines the maximum percent of regional trips to be changed per iteration. Large changes during iterations tend to have a destabilizing effect on the model convergence, hence an upper limit of ten percent is provided as a default.

var.WEIGHTING_FACTOR = 1
This parameter applies to the LinkDelay program and defines the value for the PREVIOUS_WEIGHTING_FACTOR key in that program. It defines the weight for the previous link delay during the link delay averaging process. A value of 1 implies equal weights or simple averaging and value of 2 implies a weight of 2/3 for the previous link delay and 1/3 for the current link delay.

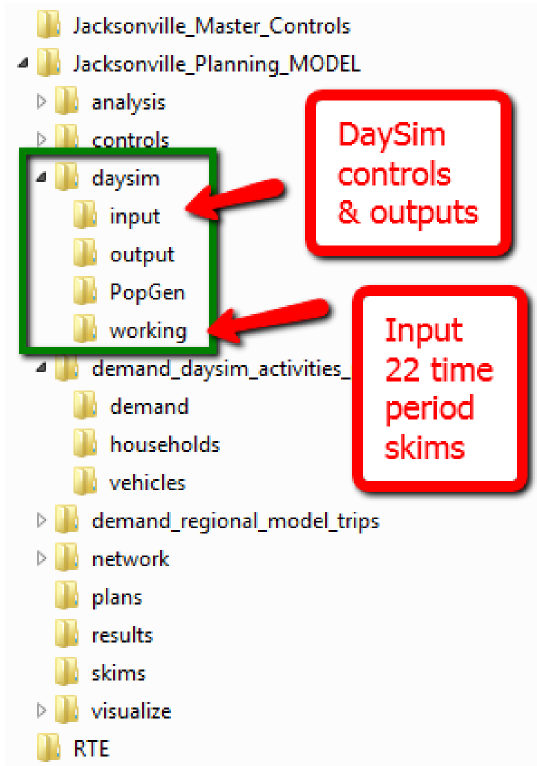
var.NUM_DELETE_PREVIOUS_RUN = 3
TRANSIMS assignments produce several gigabytes of data per iteration. A simulation based Jacksonville model produces in excess of 7 to 8 gigabytes per iteration. Since the intermediate iterations are not normally saved, this variable allows users to retain only the last few iterations at any time to conserve hard disk space. When set to '3', the three most recent iterations are preserved and the model will delete the fourth most recent iteration at the end of each assignment iteration. This deletion feature can be turned off so that all intermediate iterations can be preserved by setting this variable to a number greater than or equal to the number of expected assignment iterations.

2.8.3 DaySim Demand Component

All the inputs, controls, intermediate files and outputs of DaySim reside under the "daysim" folder of the model as shown in Figure 59. The input skims and output activities are processed and copied into and out of this folder for interfacing with TRANSIMS.

As discussed earlier, the calls to the DaySim process are placed from within the Main.py script, whereas the actual procedures for running DaySim are defined inside the DaySim.py script. Initial skims for the first DaySim run are prepared using a call to the AssembleSkims.py script. The user can choose to create these initial skims based on free-flow speeds or a link delay file. The details of the skim generation process are discussed later in the Time Period Skims section.

Figure 59. DaySim Folders



The DaySim demand creation process starts by running the DaySim executable using the initial skims in an iterative loop to prepare shadow-prices and subsequently executes the entire DaySim model system.

DaySim produces an activity and vehicle file in TRANSIMS Version 4 format. For the Jacksonville region, this process takes approximately an hour of computer processing time. The resulting activity file includes the internal travel demand generated by regional households. This demand is combined with the auxiliary trips to represent the complete travel demand for the region. To this end, the script combines the internal and auxiliary vehicle files and places the output in the "vehicles" subfolder of the folder "demand_daysim_activities_plus_auxiliary_trips". The activity file is copied to its "demand" subfolder. The DaySim activity file does not need to be merged with the auxiliary trip file because TRANSIMS programs, especially the Router, are able to read both the activity and trip file and process them in the same application.

Copies of DaySim activity and vehicle files are preserved from each global iteration for convergence analysis purposes.

2.8.4 TRANSIMS Supply Component

The TRANSIMS supply side model assigns the DaySim internal demand and auxiliary trips on the TRANSIMS network through assignment iterations designed to achieve dynamic user equilibrium convergence of the individual travel paths. The resulting network performance by time of day is used to generate zone-to-zone travel time, distance, and cost skims for twenty-two time periods for input into the next global iteration.

TRANSIMS models demand as trips between an origin and destination activity location (i.e., link offsets) at a specific time of day (i.e., seconds). The Router builds a minimum impedance path between the origin and destination based on time dependent link travel times and turning movement delays. The paths or travel plans for all of the trips over a 24 hour period are loaded onto the network and simulated by the Microsimulator. The Microsimulator considers traffic signal timing, lane changes, and vehicle interactions in estimating the volume and travel time on the network at any point in time. These data are aggregated by link and time period for feedback to the Router for path adjustments. The process continues until most travelers cannot improve their travel time by changing paths.

The temporal resolution of the link delays has traditionally been 15 minutes, but this model typically uses 5 minute link delays and has been tested using 2 minute link delays and interpolated link delays. The appropriate time increment for link delay summaries is significantly dependent on the simulation methodology. The TRANSIMS Microsimulator uses a cellular automata method that moves vehicles between link-lane cells on a second-by-second basis. A six meter cell size was used for this model. This has the effect of limiting the instantaneous speed of a vehicle at any second within the simulation to one of seven values (i.e., 0, 6, 12, 18, 24, 30, and 36 meters/second). This means that a relatively large number of vehicle-second observations are required for a given link to generate a reasonable average speed. Average speeds generated using 5 minutes worth of data appear to generate the best results.

2.8.5 Startup

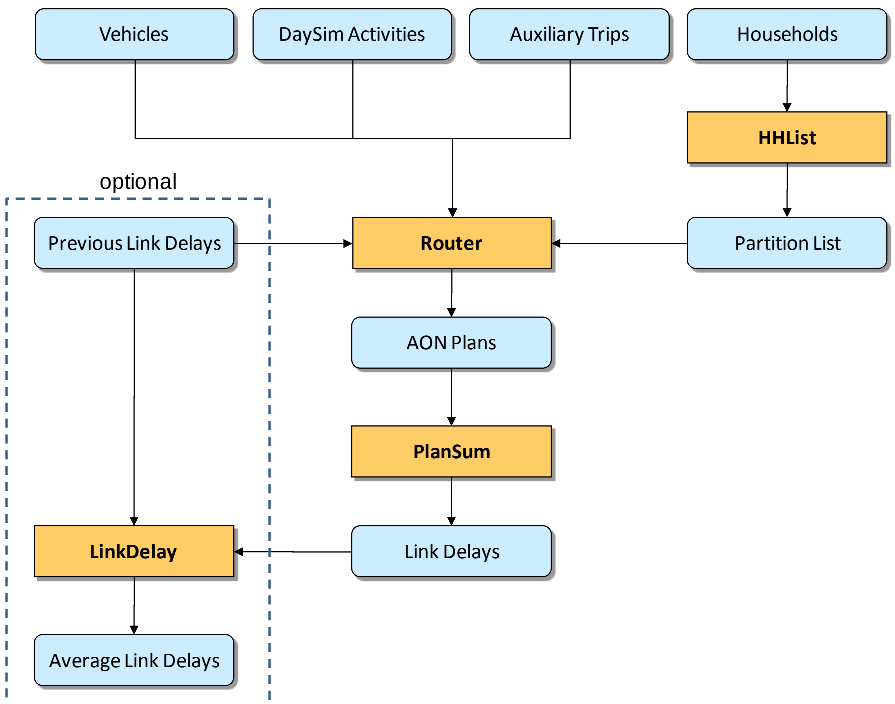
To get the assignment process started, a startup script is used to separate the steps that are either applied only once during the model execution or which require special considerations due to the lack of a previous iteration. The startup procedures are implemented as functions within the MiscUtilities.py script. This step or function is executed at the beginning of every global iteration before commencing assignment iterations and is run only once.

The primary steps executed at this stage are:

1. Partitioning the regional households,
2. Building all-or-nothing (AON) paths for each trip based on free flow speeds or loaded speeds from a previous model application,
3. Creating link delays by time of day based on volume-delay functions, and
4. If appropriate, averaging the link delays with the results of the previous model run.

This process is depicted in Figure 60.

Figure 60. TRANSIMS Assignment Startup



One of the most significant steps at this stage is the partitioning of households which impacts the rest of the model execution and the overall processing time. Typically, a higher number of partitions implies lower processing times for partitioned programs such as Router and PlanPrep. However, the number of partitions can be specified independent of the number of machines/threads or nodes available at the user's disposal. When the number of partitions is specified higher than the number of machines/threads, TRANSIMS Studio processes the partitions as sequential sets of applications based on the number of machines/threads available. For example, if 18 partitions (".tAA" thru ".tAR") are processed on a machine with 8 cores, the first set of 8 partitions (".tAA" thru ".tAH") are processed first, followed by the second set of 8 partitions (".tAI" thru ".tAP") and finally, the remaining two partitions (".tAQ" thru ".tAR") are processed using only two threads while the rest of the six threads remain idle. Therefore, for maximum efficiency, the number of partitions should be set equal to or an integer multiple of the number of machines/threads available.

In the tour-based model, the HHList program reads the household or traveler lists generated by DaySim and the auxiliary trip model to compile a master list of all travelers in the region before randomly distributing the travelers into a specific partition file. All trips made by a household are processed within the same partition. In the trip-based model, all regional trips are stored in a single file and each trip is assigned a unique household number. The trips are partitioned based on the household trip number.

2.8.6 Equilibrium Convergence

Convergence is necessary in order to ensure the behavioral integrity of the model system. The impedances or level-of-service measurements used as the basis for accessibility measures and as key inputs to the destination and mode choice models must be approximately equal to the travel times and costs produced by the final network assignment process. Model system convergence is also necessary to ensure that the model system will be useful as an analysis tool. The stability of model outputs is essential to support planning and engineering analyses, and changes to demand or supply should lead to reasonable changes in model outputs.

In the context of an integrated demand and network simulation model system, an essential precondition for pursuing overall model system is establishing network assignment convergence. Network convergence is analogous to model system convergence – the inputs to the network assignment process (the current traveler paths that give rise to the current network costs) must be approximately equal to a set of new best paths based on these current network costs.

A key focus of the C10 effort has been to identify and test different strategies for achieving both network assignment convergence as well as overall model system convergence within the context of the Daysim-TRANSIMS integrated model. This section describes the user equilibrium convergence or "gap" calculation procedures employed in the TRANSIMS supply side model. These procedures are run at the end of every iteration and do not influence the model assignment procedures. Information about the derivation and formulas involved in the gap calculation can be found in the appendix to this report.

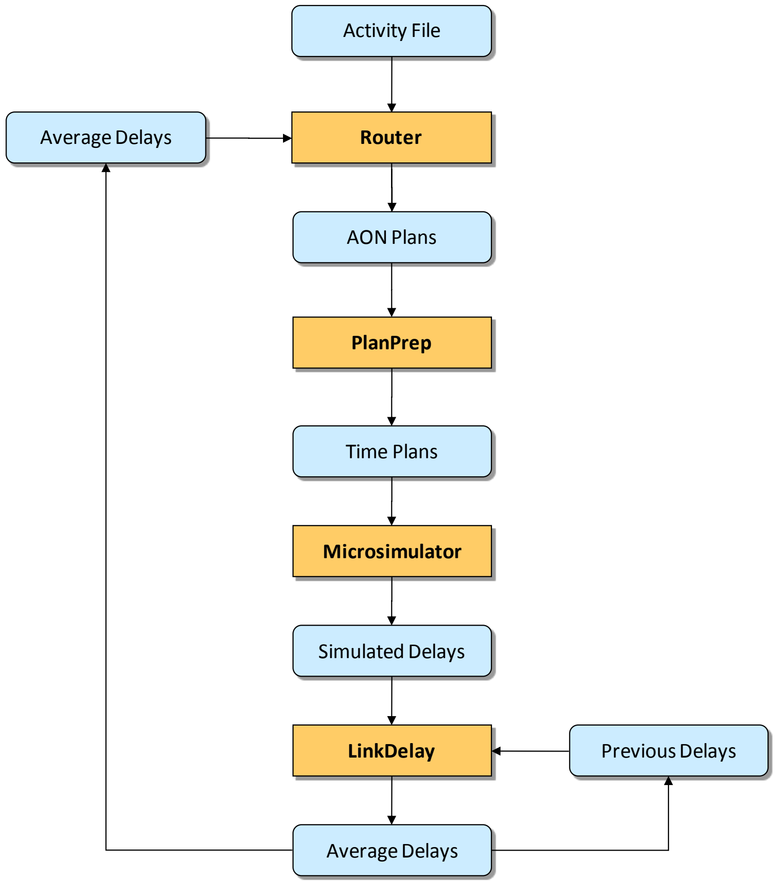
Two gap measures have been defined and employed in this model: "relative-gap" or link-based gap and "trip-gap" or traveler-based gap measures. The relative-gap is equivalent to the widely used network link-based gap measure in conventional deterministic travel demand models. The trip-gap is a newer measure enabled by the detailed information about individual travelers available within a disaggregate model like TRANSIMS.

The trip-gap concept is further divided into three estimation methods identified as re-skimmed, event-based and hybrid. The event-based and hybrid trip-gap methods require the Event file from the Microsimulator and therefore are not applicable in the Router-only assignment methods. For Router-only iterations, the model automatically switches the trip-gap measure to the re-skimmed method.

2.8.7 Microsimulation-Based Equilibrium (Planning+Operations Mode)

In coordination with a parallel Federal Highway Administration-funded TRANSIMS deployment effort researching convergence and other issues associated with advanced integrated travel demand model systems, the project team has extensively investigated and tested methods for achieving dynamic user equilibrium in the context of the TRANSIMS Router and Microsimulator. As part of this parallel effort, a peer review panel organized by the Federal Highway Administration reviewed the C10A integrated model process and identified refinements that ensured the methods implemented were consistent with current practice. A key strategy related to convergence that was deemed acceptable by the peer exchange panel was to average Microsimulator link delays in order to dampen oscillation effects, but simulate the full set of all-or-nothing paths during each iteration. This application approach is depicted in Figure 61.

Figure 61. Microsimulator-Based Equilibrium Process



This method involves generating at each iteration a new set of plans for each traveler based on average simulated delays and simulating these plans. A key concern about this process is the ability of the Microsimulator to realistically simulate all-or-nothing paths. However, empirical tests of this method in both Jacksonville and Burlington have confirmed that the Microsimulator can simulate all-or-nothing paths without creating significant congestion problems. The fact that each traveler has a unique all-or-nothing path between activity locations, starting at a specific time of day sufficiently distributes the paths to avoid the types of all-or-nothing assignment problems typically experienced by traditional modeling frameworks. These results appear to hold even with large increases in demand.

2.8.8 Router-Based Equilibrium (Planning Mode)

Microsimulator-based iterations are the preferred method of performing a user equilibrium assignment using TRANSIMS, but they are not the only way. The Router-only iterative process used in the model system "Planning Mode" is similar to the Microsimulator-based iterative process, but uses traditional volume-delay functions for computing link delays rather than the second-by-second simulation of individual vehicles. This process does not consider traffic signal operations, lane changing or vehicle interactions. It simply takes user-provided link capacity and estimated volumes to calculate the link travel time for each time increment. In this case, volume is estimated by tracing the location of each vehicle at any given time using the trip path and start time stored in the travel plan file.

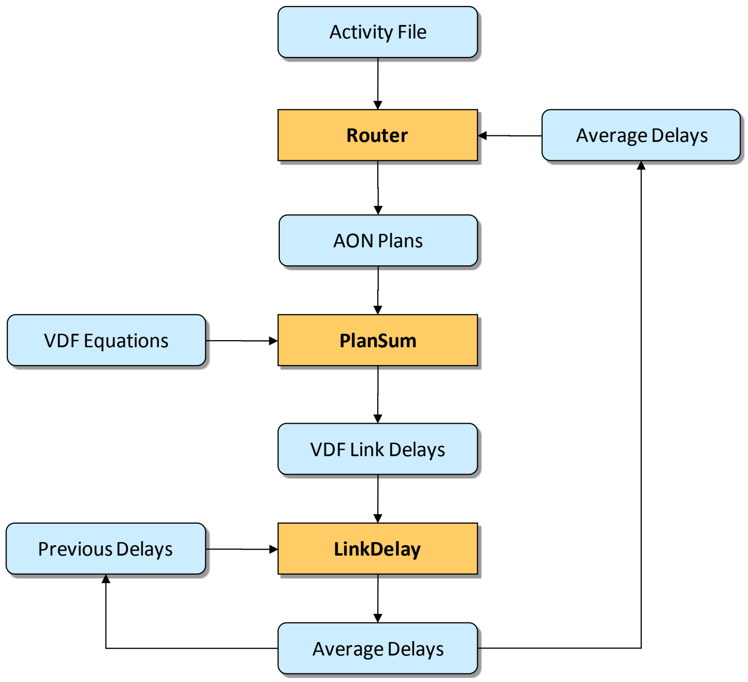
The primary reason for using Router-only iterations for all or some of the assignment process is the fact that the Microsimulator is computationally complex and therefore the most time consuming step in each iteration. The Microsimulator performance is further complicated by the fact that it is a single-threaded program in TRANSIMS Version 4 and cannot be partitioned. As a result, the Router-only approach can perform an assignment iteration in approximately one tenth of the time required for a Microsimulator-based iteration. This makes the Router-only process attractive for initializing estimates of link travel time prior to performing Microsimulator-based convergence or as a substituted for simulation for applications that can tolerate less rigorous analysis or are not focused on traffic operations.

As with the Microsimulator-based process, the Router-only iterations can be implemented within TRANSIMS in two fundamental ways. The primary method used in this project is shown in Figure 62 is conceptually consistent with a traditional assignment process. An all-or-nothing path is built based on the previous link delays and the resulting volumes are converted to travel times using a volume-delay function. The link delays are averaged using a weighted average or MSA technique for input to the next set of all-or-nothing paths. In this case the averaging technique is critical to managing the stability of the assignment and the convergence process.

Alternative approaches to seeking convergence are also feasible, such as an incremental assignment approach that compares an all-or-nothing routing of each traveler to a re-skimmed version of the previous path and selects of subset of trips with large travel time and impedance differences for inclusion in the composite plan file. This plan file is then aggregated into link volumes by time of day, delays are calculated using a facility type-based volume-delay function to the 5 or 15 minute link volumes to estimate the travel time for each time increment, and weighted averaging or method of successive averaging (MSA) procedure is used to combine these travel time estimates with previous travel time estimates for feedback to the next Router iteration.

In either approach it is important to note that volume to capacity ratios can exceed 1.0 just like they do in traditional models. In fact this problem can be even more significant in TRANSIMS when fine grained time periods are used. Volume to capacity ratios for 5 or 15 minute time periods are significantly more likely to exceed 1.0 than volume to capacity ratios calculated using peak period or daily volumes. This typically means that extra care needs to be taken in designing the parameters used in the volume-delay functions. If the travel times become excessive, the Router will have difficulty completing trips, create scheduling problems for subsequent trips within tours, and move link volumes into much later time periods of the day.

Figure 62. Router-only User Equilibrium



2.8.9 Equilibrium Convergence Measures

2.8.9.1 Relative Gap

Link-based relative gap is a convergence statistic that quantifies the difference between the simulated performance of the traffic on each link by time of day and the vehicle hours of travel that would result from each traveler taking the minimum impedance path based on the simulated travel times. The mathematically formulation is shown in Figure 63.

Figure 63. Link-based Relative Gap

$$Relative\ Gap = \frac{|\sum VE_t \times CE_t - \sum VA_t \times CE_t|}{\sum VA_t \times CE_t}$$

where:

$$where:$$

$$\Sigma = summartion\ over\ all\ network\ links$$

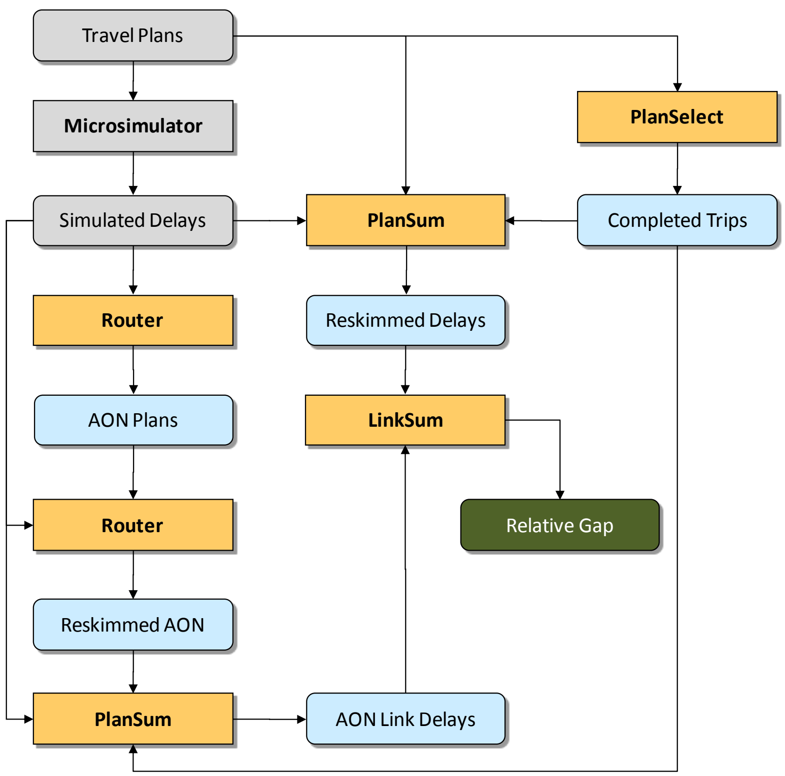
$$VE_t = the\ simulated\ volume\ on\ a\ given\ link\ and\ time\ increment$$

$$CE_t = the\ travel\ cost\ (time) associated\ with\ volume\ VE_t$$

$$VA_t = the\ link\ volume\ from\ an\ AON\ assignment\ based\ on\ CE_t$$

The data processing steps required to calculate the relative-gap convergence measure are outlined in Figure 64. The primary inputs to the process are travel plans from the current iteration and the Microsimulator performance or link delay file. These are shown in grey in Figure 64. The relative-gap process then builds all-or-nothing paths using the performance file. These paths are then re-skimmed using the same performance file. This re-skimming process updates the travel time and generalized-cost values for all travel plans in a consistent way. The all-or-nothing plans are processed by PlanSum to create an all-or-nothing link delay file. Similarly, PlanSum is used to create a link delay file using the travel plans and the Microsimulator performance file. This step, although seemingly unnecessary, is needed to create a link delay that has a complete and consistent set of volumes for all times of day. Furthermore, a common list of travelers is supplied to both PlanSum applications to limit the statistical comparison to trips that were successfully completed in both path building applications. The final two link delay files are compared using LinkSum to compute relative-gaps for every hour of the day.

Figure 64. Link-based Relative-Gap Calculation Process



2.8.9.2 Trip-Gap

TRANSIMS builds a unique path or travel plan for each trip based on the origin and destination activity locations, the trip start time, and other travel mode attributes (e.g., HOV or truck restrictions, etc.). In its simplest form, a trip-gap is computed by comparing the generalized costs for every traveler from their path in the travel plan file to their path in the all-or-nothing plan file. However, there are a number of variations and complexities in calculating this measure depending on whether the only the Router is used in assignment, or whether the Microsimulator is used, as well as arising from the nature of traffic microsimulation. Ultimately, three variations of the trip gap were developed and tested. Two of these gap measures use "experienced time" for each traveler derived directly from the Microsimulator outputs. The formula for these gaps is show in Figure 65. The third gap measure is based on "reskimmed time" and is shown in Figure 66.

The data processing steps required to calculate the three trip-gap convergence measures are outlined in Figure 67. The inputs to this process are the same as the relative-gap process. They include the travel plans and the Microsimulator performance or link delay file from the current iteration. These are shown in grey in Figure 67. All three measures are calculated in every iteration; however, only one of them is used in the model for measuring convergence. The event-based and hybrid trip-gap methods attempt to consider the actual simulation travel time for every traveler. The hybrid trip-gap method additionally attempts to correct for the influence of simulation and routing problems.

Figure 65. Trip Gap Using Experienced Costs

$$Trip\ Gap = \frac{|\sum(CE_x - CA_x\{C_{mt}\})|}{\sum CA_x\,\{C_{mt}\}}$$

$where$:

$\{C_{mt}\} = simulated\ time\ varying\ link\ costs$

$CA_x = AON\ cost\ of\ trip\ x\ based\ on\ link\ costs\ \{C_{mt}\}$

$CE_x = simulated\ cost\ of\ trip\ x\ that\ resulted\ in\ link\ costs\ \{C_{mt}\}$

The re-skimmed trip-gap is obtained by comparing the re-skimmed travel plans to the re-skimmed all-or-nothing plans. The event-based trip-gap is computed by comparing re-skimmed all-or-nothing plans and the updated travel plans incorporating the actual start and end times of every traveler in the simulation. In this comparison, problem travelers in the Microsimulator are excluded. The hybrid trip-gap goes one step further and includes the re-skimmed paths for travelers with simulation problems for comparison against the re-skimmed all-or-nothing plans.

Figure 66. Trip Gap Using Reskimmed Time

$$Trip\ Gap = \frac{|\sum(CR_x\{C_{mt}\} - CA_x\{C_{mt}\})|}{\sum CA_x\,\{C_{mt}\}}$$

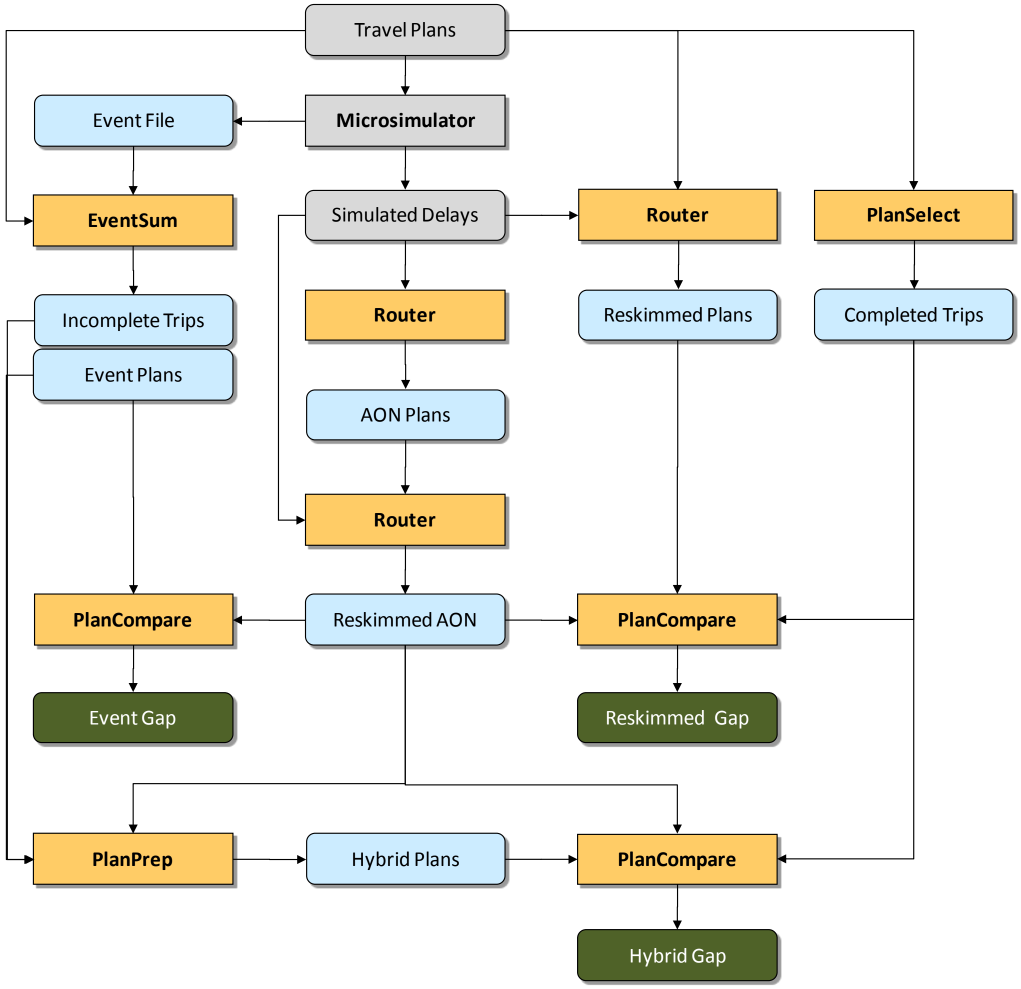
$where$:

$\{C_{mt}\} = simulated\ time\ varying\ link\ costs$

$CA_x = AON\ cost\ of\ trip\ x\ based\ on\ link\ costs\ \{C_{mt}\}$

$CR_x = reskimmed\ costs\ for\ trip\ x\ along\ the\ path\ used\ to\ generate\ \{C_{mt}\}$

Figure 67: Traveler-Based Trip-Gap(s) Calculation Process



2.8.10 DaySim-TRANSIMS Integration

As shown in earlier figures, DaySim provides trip and vehicle information to the TRANSIMS Router to perform network assignment. In its original implementation, DaySim produced person-trip records with trip-end locations defined as parcels and trip start and end times defined as specific minutes of the day. In order to integrated DaySim with TRANSIMS, a number of modifications were made to translate DaySim tour and trip records into vehicle-trip records, associate the parcel locations with the activity locations used by TRANSIMS and output the records in the format required by the Router. DaySim does not simulate vehicle type choice or allocate specific vehicles to person trips. Therefore, a simple procedure is used to treat every vehicle trip as an independent vehicle. These modifications are described below.

2.8.10.1 Activity File

The primary modification of DaySim required to integrate DaySim and TRANSIMS was the adding to DaySim the capability to create TRANSIMS input activity and vehicle files. These modifications were relatively straightforward given the comparability and the list-based nature of both DaySim's trip output and TRANSIMS activity file input.

2.8.10.2 Activity Files vs. Trip Files

Version 4 of TRANSIMS can accept two types of list-based inputs: activity files and trip files. A trip file is straightforward: it is a list of trips to be assigned to the network and includes information on the origin activity location, the destination activity location, and departure time of the trip and the mode of travel used. All of the auxiliary demand used in the integrated model, such as airport ground access trips and commercial vehicle trips are converted from SACSIM trip matrices into TRANSIMS trip files. When the TRANSIMS Router assigns a trip file to the network, each trip is considered a discrete movement, independent from any other trips in the file.

Activity files are more complex. An activity file is a list of activities undertaken by regional residents, and does not explicitly include trips. For each household, person and activity, the activity file includes a purpose, start time, end time, duration, mode, and location. Using this information, the TRANSIMS Router creates a plan for each movement required for an individual to reach their desired activity locations. These plans are essentially equivalent to trips. The critical distinction between using an activity file and a trip file is that when routing activities TRANSIMS treats each all the movements as interconnected – the "tour" structure is preserved within each person. As currently configured, the activity durations are fixed, so if it takes longer for a traveler to reach their activity location, the remainder of their trips and activities will be pushed back in time as well. This has distinct implications for the integration of the activity model with the TRANSIMS Router, and for the overall model system calibration and validation, an issue discussed in a subsequent section of this report. This approach was necessary in the initial model development in order to ensure that all trips were assigned and conserved. Future integrated model development efforts will consider how re-scheduling and time-pressures can be flexible accommodated in DaySim, TRANSIMS, or both.

Using a simple tour comprised of two trips, with a single trip on each tour leg, Table 47and Table 48 illustrate the differences between the original DaySim trip output file and the TRANSIMS activity input file. Both files indicate the household and person travelling. For each of the two given trips in Table 47, the DaySim trip output file contains information on the person tour number on which the trip occurs, the tour half (outbound or return) on which the trip occurs, and the trip number within the tour. Critical trip details are also included, such as the origin and destination parcel and TAZ numbers, the travel mode used to make the trip, the origin and destination purposes, the trip departure time and the trip arrival time.

Table 47. DaySim Trip List Output Example

SAMPN	PERSN	TOURNO	TOURHALF	TRIPNO	OTAZ	OCEL	DTAZ	DCEL	MODE
1	1	1	1	1	445	429711	1088	133524	7
1	1	1	2	1	1088	133524	445	429711	7

(continued)

OPURP	DPURP	DEPTIME	ARRTIME	EACTTIME	TRAVTIME	TRAVDIST	EXPFACT
8	4	1222	1238	1556	16.09	8.56	1.00
4	8	1556	1615	2659	18.65	8.56	1.00

Table 48. TRANSIMS Activity File Example

HHOLD	PERSON	ACTIVITY	PURPOSE	PRIOTITY	START	END	DURATION	MODE	VEHICLE	LOCATION	PASSENGER
1	1	111110	0	9	1	44520	44519	1	0	5937	0
1	1	11111	4	9	45480	57360	11880	2	1	13688	0

1	1	11121	0	9	58500	97140	38640	2	1	5937	0

Whereas the DaySim trip output contains two records representing two trips, the TRANSIMS activity file record contains three records representing three activities. The first activity represents the person's "at home" activity, where they start their day. TRANSIMS derives one trip to take the person from their home to their first activity location (shown in the second activity file record), and then derives a second trip to take the person to their next activity location, which is back home, as indicated by the common location id in the first and third activity records. Adding this initial "at home" activity was one of the key changes made to the DaySim output.

2.8.10.3 Temporal Units

Table 48 also shows that this first "at home" activity ends at time 44520. Another key change made to the DaySim output involved the conversion of the time units from hours and minutes (for example, 1222 represents 12:22 in the original DaySim trip list output), to seconds. The first activity is shown to end at 44520, and when translated from seconds to hours and minutes, this time is also 12:22. So the end time for each activity in the activity file is the same as the start time for the trip that takes the traveler to their next activity, consistent with the DaySim trip file. Note that the start time for the second activity, which is "out of home" as indicated by the new location identifier, is 45480. Subtracting 44520 from 45480 results in 960 seconds, which is consistent with the 16 minutes shown as the travel time for the first trip in the DaySim trip list output.

2.8.10.4 Spatial Resolution

Table 47 and Table 48 also illustrate the differences in the geographic resolution between the DaySim trip list output and the TRANSIMS. As previously described, DaySim uses detailed parcels as the fundamental spatial unit, but in the original DaySim implementation this parcel-level detail was aggregated to a TAZ-level prior to network assignment using traditional static equilibrium assignment methods. The DaySim trip list output contains both the origin and destination parcel and TAZ information. When integrated with TRANSIMS for network assignment, DaySim uses "activity locations." Activity locations are more fine-grained spatially than TAZs (there are approximately 25,000 activity locations in the Jacksonville region), but not as detailed as individual parcels (there are approximately 620,000 parcels in the Jacksonville region). A correspondence file between parcels and activity locations was developed in order to translate parcel information to activity locations prior to assignment in TRANSIMS. This spatial disaggregation in assignment is one of the distinguishing aspects of the integrated model.

2.8.10.5 Mode

Two changes in the configuration of DaySim to produce a TRANSIMS activity file involved the treatment of mode. The simpler of the two changes involved recoding of the travel modes used in DaySim into the pre-established mode codes used in TRANSIMS. For example, Table 47 showing the original DaySim trip list output shows the first trip using mode 7, which is "drive alone" in DaySim. Table 48 showing the new activity file shows that the second record contains mode 2, which is the TRANSIMS "drive alone" mode used to make the travelers first trip to their first out of home activity location.

This mode logic is significantly more involved for shared ride trips. In existing activity-based model implementations that have used static network assignment procedures, shared ride trips are simply aggregated to the zonal level, and divided by an assumed occupancy rate in order to calculate vehicle trips. This approach does not work in a disaggregate assignment simulation such as TRANSIMS because the goal is to preserve the details about each individual trip. It is not appropriate or logical to divide discrete shared rides trip by an occupancy rate in order to estimate vehicle trips. Instead, it is necessary to assign driver and passenger status to travelers whose mode is identified as shared ride.

Using TRANSIMS, we only want to assign auto driver tours. DaySim predicts the occupancy for auto trips – Drive Alone (DA), Shared Ride 2 (SR2) or Shared Ride 3+ (SR3), but it does not predict whether the person is the auto driver or the passenger, and it does not coordinate the driver and passengers within a household. In addition, there may be different trip modes (vehicle occupancies) for different trips within an auto tour. A detailed analysis was used to derive the most realistic and unbiased method for assigning an auto driver or passenger designation to each auto tour and trip, so that we can select which tours to send to the Router.

In order to determine which car trips are part of car driver tours, a set of rules was established to deal with mixed tours that include some car trips and some non-car trips. Car trips that are part of school bus or transit tours are typically car passenger trips where the person gets a ride in one tour direction and takes a bus in the other direction. For simplicity, it was assumed that these trips are all passenger trips and do not need to be routed in TRANSIMS. In addition, there are mixed mode auto tours that include one or more walk or bike trips. Because these are difficult to handle in TRANSIMS, and the number of them is quite small, it was assumed that those the auto trips in those tours are passenger trips. Using assumed occupancy values of 1.0, 2.0, and 3.63 for the three auto modes and the total number of trips in each tour mode, the expected number of car driver trips was calculated. A method was established for determining which tours of each occupancy level to assign as driver tours based on other trip modes used on the tour. For example, if a tour includes one or more walk or bike trips as well as shared rider trips, it is designated as a car passenger tour. In contrast, if a tour includes no walk or bike trips, but does include one or more drive alone trips, it is designated as a car driver tour. Finally, if a tour includes only shared ride tips, a certain proportion of these tours are randomly designated as car driver tours and the rest are designated as car passenger tours based on proportions derived from survey and modeled data.

One final note on mode coding is that in the TRANSIMS activity file, all of the activities that are accessed using the drive alone and shared ride-driver modes are identified as MODE=2. TRANSIMS then uses information in the PASSENGER field to determine if the trip is truly drive alone or shared ride. If PASSENGER=0, the trip is treated as a drive alone trip and assigned to the network. However, if PASSENGER>0, the trip is treated as a shared ride driver trip, and is assigned to the HOV network.

2.8.11 Vehicle File

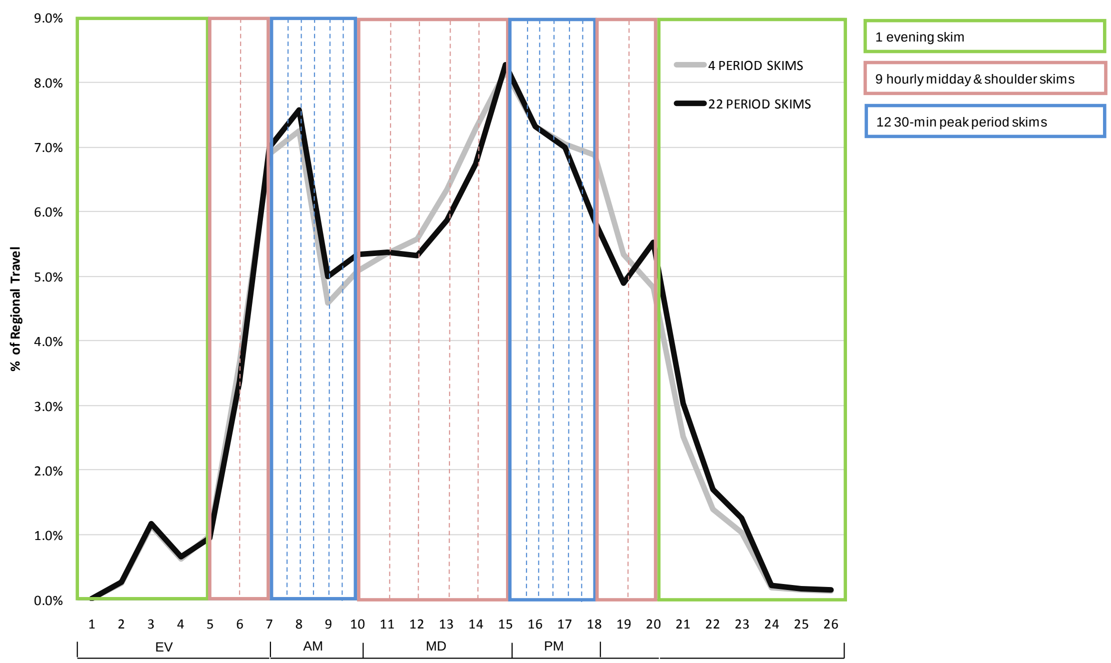
TRANSIMS has the ability to allocate or assign vehicles to individual travelers, and track these vehicles throughout the day. However, DaySim does not allocate vehicles to individual travelers. Thus, when creating the activity file, a separate vehicle is created for each auto driver tour, unconstrained by the number of vehicles each household is predicted to own or by competition amongst household members for the household vehicles. The project team anticipates enhancing DaySim so that is will assign household vehicles to each auto driver tour as part of other research efforts. This would enhance the value of the integrated model by enabling it to more realistically model vehicle usage and resulting air quality impacts in the region.

2.8.12 TRANSIMS-DaySim Integration

Network skims, or location-to-location measures of network impedances and costs, are an essential element of any travel demand forecast system. These skims are used directly or in directly in virtually every component of the DaySim model system, from calculating accessibility measures that influence long-term choices such as auto ownership and overall daily tour and trip generation, to providing direct input into short-term choices of destination, mode, and time-of-day.

These skims are generated using network assignment or simulation software based on network performance by time-of-day and are defined along a number of key dimensions such as the spatial, temporal, and modal resolution, and may be provided in a number of different file formats. In the initial implementation of the Jacksonville DaySim-TRANSIMS model system, the primary spatial unit used for skimming is the travel analysis zone (TAZ) and the temporal unit used is the detailed time period. In the current implementation, a total of 22 time period skims are generated and used in the model system, and these time periods vary in length from ½ hour during the 3-hour AM and PM peak periods to 1 hour during the midday, early morning, and early evening, to a single broad overnight time period. The DaySim-TRANSIMS model system can be configured to other levels of temporal resolution as well.

Figure 68. 22 Time Period Skim Definition



One of the significant enhancements to TRANSIMS' capabilities is the new PathSkim program, which is used to build paths and gather travel attributes between selected locations at specific times of day. In addition to significant performance improvements through multi-threading and one-to-many path building techniques, PathSkim makes selecting origins and destinations for zone-to-zone skims by time of day significantly more convenient. It automatically selects one or more activity locations near zone centroids as path origins and destinations. The locations can also be randomly or geographically distributed within the zone or provided by the modeler through a zone-location file. The zone-to-zone or district-to-district skim information is aggregated in memory and written directly as a single skim matrix or a series of matrix files for different time periods.

In PathSkim the output time periods can also vary in length or combine travel from different times of day. This is particularly important for the SHRP2-C10 model since DaySim requires 30 minute skims during peak periods, hourly skims for off-peak periods, and a night time skim that combines late evening hours with early morning hours. DaySim also uses skims to set the trip departure time based on a scheduled arrival time at the activity location. The Version 4 Router only builds paths from an origin at a specific time of day to a destination. What DaySim would prefer and what PathSkim provides is the ability to generate paths and aggregate skims based on specified arrival times at the destination. In other words, paths are built backwards in time. For the Jacksonville network, the Version 4 skimming process took over 24 hours to run and generated a "temporary" plan file close to 100 gigabytes in size. Performing the same task with a multi-threaded version of PathSkim takes approximately 15 minutes and produces no temporary files.

Although the initial C10A system architecture envisioned that the TRANSIMS tools would be used to generate and return to DaySim activity location-level measures of network impedances for a specified set of origin-destination pairs (ODs) and time period given, at present the model system is still employing TAZ-level network impedances. Ultimately, the time and cost measures may be based on more spatially detailed TRANSIMS "activity locations" and for specific times that a trip or activity may be routed.

As described earlier, the fundamental spatial unit used in the DaySim is the individual parcel, which is significantly more fine-grained than the TAZs used for network skimming. This is primarily driven by the fact that creating, storing, and accessing more spatially detailed skim data can be computationally burdensome. For example, a region such as Jacksonville with approximately 1500 TAZs necessitates the development of separate skims that for each modal attribute and time period contain 2,250,000 individual values. Were these skims to be developed at the TRANSIMS "activity location" level used in network assignment, it would be necessary to store over 400,000,000 individual values for each modal attribute and time period. However, in order to refine the TAZ level skims, DaySim incorporates some parcel level information such as the distance from each parcel to the nearest transit stop by transit submode. In the Jacksonville DaySim-TRANSIMS model, TRANSIMS creates fixed format ASCII skim files. However, DaySim can be enhanced to read and write other data file formats, such as native Cube matrix format and binary files.

2.8.13TRANSIMS-MOVES Integration

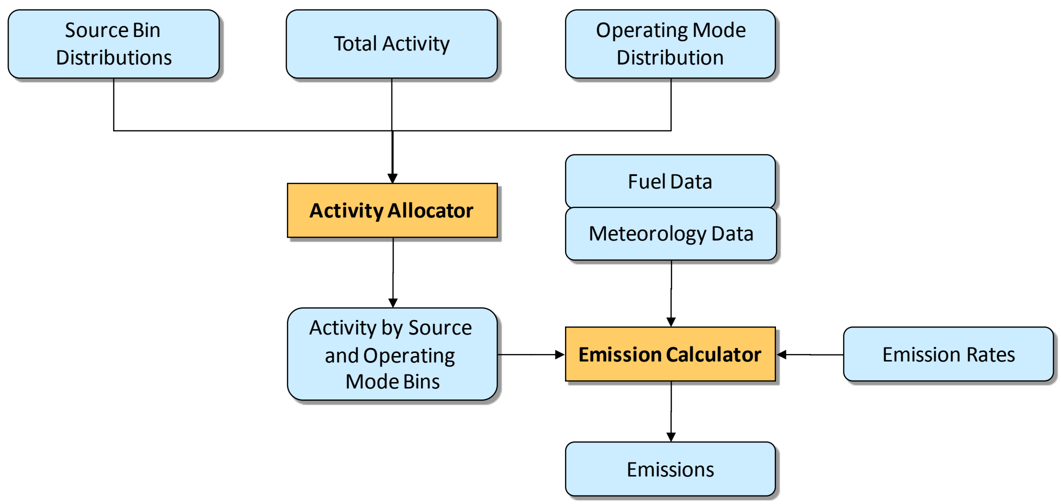
One of the objectives of this study is to estimate the air quality impacts of each of the application alternatives using the Environmental Protection Agency's (EPA) new Motor Vehicle Emission Simulator (MOVES). MOVES replaces MOBILE6 and NONROAD as the mobile source emission tool required for air quality conformity analysis and emission impact analysis. It is designed to produce county-level emission inventories for the entire nation; zone and link level emissions for State Implementation Plans and regional Conformity analysis; and micro-scale emission rates for hot-spot and project level analysis. The goal is to move away from average operating characteristics over broad geographic areas to finer analysis scales based on detailed operating characteristics of a wide variety of vehicle types at specific locations and times of day.

A brief description of MOVES in the integrated model system is provided in this section. More detailed information on MOVES implementation can be found in the Appendix to this report.

2.8.13.1 MOVES Architecture

At its lowest level, the MOVES software applies emission rates to activity data by source and operating mode bins. For transportation applications, source use types are basically equivalent to vehicle types. Figure 69 provides an overview of the core model processing. The activity data is allocated to source and operating mode bins by the Total Activity Allocator (TAA). This module processes data from three primary data sources: Source Bin Distributions, Total Activity, and Operating Mode Distributions. The activity data is passed into the emission calculator along with emissions rates that are adjusted using local and seasonal fuel and meteorology data.

Figure 69. MOVES Core Model



From the MOVES core model perspective:

- Fuel data are the supply of fuel by time, location, and type.
- Meteorology data are the temperature and humidity by time and location.
- Total activity is the quantity of emission generating activity by time, location, and source use type.
- Operating mode distributions distribute average operating characteristics such as average speed to vehicle specific power (VSP) bins by time, location, and roadway type.
- Source bin distributions convert vehicle data to emission specific classifications. These distributions can vary by source use type, but not by time or location.

2.8.13.2 Total Activity Data Generator

To support the core model and facilitate a wide variety of user inputs, the MOVES architecture includes a number of data generators to manipulate and format the required information. The Total Activity Generator (TAG) is designed to produce the activity for the source use types by time and location. This includes the vehicles and the operation of these vehicles. The MOVES process prepares the Total Activity through a series of steps that perform calculations at increasing levels of detail and specificity. These steps are controlled through the data input interface. Table 49 provides a high level overview of the calculations performed by the Total Activity Generator.

Table 49. Total Activity Generator Steps

Step	Calculation
TAG-0	Determine the base year
TAG-1	Calculate base year vehicle population
TAG-2	Grow vehicle population to analysis year
TAG-3	Calculate analysis year travel fractions
TAG-4	Calculate analysis year VMT
TAG-5	Allocate analysis year VMT by roadway type, use type and age
TAG-6	Allocate annual VMT to hour by roadway type, use type and age
TAG-7	Convert to total activity basis by process
TAG-8	Allocate total activity basis by zone location
TAG-9	Calculate distance traveled

The first few steps in the Total Activity Generator (TAG) are designed to define the array of vehicle types that are included in the analysis domain or region. MOVES defines vehicles as a source use type and tends to define vehicle types at a much higher level of specificity than a transportation planner. From a MOVES perspective every combination of vehicle weight, fuel, technology, emission standard, and engine size represents a different source bin for a given vehicle age.

2.8.13.3 Operating Mode Distribution Data Generator

The Operating Mode Distribution Generator provides a mechanism to define the distribution of operating modes used to calculate emissions. For exhaust running emissions and energy consumption, this is a distribution of vehicle specific power (VSP) bins by time, location, and roadway type. This would typically be used to create the default or user specific profiles by roadway type to account for link or roadway specific characteristics. These characteristics might provide specific information for additional roadway types (e.g., ramps), the grade of the link, or the distribution of the drive schedules around a specific average speed value. Adjustments to parameters other than average speed are probably limited to applications involving detailed operational simulations. Table 50 lists the calculations performed by the Operating Mode Distribution Generator.

Table 50. Operating Mode Distribution Generator Steps

Step	Calculation
OMDG-1	Define drive schedules
OMDG-2	Define the distribution of drive schedules by average speed
OMDG-3	Calculate the distribution of drive schedules for a given link
OMDG-4	Calculate the second-by-second vehicle specific power
OMDG-5	Determine operating mode bin for each second
OMDG-6	Calculate operating mode fractions for each drive schedule
OMDG-7	Calculate operating mode fractions for each link
OMDG-8	Adjust operating mode fractions based on the grade of the link

2.8.13.4 Application Modes

MOVES supports three primary levels of analysis: macroscopic, mesoscopic, and microscopic. The macroscopic analysis is used by EPA to perform national-level estimates of energy consumption and policy-related studies. The mesoscopic analysis is focused on generating State Implementation Plans and Emissions Inventories for state and regional agencies responsible for air quality conformity analysis. The microscopic analysis or project level analysis is designed for hot-spot analysis of local project that have air quality implications. This level of analysis is intended to support Environmental Impact Statements.

Based on the primary purposes of the C10A model development effort, the mesoscopic or county-level analysis option is most appropriate. This level of analysis focuses on one or more counties within a region or state. Counties can be modeled independently or grouped into "custom domains". If modeled independently, the input tables containing total activity (VMT), fleet mix and age distributions, vehicle inspection programs, fuel types, and meteorology can be different for each county. If a custom domain is used, a single set of inputs are provided for the domain.

The decision to use individual counties or custom domains depends on available data and significant variations within the region. If vehicle populations, fleet and fuel data are only available at the regional level, a custom domain may be the only option. On the other hand, if the region includes different states or counties with different vehicle inspection programs, it may be necessary to subdivide the analysis. In this case, it is often desirable to define multiple custom domains. Custom domains can simplify input data preparation, but they also reduce or avoid the need for output processing. For a region that includes a large number of counties, merging the output databases into a single answer can be time consuming.

Within the mesoscopic or county-level analysis, MOVES supports two primary application methods. In the "Lookup Method", MOVES generates a table of emissions rates by vehicle type, facility type, speed bin, and various other classifications. This table can then be used by custom software to calculate and aggregate emissions from individual links. In the "Inventory Method", the transportation data are processed to generate a series of tables and distribution factors that MOVES importers can read for a MOVES emissions inventory application.

The primary advantage of the inventory method is that the emissions estimates can be more accurate by minimizing rounding and interpolation errors within the internal calculations. The application process can also include greater policy sensitivity by permitting the user to adjust the emissions assumptions and input data for each run. The primary disadvantage of this approach is the need to customize inputs and outputs for each application; the time required to run the MOVES software for one or more custom domains or counties; and the need for transportation modelers to possess skills and be trained in MOVES applications.

The lookup method has the advantage that the MOVES software is run once and the rates are applied to multiple scenarios or alternatives. Transportation modelers can focus on applying the rates and leave the details of setting up and applying the MOVES software to air quality experts. The process also has the advantage that the rates can be applied to individual links and aggregated in standard ways. It is also very similar to the way most MPO's applied the MOBILE6 software in the past.

The primary disadvantage of the lookup method is that the resulting rate tables may be too big and bulky for practical use. The attribute details created by MOVES are often not helpful and need to be restructured for efficient application. This restructuring can be challenging because individual rates need to be properly weighted to create aggregate rates that match the available transportation data.

Using the lookup emission rates with operational simulation models can also be problematic. Rates are available in 5 mph increments, but they should not be applied to instantaneous speeds. The rates are based on VMT distributions of average speeds. Since stopped vehicles do not generate VMT, there is no way to apply the lowest emission rate to a vehicle that is stopped for one or more seconds. The rate needs to be applied to the average speed of the vehicle over the length of the link or some other measure of distance.

2.8.13.5 TRANSIMS Interface

The TRANSIMS Emissions program is designed to support both county-level and project-level applications of the MOVES software. For county-level analysis, the program supports both lookup table and inventory application methods.

Figure 70 shows the TRANSIMS interface using MOVES lookup tables. In this approach, MOVES is applied once with appropriate local county-specific data to generate one or more lookup tables. The TRANSIMS Microsimulator is executed to generate speed bin files for each vehicle type. These files contain the number of seconds over each 15 minute period that each 30 meter segment of roadway has vehicles of the specified type traveling in each of six speed bins. The TRANSIMS Emission program is then executed with various parameters to aggregate some values and disaggregate other values in the MOVES lookup table and output the resulting composite rates. These rates can then replace the MOVES lookup table for subsequent applications. In addition, the Emissions program applies the composite rates to each record and aggregates the resulting emissions by facility type, vehicle type, and/or summary district.

For MOVES inventory applications, the process shown in Figure 71 is used. In this case the TRANSIMS Microsimulator generates the speed bin files and a link delay file. The link delay file contains the volume and speed on each link in 15 minute increments. The LinkSum program aggregates this information to generate the VMT by HPMS vehicle types and the distribution of VMT by MOVES facility types. The Emissions program in this case is configured to output VMT distributions and average speed bin distributions by hour of the day. These tables are generated in the format required by MOVES importers that insert the data into the MOVES MySQL database for emissions inventory processing.

Figure 70. Emission Rate Lookup Table Method

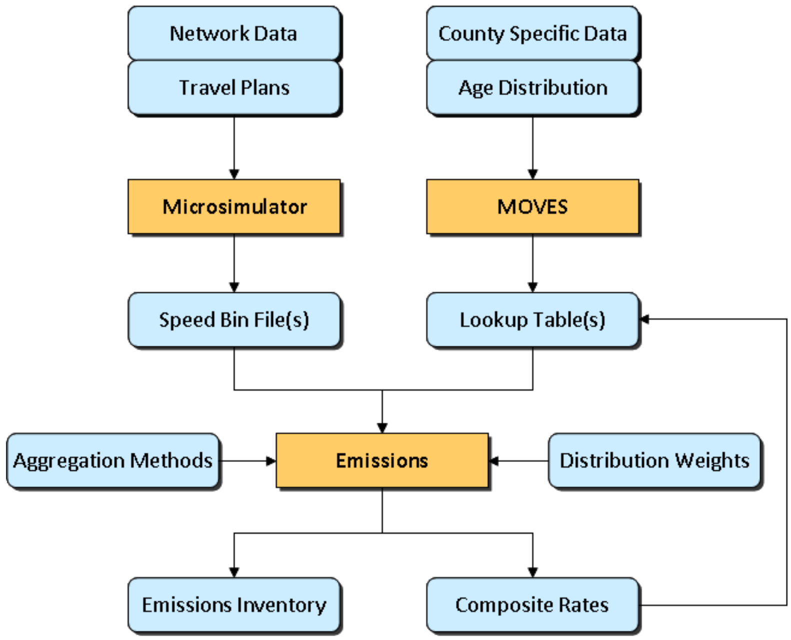
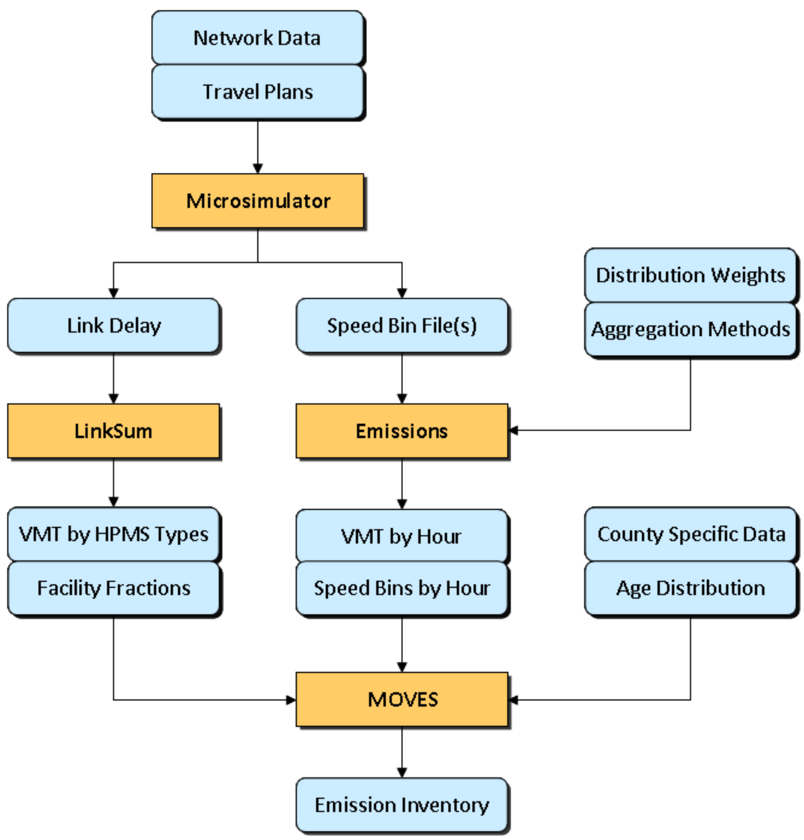


Figure 71. MOVES Emissions Inventory Method



Microsimulator Outputs

One of the key inputs to the emissions estimate is a set of Speed Bin files output from the TRANSIMS Microsimulator. Speed Bin files are generated for each vehicle type included in the network demand (i.e., travel plans). The control keys are listed in Table 51. At a minimum Speed Bin files are generated for autos and trucks. If the region includes different vehicle inspection programs for different subregions, the auto vehicle types should include separate vehicle type or subtype to aggregate the travel separately. For detailed emissions estimates the sample rate is once per second in 15 minute time increments and the links are subdivided into 30 meter segments (box length).

Table 51. Microsimulator Control Keys

Control Key	Description
OUTPUT_SPEED_FILE	File name to be created within the project directory
OUTPUT_SPEED_FORMAT	File format to be created (default Version3)
OUTPUT_SPEED_VEHICLE_TYPE	A vehicle type code number (default 0 = ALL)
OUTPUT_SPEED_FILTER	Minimum number of vehicles per time increment (default 1)
OUTPUT_SPEED_TIME_FORMAT	Output time format (default seconds)
OUTPUT_SPEED_INCREMENT	Time increment duration (default 24 hours)
OUTPUT_SPEED_TIME_RANGE	Time period range (default ALL)
OUTPUT_SPEED_LINK_RANGE	Link number range (default ALL)
OUTPUT_SPEED_SAMPLE_TIME	The time frequency in seconds at which the speed bins will be summarized (default 1 second)
OUTPUT_SPEED_BOX_LENGTH	The length in meters of the link segments for which speed bins are summarized (default = 0 = full link length)
OUTPUT_SPEED_NUM_BINS	The number of speed bins that are summarized (default = 6)

The output Speed Bin files include the metadata header record shown in Table 52 and the data fields listed in

Table 53.

Table 52. TRANSIMS Speed Bin MetaData

MetaData	Description
TIME_STAMP	The time and date when the file was created
BOX_LENGTH	The segment length in meters
CELL_LENGTH	The cell length used in the simulation
SAMPLE_TIME	The frequency in which data is collected (seconds)
INCREMENT	The summary time increment
VEHICLE_SUBTYPE	The subtype of the vehicle type summarized in the file
VEHICLE_TYPE	The vehicle type code summarized in the file
VELOCITY_BINS	The number of speed bins
VELOCITY_MAX	The maximum speed in meters per second

Table 53. TRANSIMS Speed Bin Data Fields

Field Name	Description
LINK	Link number
DIR	Direction of travel (0 = AB, 1 = BA)
OFFSET	Distance from the beginning of the link to the end of the segment
TIME	Ending time of the time increment
SPEED0	Total vehicle seconds at speed zero cells/second
SPEED1	Total vehicle seconds at speed one cell/second
SPEED2	Total vehicle seconds at speed two cells/second
SPEED3	Total vehicle seconds at speed three cells/second
SPEED4	Total vehicle seconds at speed four cells/second
SPEED5	Total vehicle seconds at speed five cells/second
SPEED6	Total vehicle seconds at speed six cells/second

Control keys for the TRANSIMS Emissions program provide the tools necessary to compress MOVES emissions rates into values that correspond to the transportation data and the analysis requirements. This includes selecting columns, column attributes for selecting rows, and providing weighting factors for combining emissions rates into weighted average values. Once the table is collapsed it can be output as a new emissions rate table for use as input to subsequent applications. An example of a collapsed emissions rate table is shown in Table 54.

Table 54. Collapse Emission Rate Table

yearID	monthID	sourceTypeID	roadTypeID	pollutantID	processID	avgSpeedBinID	emissionRate
2008	1	21	2	1	1	1	1.91824
2008	1	21	2	1	1	2	1.02998
2008	1	21	2	1	1	3	0.608886
2008	1	21	2	1	1	4	0.430296
2008	1	21	2	1	1	5	0.37313
2008	1	21	2	1	1	6	0.318093
2008	1	21	2	1	1	7	0.2814
2008	1	21	2	1	1	8	0.258368
2008	1	21	2	1	1	9	0.241404
2008	1	21	2	1	1	10	0.228209
2008	1	21	2	1	1	11	0.217652
2008	1	21	2	1	1	12	0.206746
2008	1	21	2	1	1	13	0.195869
2008	1	21	2	1	1	14	0.208396
2008	1	21	2	1	1	15	0.237551
2008	1	21	2	1	1	16	0.276734

The TRANSIMS Emissions program applies these rates to the Microsimulator Speed Bin data. This involves mapping TRANSIMS facility and area types to MOVES road types and TRANSIMS vehicle types to MOVES source types. Summary years, months and weekend travel factors are specified. As the Speed Bin data are read for each link segment and time period, the vehicle seconds in each TRANSIMS speed bin is distributed to the 16 speed bins defined by MOVES. This distribution process ensures that the total VMT and VHT included in the TRANSIMS speed bins is equal to the total VMT and VHT represented in the MOVES speed bins. The appropriate emissions rates are applied to the VMT in each speed bin and summarized as requested. Emissions summary reports can be generated by area type, facility type, vehicle type, road type, area and facility types, area and vehicle types, facility and vehicle types, road and vehicle types, and total emissions. Emissions summary data can also be written to a file for additional processing.

Emissions Program Applications – Emission Inventory Method

The TRANSIMS Emission program can also be used to generate input tables in the format required by the MOVES county-level data importers. The TRANSIMS Speed Bin and Link Delay datasets provide the information needed for five of the MOVES input tables. County or custom domain attributes such as temperature and relative humidity and vehicle population data related to fuel and age distributions need to be provided from other sources.

The interface includes many of the same elements as a lookup table application, but the process is reversed. Rather than collapse or convert MOVES emissions rates to TRANSIMS data elements, this process expands or converts TRANSIMS data to MOVES data classifications. For example, TRANSIMS facility and area types are collapsed to MOVES road types; TRANSIMS vehicle types are expanded to MOVES source types; and TRANSIMS speed bins are distributed to MOVES speed bins. As each TRANSIMS Speed Bin record by link segment and time period is read, the data is converted and aggregated into an appropriate MOVES-related data structure. The MOVES data is processed, formatted, and output as tab delimited data files.

The first table required by the MOVES Emissions Inventory process is a distribution of annual VMT by HPMS vehicle types. A mapping between TRANSIMS vehicle types and HPMS vehicle types is provided along with distribution fraction as necessary. An expansion factor is provided to convert the daily TRANSIMS VMT to annual VMT. This factor may also include some consideration for travel on roadways not included in the TRANSIMS network. Table 55 shows an example of the HPMS VMT distribution.

Table 55. VMT by HPMS Vehicle Type

HPMSVtypeID	yearID	VMTGrowthFactor	HPMSBaseYearVMT	baseYearOffNetVMT
10	2008	0	0	0
20	2008	0	7528117453	0
30	2008	0	3051203213	0
40	2008	0	361066.94	0
50	2008	0	3266014.62	0
60	2008	0	5935721.71	0

The next table provides factors to distribute the VMT assigned to each MOVES source type to road types. In addition to mapping TRANSIMS vehicle types to HPMS vehicle types, TRANSIMS vehicle types also need to be mapped to MOVES source types. TRANSIMS facility and area types are also mapped to MOVES road types. The VMT by vehicle type is summed by road type and then the road type fractions are calculated for each vehicle type. Since the purpose of this table is to distribute VMT assigned to a given source type to road types, the same fractions can be used for each source type associated with a given vehicle type. Table 56 provides an example of road type fractions for two source types.

Table 56. VMT Road Type Fractions

sourceTypeID	roadTypeID	roadTypeVMTFraction
21	1	0
21	2	0.12795
21	3	0.095435
21	4	0.399313
21	5	0.377302
41	1	0
41	2	0.201374
41	3	0.030168
41	4	0.696728
41	5	0.07173

Since MOVES road types are limited to restricted access and unrestricted access categories and speed profiles on freeways are considerably different than ramps, MOVES splits the VMT assigned to the restricted access road type into freeways and ramps using a ramp fraction table. This fraction is simply the total VMT on ramps in urban or rural area types divided by the total VMT on ramps plus freeways (and expressways). An example of the ramp fractions file is shown in Table 57.

Table 57. Ramp Fractions

roadTypeID	roadDesc	rampFraction
1	Off-Network	0
2	Rural Restricted Access	0.056354
3	Rural Unrestricted Access	0
4	Urban Restricted Access	0.084319
5	Urban Unrestricted Access	0

The fourth table distributes daily VMT associated with a given source type and road type to VMT by hour of the day. The fractions can also vary for weekdays and weekends. The dayID field distinguishes a weekday (5) from a weekend (2). MOVES uses one distribution for Monday through Friday and the other distribution for Saturday and Sunday. Total weekend VMT is modeled as a fraction of total weekday VMT. Table 58 shows an example of the hourly distribution of VMT assigned to a given combination of source type, road type, and day type.

Table 58. VMT Hour Fractions

sourceTypeID	roadTypeID	dayID	hourID	hourVMTFraction
21	2	2	1	0.004541
21	2	2	2	0.003671
21	2	2	3	0.003132
21	2	2	4	0.003283
21	2	2	5	0.005489
21	2	2	6	0.015789
21	2	2	7	0.039964
21	2	2	8	0.069848
21	2	2	9	0.073569
21	2	2	10	0.057104
21	2	2	11	0.051712
21	2	2	12	0.055294
21	2	2	13	0.061564
21	2	2	14	0.061838
21	2	2	15	0.064897
21	2	2	16	0.073156
21	2	2	17	0.081059
21	2	2	18	0.083631
21	2	2	19	0.066849
21	2	2	20	0.045871
21	2	2	21	0.028413
21	2	2	22	0.023375
21	2	2	23	0.015719
21	2	2	24	0.010232

The fifth table is perhaps the most important. It distributes the VMT assigned to each combination of source types, road type, day type, and hour of the day to average speed bins. MOVES includes 16 speed bins in 5 mph increments. The amount of VMT assigned to each speed bin is critical to the emissions calculations. The shape of the distribution defines the operating mode distribution, driving schedules, and the vehicle specific power bins used to calculate emissions.

The data from the TRANSIMS Speed Bin files are distributed to source types using the vehicle type to source type map and source type factors. The link facility and area type attributes map the link segment to a MOVES road type. The 15 minute time periods are summed to hours of the day. The vehicle seconds in each TRANSIMS speed bin is then distributed to the 16 MOVES speed bins. This distribution process ensures that the total VMT and VHT included in the TRANSIMS speed bins is equal to the total VMT and VHT represented in the MOVES speed bins. The VMT in each speed bin is divided by the total VMT for the hour to set the average speed fraction. An example of the speed bin distribution for one classification category is shown in Table 59.

Table 59. Average Speed Bin Distribution

sourceTypeID	roadTypeID	hourDayID	avgSpeedBinID	avgSpeedFraction
21	2	12	1	0.004948
21	2	12	2	0.004122
21	2	12	3	0.003
21	2	12	4	0.002265
21	2	12	5	0.002105
21	2	12	6	0.003277
21	2	12	7	0.00927
21	2	12	8	0.019876
21	2	12	9	0.04253
21	2	12	10	0.093737
21	2	12	11	0.152748
21	2	12	12	0.169864
21	2	12	13	0.125502
21	2	12	14	0.072482
21	2	12	15	0.065015
21	2	12	16	0.229258

The tables created by the TRANSIMS Emissions program are then imported into the MOVES database, various MOVES parameters are set, and a MOVES run is executed. If multiple counties or custom domains are required, the MOVES databases will need to be combined to create the total emissions inventory. Data can then be selected from the tables to generate summary reports.

3.0 Model Calibration / Validation

3.1 Model System Calibration / Validation Process

Because the Jacksonville DaySim implementation was "transferred" from Sacramento, and the model coefficients and alternative-specific constants were initially estimated and calibrated for the Sacramento region, it was necessary to recalibrate the core model components to reflect Jacksonville region-specific travel patterns. Calibration and validation of the entire model system is a highly iterative process that involves making changes to individual model components to better match observed data sources, and evaluating the impacts of these changes on other model components and on overall model system performance. One of the advantages of disaggregate nature of activity-based microsimulation models such as DaySim is that they support more flexibility and realistic calibration adjustments than is possible with aggregate trip-based models. Note that a calibration effort was not performed for the Burlington implementation.

3.1.1 Observed Data Sources

Prior to the calibration of the core behavioral components, it was necessary to prepare observed datasets against which to compare the model outputs. The primary observed data source for the calibration of the core DaySim component models was the National Household Travel Survey (NHTS) data collected in 2008-2009. For some model components such as household vehicle availability model and the work tour destination model, the NHTS was supplemented with observed information from the 2005-2009 American Community Survey (ACS) data. Because the focus of the C10A effort is on a region in which choices of non-highway modes are limited, and thus the dynamic integrated model represents behavioral changes in response to roadway conditions primarily, detailed transit information such as an on-board survey was not used in this effort. The observed transit mode share for the Jacksonville region is less than one percent.

In order to support the calibration of the DaySim models, it was necessary to first process the NHTS household, person and trip records to create a new tour record file and to append additional information to the existing NHTS household, person, and trip files. A summary of the NHTS data available for calibration is shown in Table 60.

Although additional NHTS "add on" survey data was collected in the Jacksonville region, the overall number of households, persons, tours and trips was relatively small. Because DaySim models travel behavior for a "typical weekday," it was necessary to remove weekend days from the dataset, further reducing the sample size. Although the NHTS contains all the data items required for activity-based model system development, such a small regional sample is insufficient to completely estimate the coefficients in the DaySim component models. However, in the absence of any other datasets contain the information required for activity-based model development, it was deemed acceptable to use the NHTS to derive calibration targets.

Table 60. NHTS Summary Statistics for Jacksonville Region

County	HHs	Persons	Tours Total	Tours Weekday	Trips Total	Trips Weekday
Clay	658	1,365	1,717	1,304	4,630	4,108
Duval	205	448	599	438	1,628	1,428
Nassau	198	415	476	373	1,287	1,153
St Johns	79	171	203	172	580	548
Total	**1,140**	**2,399**	**2,995**	**2,287**	**8,125**	**7,237**

In addition to the relatively small sample size, a number of other issues arose when using the NHTS data. These issues included:

- The absence of any person, tour or trip information for children under 5 years of age: Although these people are reflect in the household size information, no travel behavior was recorded. As a result, all of the summaries of DaySim estimated travel behaviors also excludes travel by young children, in order to facilitate comparisons.
- Missing travel information for some members of the household: Typically all travel made by all members of each household is collected during the household survey data collection process. Having this complete set of travel demand is even more critical in the context of advanced activity-based model systems which consider all travel by all household members across all times of day, and which may also explicitly consider travel made jointly by members of the household. In order to address these missing persons, it was necessary to adjust the person weights so as to match regional controls of persons by person type developed to support creation of the synthetic population.
- Inconsistent expansion factors: When the household and person expansion factors provided with the NHTS were applied, significant discrepancies with other regional

person and household totals were observed, necessitating further adjustments to the expansion factors.

Finally, a number of the NHTS-derived summaries seemed inconsistent with other household travel surveys used for DaySim development. For example, the NHTS seemed to show relatively high shares of workers choosing to work at home. Nothwithstanding these issues, the NHTS was used as the primary data source in the absence of any viable alternatives.

3.1.2 Calibration Results

The following sections present the results of the initial calibration of the Jacksonville DaySim implementation. Although all model calibration adjustments have a simultaneous impact on the model predictions, the calibrate effort typically follows a sequentially process from the top to the bottom of the DaySim model hierarchy, because adjustments to upper level models will tend to impact lower level model predictions more than the reverse. The calibration results shown below follow this hierarchy.

It should be noted that these calibration results should not be considered "final" due to the fact that the project has involved the use of multiple sets of regional skims at different temporal resolutions and using different network simulation methods at different points in the project. For example, an initial calibration was performed using skims for four broad time periods. Subsequently, this calibration was revisited when Microsimulator-based skims for twenty-two time periods became available. The calibration was further revised when Router-based skims from the fully integrated model system were developed. It should also be note that the summaries shown are not exhaustive, and that additional summaries have been prepared and used in the calibration process.

3.1.2.1 Usual Work / School Location

The usual work and school location models are the first models in the DaySim model system, and predict the usual destination parcels for work and school tours. Information on workplace locations can then be used in subsequent model components, such auto ownership. The work and school location models, as well as all the tour destination choice models, assume a single anchor point, the tour origin, from which impedance is measured, without direct consideration of the impedance for stops on the way to and from the tour destination. For the usual work and school location models, the anchor is the person's home. In these models, the home location is treated as a special location, because it occurs with greater frequency than any given non-home location, and size and impedance are not meaningful attributes. In addition, the model incorporates availability constraints – for example, only parcels with grade enrollment are available as school tour destinations for children.

Table 62 shows that the model system achieves a relative good match between overall average work tour lengths. DaySim predicts shorter work tours for part-time workers than observed in the NHTS, and longer student work tours. Figure 72 illustrates that the overall distribution of estimated work tour lengths matches the distribution of observed work tours. This figure also illustrates that, even with relative large travel markets such as work purpose tours, the observed data derived from NHTS shows a fair amount of variation. Finally, Table 62 summarized commute flows from the 2005-2009 ACS data. This table demonstrates that the model system is doing a reasonable job of capturing these flows, although Duval County (which contains Jacksonville) is slightly over-predicted as a commute destination while St Johns county (which is along the coast) is underpredicted. Note that a "cleaned" business employment database was not available for St Johns County and it was necessary to adjust the

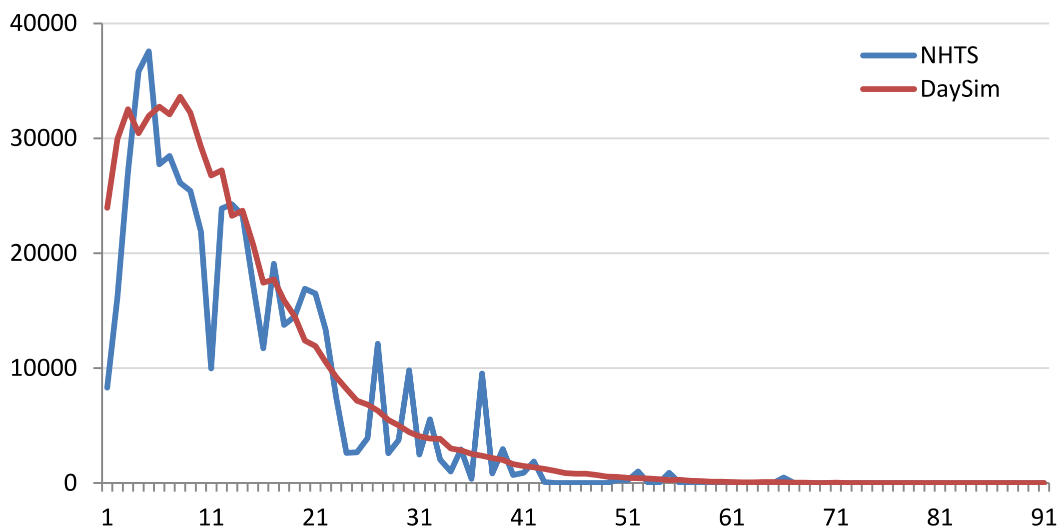
St Johns existing employment to match estimates of county employment derived from external sources.

Table 61. Average Work Tour Length (miles)

	NHTS	DaySim
Full time workers	13.67	13.51
Part time workers	9.75	8.10
Students	5.71	6.84
Total	12.99	12.53

Sources: DaySim, NHTS

Figure 72. Distribution of Work Tour Lengths



Sources: DaySim, NHTS

Table 62. Worker Flows by County (2005-2009 ACS)

O/D	Clay	Duval	Nassau	St Johns	Total
Clay	-1.2%	1.2%	0.0%	0.0%	0.0%
Duval	0.5%	-1.0%	0.4%	0.2%	0.0%
Nassau	0.0%	0.5%	-0.5%	0.0%	0.0%
St Johns	0.3%	0.8%	0.0%	-1.2%	0.0%
Total	**-0.4%**	**1.5%**	**-0.1%**	**-1.0%**	**0.0%**

Sources: DaySim, 2005-2009 ACS

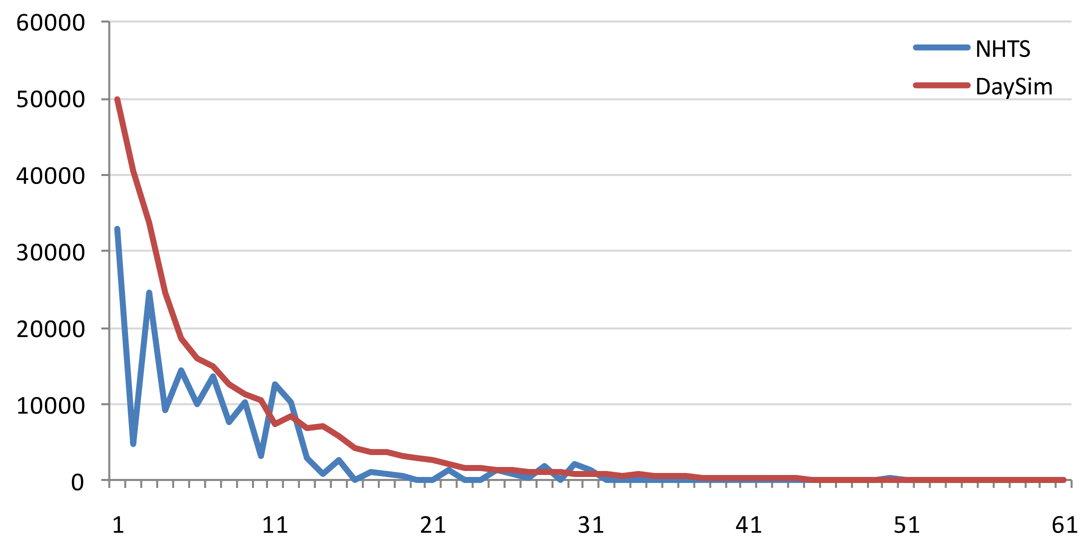
Table 63 and Figure 73 summarize the school usual location model results. Overall, the current calibration of DaySim is producing longer school tours than observed. DaySim predicts slightly longer school tours for grade school students and university students, and slightly shorter school tours for high school students. As seen with the work tours, the observed data derived from NHTS shows a fair amount of variation in school tour lengths.

Table 63. Usual School Location Average Distance

	NHTS	DaySim
Grade School	5.25	5.41
High School	6.56	5.87
University	11.67	12.67
Total	**6.64**	**7.14**

Sources: DaySim, NHTS

Figure 73. Distribution of School Tour Lengths



Sources: DaySim, NHTS

3.1.2.2 Vehicle Availability

The vehicle availability model predicts the number of motorized vehicles owned, leased, or otherwise belonging to the fleet of vehicles possessed by the household. The vehicle availability model takes as given the household characteristics, as well as the regular work location information of all workers in the household.

To calibrate and validate the model, the estimated share of households in each vehicle availability category was compared to the observed shares of households along three primary dimensions: household potential drivers, household income and household residence county. Two primary sources for observed data were identified and summarized: the 2005-2009 American Community Survey data and 2009 NHTS. Table 66 shows the difference between the original estimated and observed (ACS) shares of households by auto availability and county of residence. This table illustrates that the model is underpredicting 0-vehicle and 2-vehicle households and overpredicting 1-vehicle and 3+ vehicle households, suggesting that further calibration of this model is warranted.

Table 64. Observed Households by Vehicle Availability

County	0	1	2	3	4+	Total
Clay	0.3%	3.3%	5.7%	2.2%	0.8%	12.3%
Duval	5.4%	26.3%	27.0%	7.7%	2.7%	69.1%
Nassau	0.2%	1.4%	2.1%	1.0%	0.4%	5.0%
St Johns	0.5%	4.2%	6.6%	1.6%	0.6%	13.6%
Total	**6.4%**	**35.1%**	**41.4%**	**12.6%**	**4.4%**	**100.0%**

Sources: 2005-2009 ACS

Table 65. Estimated Households by Vehicle Availability

County	0	1	2	3	4+	Total
Clay	0.1%	3.8%	5.5%	2.0%	0.9%	12.3%
Duval	3.7%	27.0%	26.3%	8.6%	3.5%	69.1%
Nassau	0.0%	1.7%	2.2%	0.7%	0.3%	5.0%
St Johns	0.1%	4.9%	5.9%	1.9%	0.8%	13.6%
Total	**3.9%**	**37.4%**	**40.0%**	**13.3%**	**5.5%**	**100.0%**

Sources: DaySim

Table 66. Difference in Households by Vehicle Availability

County	0	1	2	3	4+	Total
Clay	-0.2%	0.5%	-0.1%	-0.3%	0.1%	0.0%
Duval	-1.7%	0.7%	-0.7%	0.9%	0.8%	0.0%
Nassau	-0.2%	0.3%	0.1%	-0.2%	0.0%	0.0%
St Johns	-0.4%	0.7%	-0.7%	0.3%	0.2%	0.0%
Total	**-2.5%**	**2.3%**	**-1.5%**	**0.6%**	**1.1%**	**0.0%**

Sources: DaySim, ACS

3.1.2.3 Day Pattern

The day pattern model predicts the number and purpose of tours and intermediate stops made by each individual. These predictions arise from a series of sequential sub-models that address different aspects each individual's daily activity pattern. The main activity pattern model predicts whether or not a person participates in any tours and intermediate stops for each of the seven different activity purposes, and then the exact number of tours made for that purpose during the full day. Another submodel predicts the number and purpose of work-based subtours, while a final submodel predicts the number and purpose of intermediate stops:

Calibration targets for the day pattern model calibration were derived from the NHTS. The full set of targets addressed tours by person type, tour and stop combinations by person type, exact numbers of tours and stops by purpose and person type, exact number and purpose of work-based sub-tours by person type, numbers of stops by tour purpose, and the exact numbers of tours and stops by person type were prepared. The estimated results produced by the activity generator were then compared to these targets. The calibration and validation process primarily involved making adjustments to alternative specific constants and reviewing and revising estimated parameters to ensure reasonability and consistency.

Table 67 compares the total number of tours for each of the destination purpose predicted by DaySim to the NHTS observed tours by destination purpose. This table illustrates that overall DaySim is matching overall regional tours relatively well, with 4% too many tours across all purposes. Tours by individual destination purpose match reasonably well, with the exception of work tours, which are overpredicted by 12%. Further adjustment to the calibration to address this overprediction is recommended.

Table 67. Tours By Destination Purpose

Purpose	NHTS	DaySim	Diff	% Diff
work	432,006	485,234	53,228	12%
school	176,802	184,635	7,833	4%
escort	135,493	128,014	-7,479	-6%
pers.bus	111,630	108,583	-3,047	-3%
shop	216,455	225,625	9,170	4%
meal	59,408	59,031	-377	-1%
soc/rec	244,219	242,368	-1,851	-1%
Total	**1,376,013**	**1,433,490**	**57,477**	**4%**

Sources: DaySim, NHTS

Table 68 summarizes the estimated and observed trips by destination purpose. This table demonstrates that DaySim is matching overall trips extremely well, although some of the individual purposes could use further refinement. Specifically, work trips are overpredicted, which is consistent with the overprediction of work tours, while school trips are underpredicted.

Table 68. Trips by Destination Purpose

Purpose	NHTS	DaySim	Diff	% Diff
work	730,988	797,150	66,162	9%
school	209,466	189,977	-19,489	-9%
escort	265,299	254,821	-10,478	-4%
pers.bus	251,000	250,055	-945	0%
shop	646,348	651,131	4,783	1%
meal	199,045	203,632	4,587	2%
soc/rec	395,372	375,919	-19,453	-5%
home	1,467,457	1,434,811	-32,646	-2%
Total	**4,164,975**	**4,157,496**	**-7,479**	**0%**

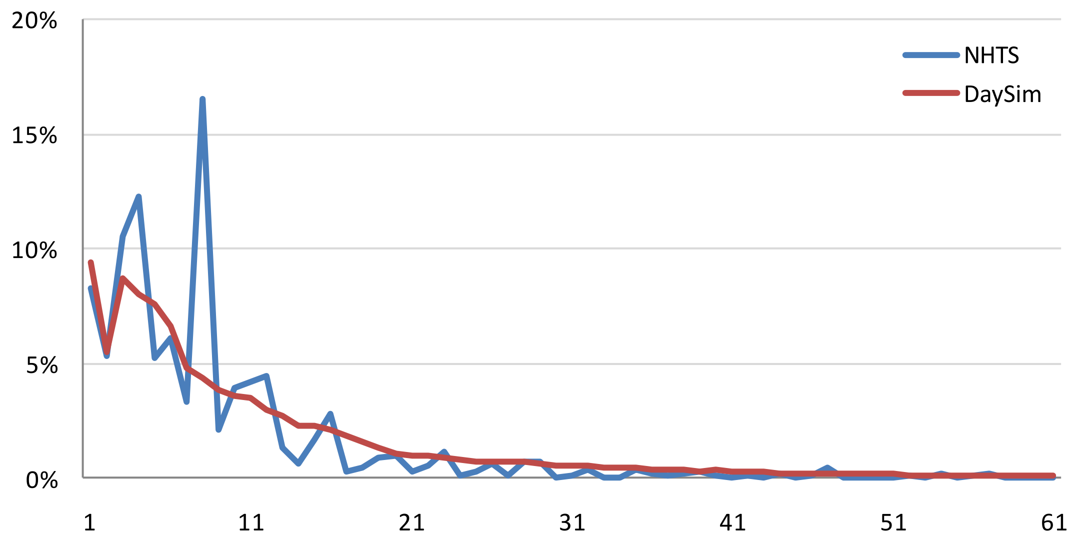
Sources: DaySim, NHTS

3.1.2.4 Tour / Stop Destination

Destinations for each of the tours predicted by the daily activity pattern models area predicted by a purpose-segmented destination choice models, using information about network impedances, purpose-specific size terms, and household and person attributes. As described in Section 3.1.2.1, the tour destination choice models predict a specific parcel as a destination, and assume a single home anchor point, the tour origin, from which impedance is measured, without direct consideration of the impedance for stops on the way to and from the tour destination. Unlike the usual work location, which has a nested structure to reflect the treatment of home as a special location, the destination choice models consider all parcels in a multinomial structure, subject to availability constraints.

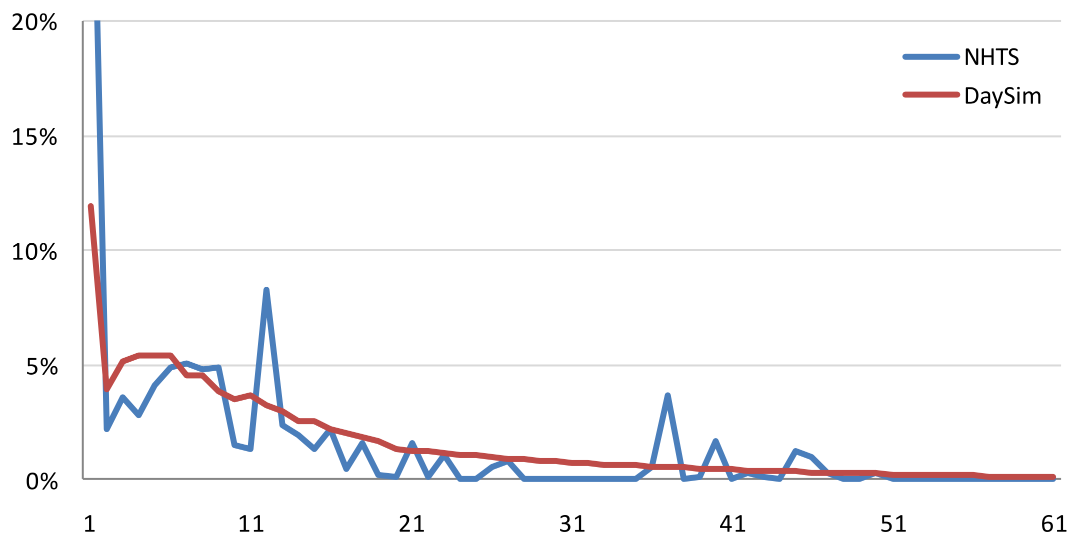
Figure 74 through Figure 78 compare the observed NHTS tour length frequencies to the estimated DaySim tour length frequencies. These figures demonstrate that DaySim matches observed data reasonably well, but they also illustrate the variations in some of the NHTS data, especially for purpose not well represented in the NHTS, such as meal tours.

Figure 74. Distribution of Shop Tour Lengths



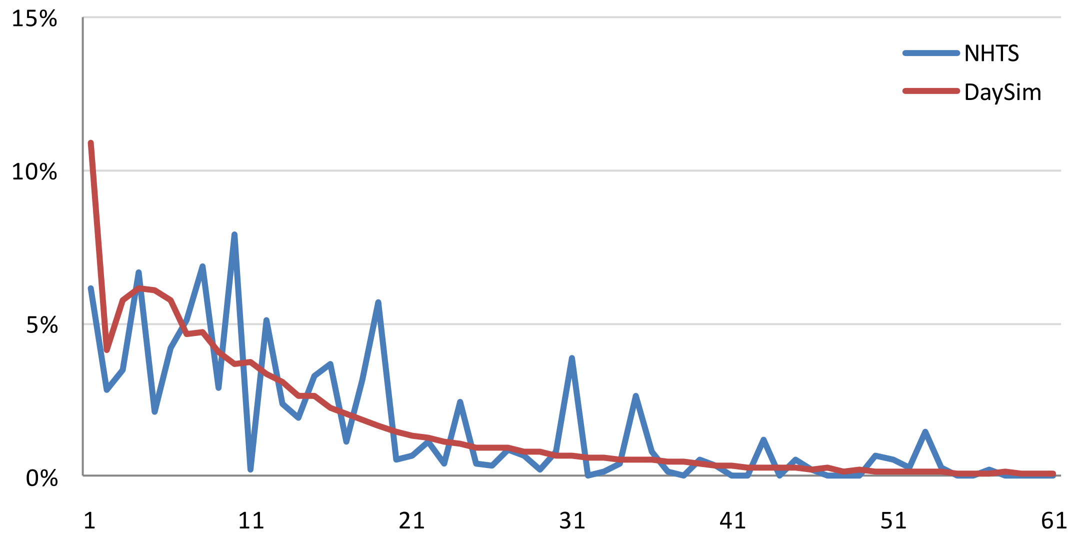
Sources: DaySim, NHTS

Figure 75. Distribution of Social/Recreation Tour Lengths



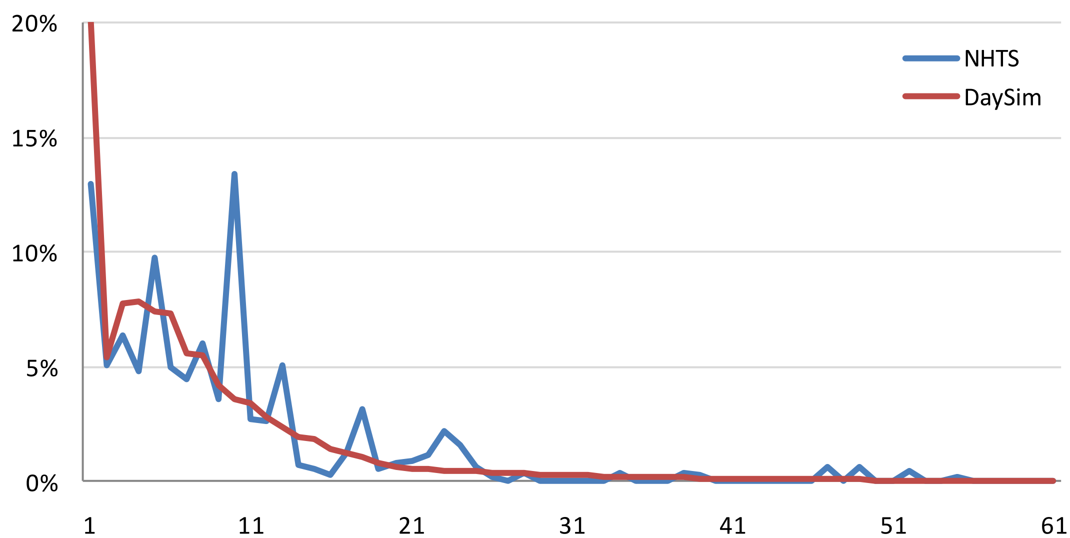
Sources: DaySim, NHTS

Figure 76. Distribution of Personal Business Tour Lengths



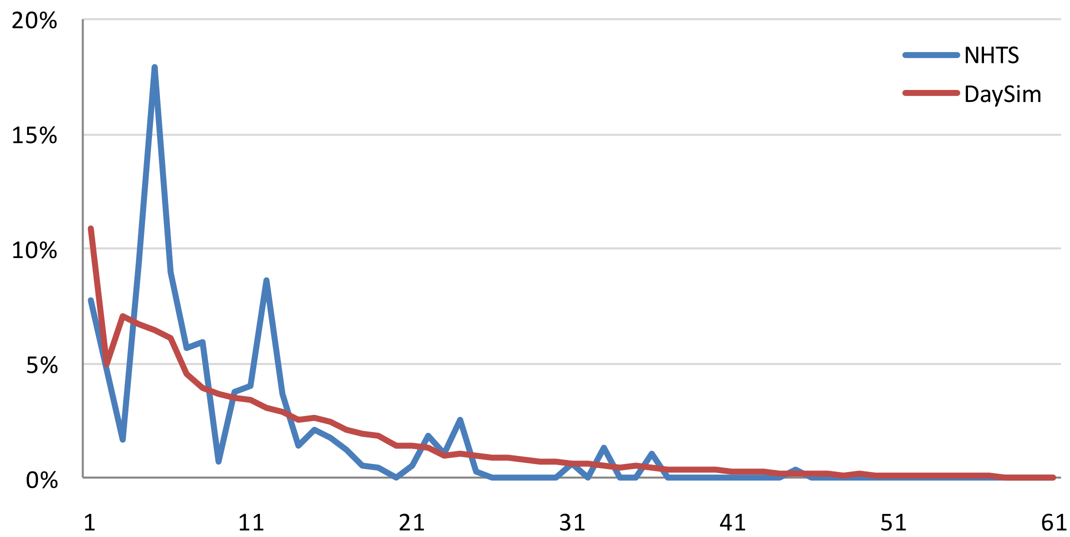
Sources: DaySim, NHTS

Figure 77. Distribution of Escort Tour Lengths



Sources: DaySim, NHTS

Figure 78. Distribution of Meal Tour Lengths



Sources: DaySim, NHTS

3.1.2.5 Trip Mode

The DaySim model system incorporates two sets of mode choice models. The tour mode choice model predicts the primary mode used for a tour, while the trip mode choice model predicts the mode used for each individual trip on the tour, constrained by the tour mode. The tour and trip mode choice models incorporate a variety of network impedance, household and purpose attributes, and even land use attributes. The core mode choice models incorporate the following modes:

- Drive to Transit
- Walk to Transit
- School Bus
- Shared Ride 2
- Drive Alone
- Bike
- Walk

Table 69 through Table 71 summarize the observed and estimation mode shares by trip destination purpose. Although the mode choice model calibration process involves making adjustments to both the tour models and trip models, it is the trip mode choice model outputs that are used directly in the network assignment process, and thus only the trip mode results are reported below. These tables demonstrate that DaySim does a reasonably good job at matching aggregate mode shares, although the calibration by purpose could be improved. Overall, drive alone shares are overpredicted by 3.6%, while walk shares are underpredicted by -2.4%.

Note that the tables below include additional modes not included in the list above. In the context of existing activity-based model implementations that have used static network assignment procedures, shared ride trips are simply aggregated to the zonal level, and divided by an assumed occupancy rate in order to calculate vehicle trips. This approach does not work in a disaggregate assignment simulation such as TRANSIMS because the goal is to preserve the details about each individual trip. Because it isn't possible to divide discrete shared ride trips by an occupancy rate in order to estimate vehicle trips, it is necessary to assign driver and passenger status to travelers whose mode is identified as shared ride. DaySim does not predict whether a person is an auto driver or the passenger for shared ride tours and trips, so a detailed analysis was used to derive a method for assigning an auto driver or passenger designation to each auto tour and trip, based on the modes used on a given tour.

Table 69. Observed Trip Mode Shares by Destination Purpose

Mode	Work	School	Escort	PersBus	Shop	Meal	SocRec	All Purp
Drive Alone	80.7%	17.9%	29.9%	48.8%	56.0%	35.1%	24.3%	52.8%
SR2 - Driver	8.0%	7.3%	33.4%	17.8%	14.2%	24.6%	8.7%	12.8%
SR2 - Passenger	4.6%	20.8%	4.9%	19.4%	13.3%	8.2%	7.5%	9.1%
SR3+ - Driver	3.3%	3.0%	18.8%	3.3%	5.3%	9.6%	9.0%	6.2%
SR3+ - Passenger	1.3%	39.0%	8.9%	8.7%	5.5%	13.0%	15.3%	9.9%
Drive-Transit-Walk	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Walk-Transit-Drive	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Walk-Transit-Walk	0.0%	0.0%	0.0%	0.6%	0.0%	0.0%	2.4%	0.4%
School Bus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Bike	0.0%	3.0%	1.4%	0.1%	0.3%	2.5%	6.0%	1.3%
Walk	2.1%	8.9%	2.6%	1.3%	5.5%	7.0%	26.7%	7.5%
Total	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**

Source: NHTS

Table 70. Estimated Trip Mode Shares by Destination Purpose

Mode	Work	School	Escort	PersBus	Shop	Meal	SocRec	All Purp
Drive Alone	83.5%	14.4%	22.4%	57.4%	52.6%	28.4%	43.7%	56.4%
SR2 - Driver	8.2%	6.9%	27.5%	16.1%	18.4%	32.6%	18.1%	14.3%
SR2 - Passenger	1.3%	22.3%	10.3%	10.7%	11.7%	14.8%	15.4%	9.0%
SR3+ - Driver	3.5%	6.8%	17.1%	5.3%	6.5%	9.9%	6.2%	6.1%
SR3+ - Passenger	0.8%	30.1%	11.0%	7.0%	7.7%	10.6%	10.7%	7.9%
Drive-Transit-Walk	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Walk-Transit-Drive	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Walk-Transit-Walk	0.2%	0.5%	0.0%	0.2%	0.0%	0.2%	0.1%	0.2%
School Bus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Bike	0.8%	4.0%	0.2%	0.6%	0.5%	0.2%	0.7%	0.9%
Walk	1.7%	15.0%	11.5%	2.8%	2.5%	3.3%	5.2%	5.1%
Total	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**

Source: DaySim

Table 71. Difference in Trip Mode Shares by Destination Purpose

Mode	Work	School	Escort	PersBus	Shop	Meal	SocRec	All Purp
Drive Alone	2.8%	-3.5%	-7.5%	8.6%	-3.4%	-6.7%	19.3%	3.6%
SR2 - Driver	0.2%	-0.4%	-5.9%	-1.7%	4.2%	8.0%	9.4%	1.6%
SR2 - Passenger	-3.3%	1.5%	5.4%	-8.7%	-1.5%	6.6%	7.9%	-0.1%
SR3+ - Driver	0.2%	3.8%	-1.7%	2.0%	1.3%	0.3%	-2.8%	0.0%
SR3+ - Passenger	-0.5%	-8.9%	2.1%	-1.8%	2.2%	-2.4%	-4.6%	-2.0%
Drive-Transit-Walk	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Walk-Transit-Drive	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Walk-Transit-Walk	0.2%	0.5%	0.0%	-0.4%	0.0%	0.2%	-2.4%	-0.3%
School Bus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Bike	0.8%	1.0%	-1.2%	0.4%	0.2%	-2.3%	-5.2%	-0.4%
Walk	-0.4%	6.1%	8.9%	1.6%	-3.0%	-3.7%	-21.6%	-2.4%
Total	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**

Sources: DaySim, NHTS

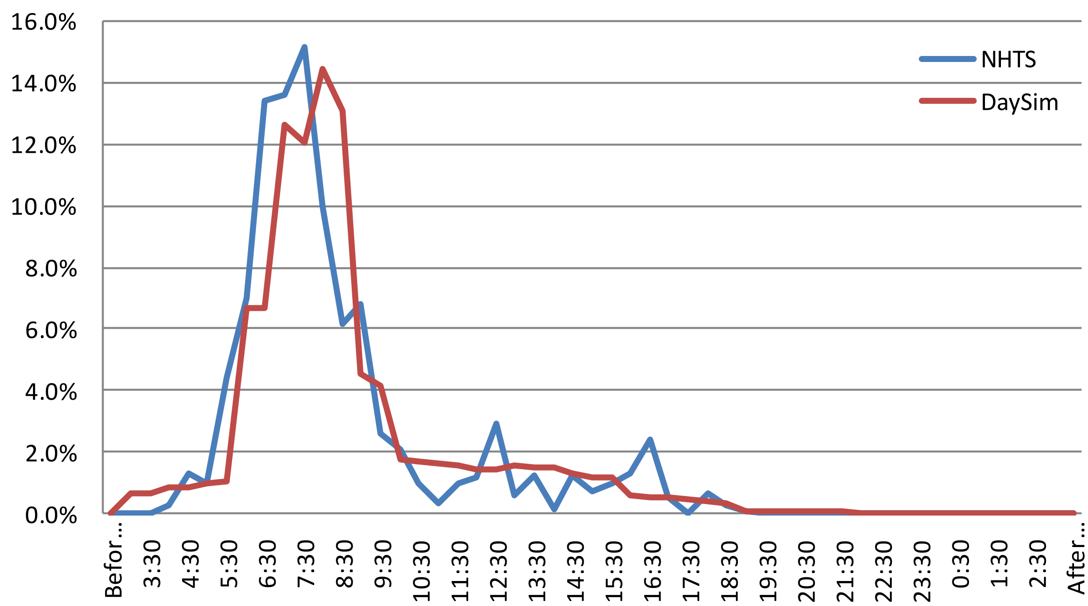
3.1.2.6 Tour / Trip Time-of-Day

One of the most compelling features of activity-based model approaches is that they have the capability to treat time explicitly and consistently across all travel choice dimensions. Rather than using fixed factors, or broad time periods, activity-based models can consider detailed time periods, as well as desired arrival times, departure times, and activity durations. This capability is essential to fulfilling goals of the C10A project, which is to make operational a dynamic integrated model that is sensitive to the dynamic interplay between travel behavior and network conditions.

In order to provide incorporate this sensitivity, DaySim includes two types of time-of-day models. Tour arrival and departure time at the primary destination models predict the time that the person arrives at the tour primary destination and the time that the person leaves that primary destination. Intermediate stop arrival or departure time models predict the time that the person arrives at the stop location (on the first half tour), or else the time that the person departs from the stop location (on the second half tour). The time of day models operate at a 30 minute time resolution – using 48 half-hour periods of the day. In addition, the models employ "time windows" when scheduling, so that when a tour or stop is scheduled, the portion of the window that it does not fill is left as two separate and smaller time windows.

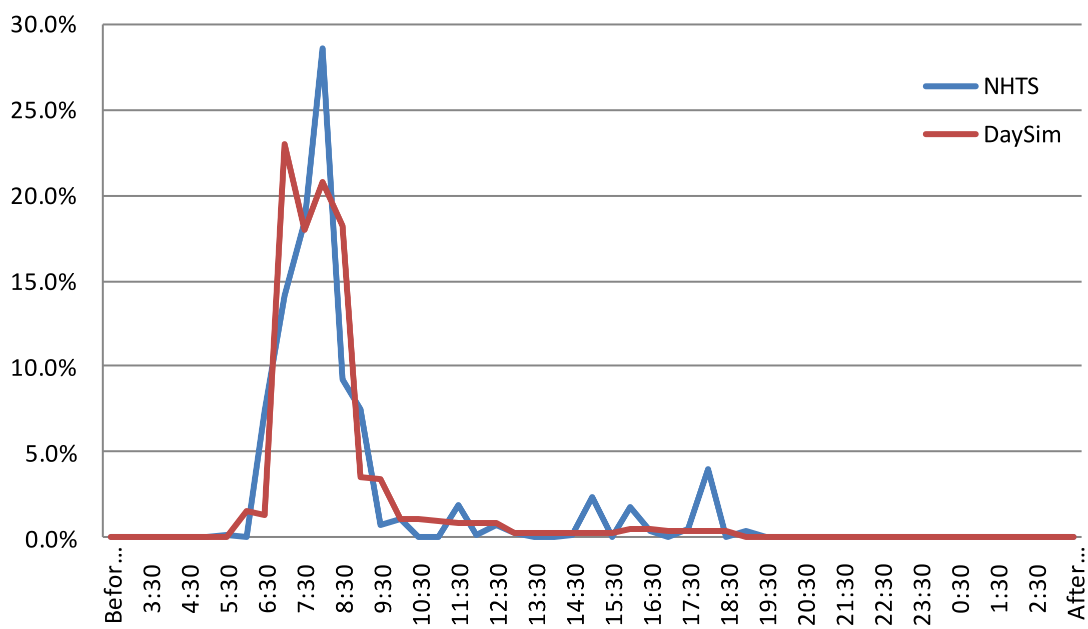
Figure 79 through Figure 82 compare the estimated and observed arrival times at tour primary destinations by purpose. Figure 79 shows the strong AM peak for arrival times at work, while Figure 80 shows an even stronger AM peak for arrival times at school. For both these tour purposes, the estimated and observed distributions of tour arrivals by half-hour are very similar. Figure 81 illustrates the tour arrival times for other purposes (shop, meal, escort, social/recreational, and personal business) are more evenly distributed across the day, and that the estimated and observed distributions are similar. The estimated and observed work-based subtours, shown in Figure 82, do not match as closely. This subset of tours uses the work locations as the anchors rather than the home locations. The observed NHTS data shows a strong peak at the midday, corresponding the lunch, while DaySim predicts more of these tours at other times of day.

Figure 79. Work Tour Arrival Times



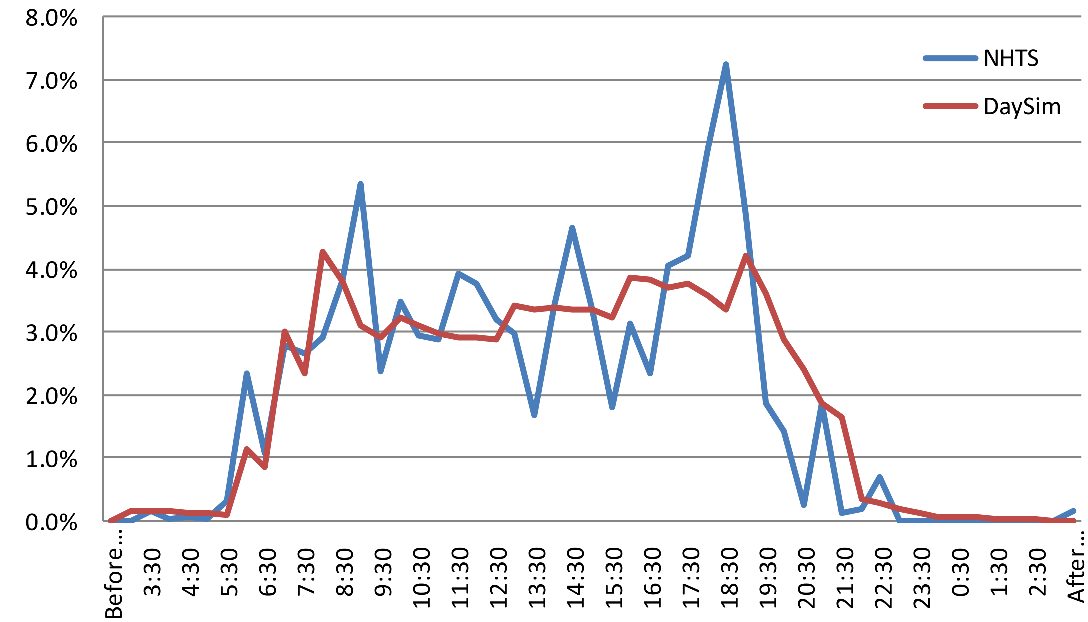
Sources: DaySim, NHTS

Figure 80. School Tour Arrival Times



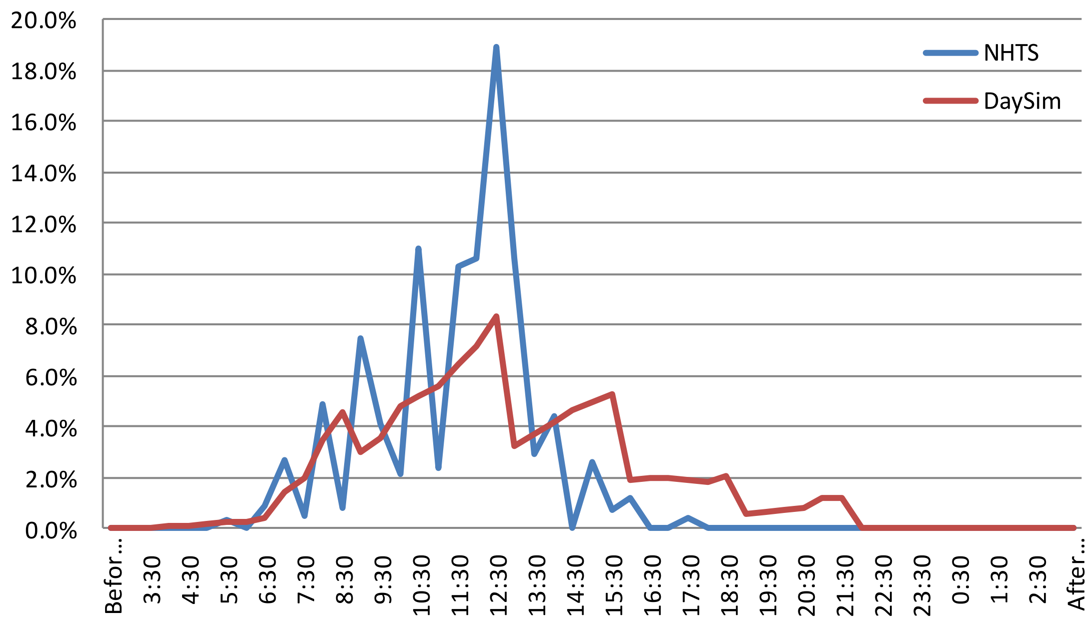
Sources: DaySim, NHTS

Figure 81. Other Arrival Times



Sources: DaySim, NHTS

Figure 82. Workbased Arrival Times

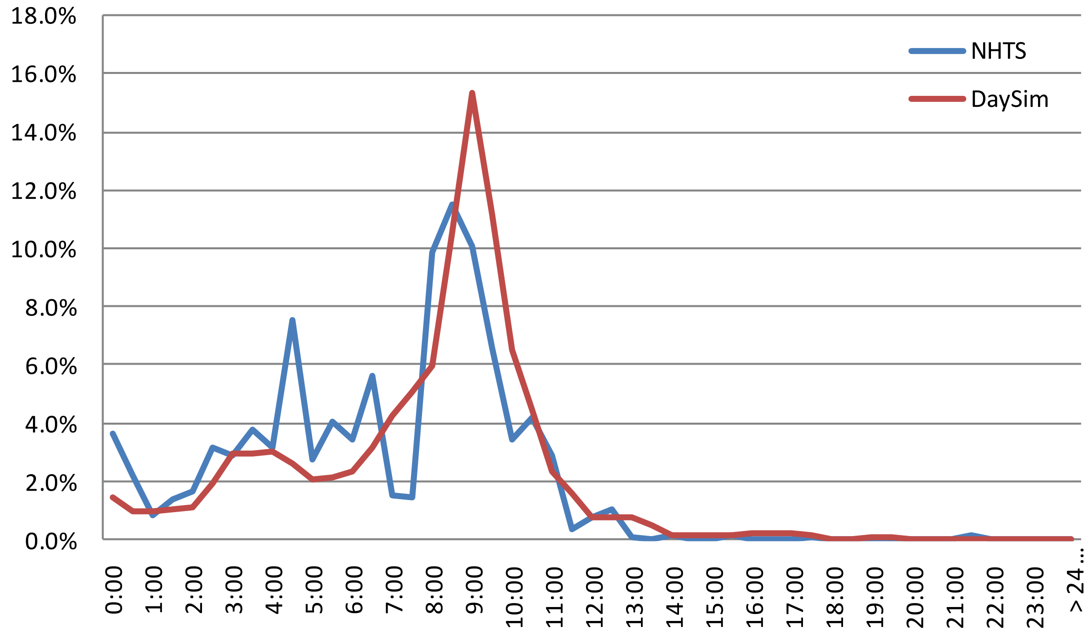


Sources: DaySim, NHTS

In addition to considering tour arrival and departure time, DaySim incorporates parameters related to the durations of activities. Figure 83 through

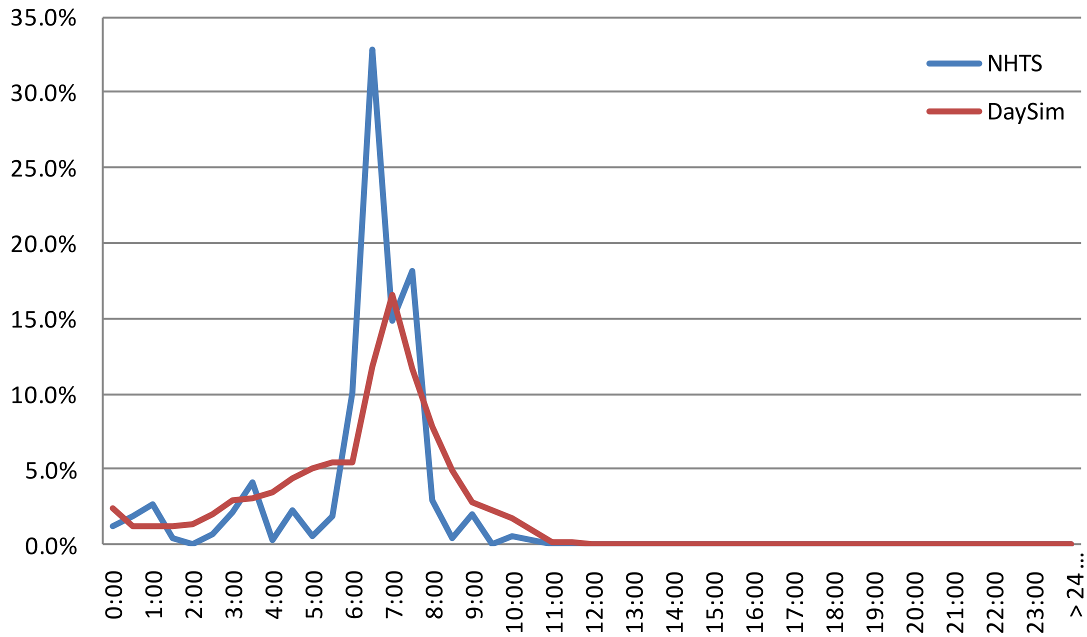
Figure 86 show the estimated and observed tour durations. Overall, the estimated and observed results are similar, although DaySim predicts a stronger peak at a 9 hour work tour activity, while the NHTS shows a stronger peak in school tour durations at 6 hours.

Figure 83. Work Tour Durations



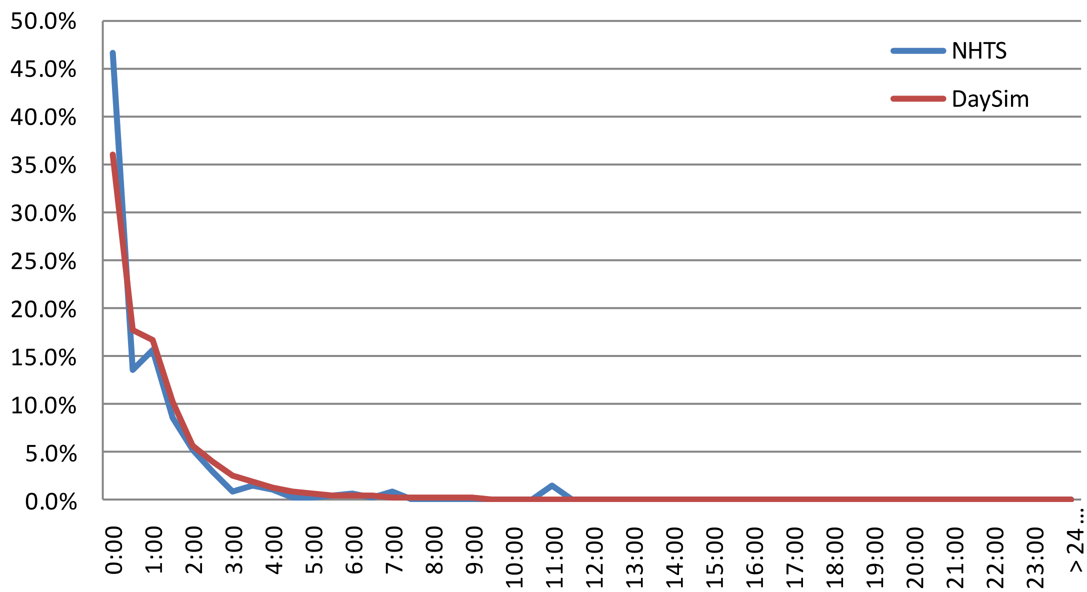
Sources: DaySim, NHTS

Figure 84. School Tour Durations



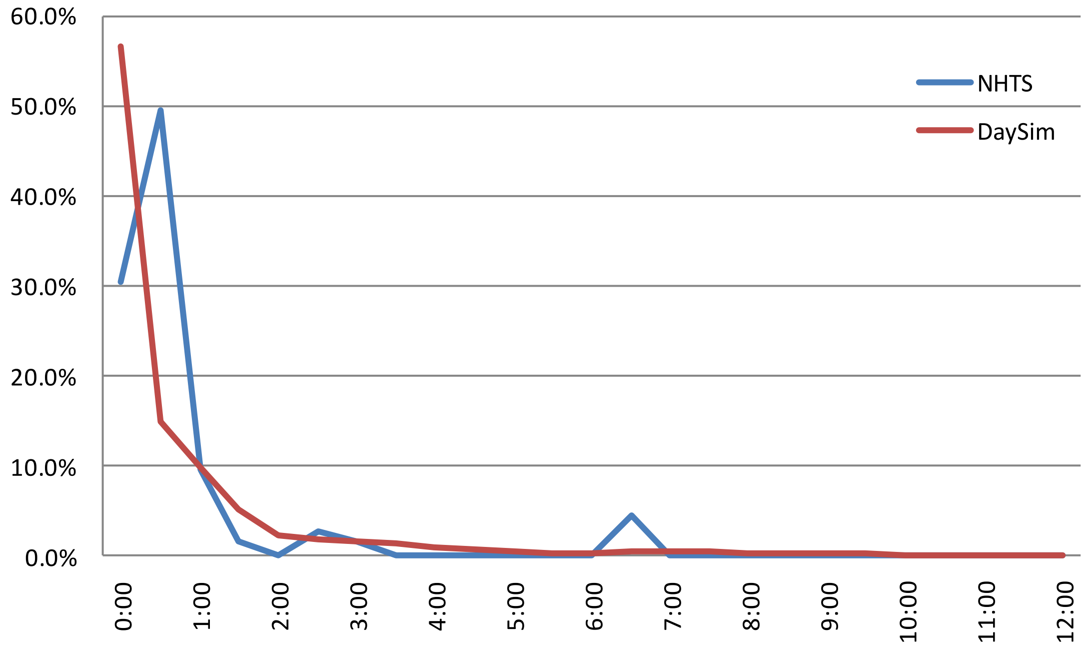
Sources: DaySim, NHTS

Figure 85. Other Tour Durations



Sources: DaySim, NHTS

Figure 86. Work-based Tour Durations



Sources: DaySim, NHTS

3.2 TRANSIMS Validation Process

This chapter documents the calibration and validation results of assigning the NERPM trip tables to the PLANNING level regional network. Validation tests were performed using all three network resolutions described earlier in this report, with a particular focus on the PLANNING and FINEGRAINED resolution networks. However, due to the significantly longer runtimes associated with the FINEGRAINED and ALLSTREETS networks, the PLANNING network has been the primary network resolution used in the integrated model system.

3.2.1 Observed Data Sources

The following three sources of 15-min count and data were compiled from the Florida Department of Transportation (FDOT):

1. Intelligent Traffic Systems (ITS) detectors on I-295 and I-95;
2. Portable Traffic Monitoring Stations (PTMS) on arterials and freeways; and
3. Telemetered Traffic Monitoring Stations (TTMS) on arterials and freeways.

The PTMS and TTMS data were obtained from FDOT's Transportation Statistics Office (TRANSTAT) and included only 15-minute vehicle counts. The ITS data included 15-minute count and speed data. All of these count data were collected in the years 2008-2009 and were processed as described later for comparison with the 2005 model year assignments.

With the help of the project team members at Florida International University (FIU), the observed traffic counts were scrubbed for spatial and temporal consistency. The data were then tagged to the NERPM master "merged" network to identify corresponding links in the 2005 NERPM network. The tagging process involved identifying a pair of nodes – Anode and Bnode – from the merged network for each of the count or speed locations. Tagging to the merged network ensured that the data can be transferred to the different network resolutions and modeling years without duplication of work. This process resulted in what is called a directional-data set where each record corresponds to a link direction, whether or not it is represented as a two way link in TRANSIMS.

During the network conversion process a link-node equivalence file is created by TransimsNet which lists the sequence of nodes that were merged to create each TRANSIMS link. This file is used by the LinkData program to transfer the directional count and speed datasets from the NERPM network links to the link numbers created for the PLANNING network.

During this data processing, issues with the original tagging process were identified and addressed, and further data checks were performed. Some data points needed to the merged because they were located on the same merged link. Similarly, a few data points needed to be dropped due to incomplete or erroneous data for all times of day or due to problems with proper identification of links.

The number of data points from each data source included in each step of the process is shown in Table 72.

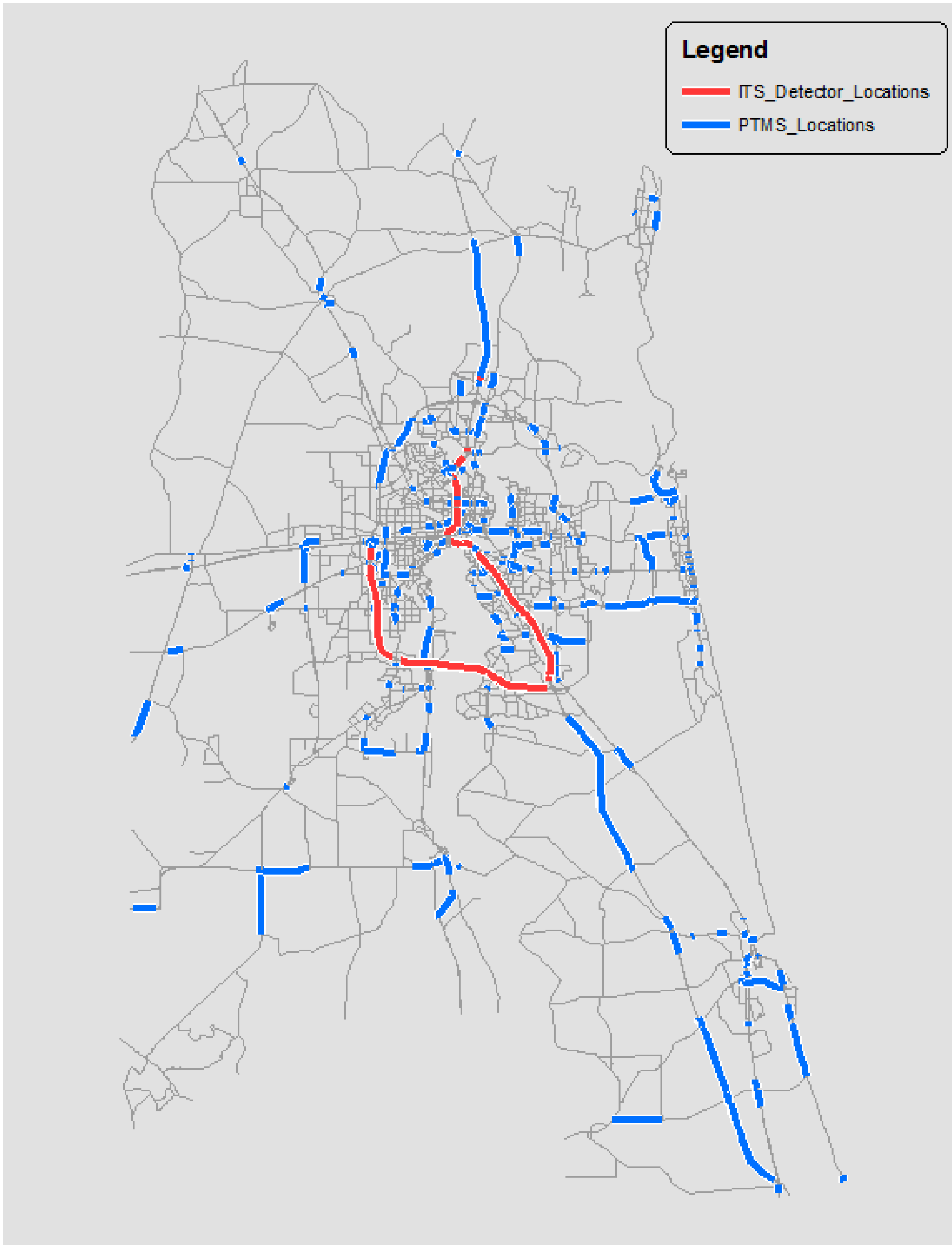
Table 72. Number of Locations for the PTMS, TTMS and ITS Data

Data Source	Native Format	Tagged to the Merged Network	Valid Tags	Transferred to the TRANSIMS Network
PTMS	923	923	897	562
TTMS	20	7	7	7
ITS	190	123	122	84

The TTMS data points were very limited in comparison to the other two sources. However, being a permanent traffic monitoring station, traffic counts throughout the year at the 15-minute resolution were available. This information can be used for studying the seasonal and special event traffic patterns at these select locations. Since the PTMS and ITS data provided the bulk of the validate data set, these two sources were relied upon heavily for comparisons and are presented in this chapter.

The locations of the PTMS and ITS data points for the Jacksonville region are shown in Figure 87.

Figure 87. Locations of PTMS and ITS Data shown on the PLANNING Network



3.2.2 Validation Results

Table 73 and Table 74 summarize the initial daily validation by facility type and area type. These tables demonstrate that overall daily volumes match relatively well, with estimated volumes approximately 3.4% higher than observed volumes and a regional %RMSE of 38.6. higher level facilities are generally overpredicted, with lower level facilities underpredicted. This is primarily attributable adjustments made to parameters affecting the circuity of routes during the TRANSIMS microsimulation calibration to better match highway volumes and speeds.

Table 73. Daily Validation by Facility Type

Facility Type	# Obs	Est Vehicles	Obs Vehicles	Diff	% Diff	% RMSE
Freeway	128	5,475,280	5,136,426	338,854	6.6	32.2
Expressway	30	929,697	767,990	161,707	21.1	35.1
Principal Arterial	52	580,435	576,392	4,043	0.7	26
Major Arterial	293	5,150,743	5,040,281	110,462	2.2	28.3
Minor Arterial	99	900,580	900,885	-305	0	48.8
Collector	18	140,057	123,639	16,418	13.3	46.6
Local Street	8	52,533	104,271	-51,738	-49.6	88.3
Ramp	80	492,855	615,063	-122,208	-19.9	72.7
External	2	3,819	5,287	-1,468	-27.8	28.3
Total	**710**	**13,725,999**	**13,270,234**	**455,765**	**3.4**	**38.6**

Table 74 indicates that the volumes in the denser regional core (area types 1 and 2) are generally overpredicted, with more suburban are rural areas slightly underpredicted.

Table 74. Daily Validation by Area Type

Area Type	# Obs	Est Vehicles	Obs Vehicles	Diff	% Diff	% RMSE
Area Type 1	45	665,453	594,015	71,438	12	32.2
Area Type 2	118	2,651,881	2,419,343	232,538	9.6	35.1
Area Type 3	451	9,004,366	8,775,450	228,916	2.6	26
Area Type 4	47	1,074,084	1,141,492	-67,408	-5.9	28.3
Area Type 5	49	330,215	339,934	-9,719	-2.9	48.8
Total	**710**	**13,725,999**	**13,270,234**	**455,765**	**3.4**	**38.6**

Estimated roadway volumes were also compared to observed volumes for four broad time periods. It should be noted that, unlike a traditional static assignment model, the TRANSIMS assignment process does not assign demand to the network by broad time period. Rather, the entire day demand of individual trips are loaded onto the network using minute-level departure time information provided by DaySim. However, time period summaries are still helpful in assessing the performance of the assignment model and informing adjustments to be made to both the DaySim demand and TRANSIMS supply models.

The time period summaries illustrate that the integrated model results for the AM, midday and PM periods look reasonably good both in terms of matching aggregate volumes by facility type and in terms of % RMSE. The evening time period looks more problematic, and will require additional investigation. To some extent this may be reflective of the "cascade effect." Due to the DaySim-TRANSIMS model preserving the integrity and linked nature of the individual trips on a tour across both the demand and assignment simulations, if it takes longer in the assignment to reach a given activity location given network travel times than expected when the demand was scheduled, the start time and the end time for that activity will be delayed, causing a cascade effect through the traveler's entire daily activity pattern, with trips being pushed later and later in the day. However, this effect is typically manifest in both the PM and evening periods, and in the current Jacksonville model the PM is actually underpredicted.

Table 75. AM Validation by Facility Type

Facility Type	# Obs	Est	Obs	Diff	% Diff	% RMSE
Freeway	128	930,670	986,919	-56,249	-5.7	34.7
Expressway	30	153,556	143,126	10,430	7.3	30
Principal Arterial	52	125,920	107,029	18,891	17.7	52.1
Major Arterial	293	999,174	898,274	100,900	11.2	35.1
Minor Arterial	99	187,351	177,146	10,205	5.8	59.4
Collector	18	23,813	22,902	911	4	36.5
Local Street	8	11,170	21,925	-10,755	-49.1	83.5
Ramp	80	105,651	122,742	-17,091	-13.9	84.8
External	2	831	750	81	10.8	33.1
TOTAL	**710**	**2,538,136**	**2,480,813**	**57,323**	**2.3**	**43.9**

Table 76. Midday Validation by Facility Type

Facility Type	# Obs	Est	Obs	Diff	% Diff	% RMSE
Freeway	127	1,483,157	1,405,075	78,082	5.6	2978
Expressway	30	288,789	204,735	84,054	41.1	3090
Principal Arterial	52	147,854	167,971	-20,117	-12	830
Major Arterial	293	1,455,803	1,548,916	-93,113	-6	1318
Minor Arterial	99	270,473	270,067	406	0.2	1016
Collector	18	39,580	36,260	3,320	9.2	641
Local Street	8	14,749	28,670	-13,921	-48.6	1862
Ramp	80	134,675	173,331	-38,656	-22.3	1132
External	2	1,073	1,890	-817	-43.2	409
Total	**709**	**3,836,153**	**3,836,915**	**-762**	**0**	**42.4**

Table 77. PM Validation by Facility Type

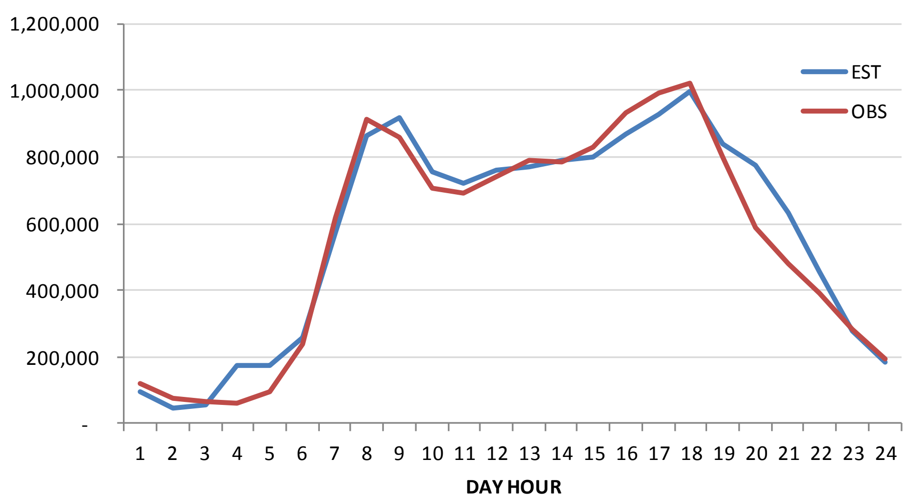
Facility Type	# Obs	Est	Obs	Diff	% Diff	% RMSE
Freeway	127	1,107,510	1,128,918	-21,408	-1.9	2029
Expressway	30	186,698	172,605	14,093	8.2	1325
Principal Arterial	52	114,778	127,018	-12,240	-9.6	525
Major Arterial	293	1,046,420	1,126,449	-80,029	-7.1	826
Minor Arterial	99	199,800	205,659	-5,859	-2.8	681
Collector	18	26,379	28,242	-1,863	-6.6	454
Local Street	8	10,398	23,008	-12,610	-54.8	1702
Ramp	80	97,834	134,869	-37,035	-27.5	741
External	2	847	1,207	-360	-29.8	180
Total	**709**	**2,790,664**	**2,947,975**	**-157,311**	**-5.3**	**35.3**

Table 78. Evening Validation by Facility Type

Facility Type	# Obs	Est	Obs	Diff	% Diff	% RMSE
Freeway	256	1,937,651	1,615,514	322,137	19.9	2542
Expressway	60	300,654	247,524	53,130	21.5	1556
Principal Arterial	104	191,883	174,374	17,509	10	613
Major Arterial	586	1,649,346	1,466,642	182,704	12.5	852
Minor Arterial	198	242,956	248,013	-5,057	-2	607
Collector	36	50,285	36,235	14,050	38.8	616
Local Street	16	16,216	30,668	-14,452	-47.1	1180
Ramp	160	154,695	184,121	-29,426	-16	617
External	4	1,068	1,440	-372	-25.8	166
Total	**1420**	**4,544,754**	**4,004,531**	**540,223**	**13.5**	**65.5**

Due to the continuous assignment of demand to the network across the entire day, it is possible to compare estimated to observed volumes using any temporal resolution. Figure 88 illustrates the estimated and observed volumes by hour. This chart clearly demonstrates that the integrated model is assigning more demand from 8pm (hour 20) through 10pm (hour 22) than observed.

Figure 88. Estimated and Observed Total Volumes by Hour



4.0 Model Sensitivity Testing

4.1 Purposes

Travel demand forecasting model systems are only able to test the effects of policies and assumptions which have been explicitly included when designing and implementing the model system, and are not intrinsically sensitive to the increasingly broad range of transportation policies and improvements of interest to decision-makers. While most regional models are sensitive to large-scale assumptions about land use and demographics, few are sensitive to more detailed assumptions about pricing policies, or to traffic or travel demand management strategies. Even where models have the capability to address these types of policies, they are typically not sufficiently sensitive to the dynamic interplay between travel behavior and network conditions by time-of-day, and are unable to reasonably represent the effects of road pricing, travel demand management, and other policies. A key goal of the SHRP2 C10A project is to make operational a dynamic integrated model–an integrated, advanced travel-demand model with a fine-grained, time-dependent network, and to demonstrate the model's performance through sensitivity tests and policy analyses.

Sensitivity testing of model systems involves the evaluation of the effects of changes in model inputs on model outputs. Although sensitivity testing of model systems can be performed in many ways, two approaches to sensitivity testing of travel demand forecast models are often employed. In the first method, the sensitivity of individual model components is evaluated by adjusting model inputs and documenting the effect on outputs. Elasticities are calculated and evaluated relative to established standards. In the second method, the focus of the sensitivity testing is on the overall model system. The focus of these efforts was on reporting the sensitivities of the model system.

4.2 Sensitivity Tests

A key motivating force behind the SHRP2 C10A project is the need to address transportation policies that are being considered in metropolitan planning organizations around the U.S. These policies are not adequately addressed by the current state of the practice travel forecasting models and so the integrated modeling tool developed for this project seeks to improve how these policies are addressed. In order to assess the increased sensitivity of the integrated model system, a set of tests were designed, implemented and evaluated. These tests were designed to illustrate the unique capabilities of the model system and included:

- Pricing: Pricing strategies are costs that are imposed on travelers using certain roads, traversing certain screenlines, or travelling to certain areas (tolling, cordon pricing or area pricing). These costs may be either fixed or vary by time-of-day or in response to congestion. Two types of pricing tests were evaluated as part of this effort. In the first, a number of scenarios were defined in which freeway tolls that varied by time of day. In the second, a number of scenarios were defined in which auto operating costs were modified from the "baseline" condition.
- Travel Demand Management: TDM approaches incorporate a wide range of strategies aimed at changing travel behavior to reduce congestion and improve mobility, such as increasing the frequency and numbers of people who work at home, adjusting work schedules to travel in off-peak, less congested conditions, or increasing the number of people who carpool to work. This sensitivity testing focused the impacts of a flexible work schedule in which workers worked fewer days but longer hours on those days. The overall time spent in work activities was held fixed.
- Operations: Operational strategies, also known as transportation system management (TSM) also address a wide range of projects and changes, including bottleneck improvements, corridor improvements and parking strategies. For this project, the sensitivity testing was focused on a scenario in which signals were coordinated along three primary regional corridors with the goal of reducing bottlenecks and improving the overall traffic flow.

The sensitivity tests documented here were performed using the Burlington implementation of the model system. Use of this smaller region allowed for more rapid testing and debugging of a greater number of scenarios. In order to ensure that there was sufficient congestion-related delay on the Burlington regional network (which does not have significant congestion), the socioeconomic inputs to the model system were scaled up by 50%. This increase in the population, employment, and all related inputs exceeded the forecast growth for 2030 in the Burlington region. Two notes about the charts and tables that illustrate the results of the sensitivity tests. First, both DaySim and TRANSIMS were used to generate summaries of travel demand and network performance measures, respectively. Second, many of the charts employ time-of-day along the X-axis in order to highlight one of the distinguishing features of the integrated model system: the exchange of information between DaySim and TRANSIMS by detailed time-of-day.

4.2.1 Pricing

4.2.1.1 Freeway Tolling

The first set of sensitivity test scenarios that were evaluated using the model system involved assessing the effects of freeway tolling by time-of-day. For these sensitivity tests, a set of 3 scenarios were evaluated and compared against the baseline alternative. These scenarios were based on pricing alternatives tested in a pricing experiment conducted by the Puget Sound Regional Council in Seattle to observe travel behavior and better understand regional pricing analyses. In the baseline alternative, no costs were assessed at any time. In the "Pricing_3" scenario, a fixed $0.25/ mile charge was assessed for anyone using the freeways during the peak periods. In the "Pricing_4" scenario, these fixed peak charges were maintained, while a fixed $0.10/freeway mile charge was added in the midday. Finally, in the "Pricing_5" scenario the fixed peak period charges were increased to $1.00/freeway mile and the fixed midday charges were increased to $0.50/freeway mile. When testing the sensitivity of the model system, it is often useful to test extreme cases or scenarios such as Pricing_5, even if they are unlikely to ever be implemented in reality.

Based on the structure and linkages of the DaySim and TRANSIMS models, one would expect that, in general, increases in tolls on facilities at certain times of day will result in overall increases in user costs (unless these tolls are optimized, which was not performed as part of these tests). These changes may be reflected in decreases in the overall level of activity generation thorough the upward feedback of aggregate logsum measures, although this effect would likely be small. We expect the effect on the distribution of travel demand by time-of-day to be more pronounced, with travelers choosing to reduce travel distances during the tolled time periods. The effect on mode choice would likely be small due to the relatively small transit services offered in Burlington, though one would expect that overall freeway volumes would decrease noticeably during the peak periods.

Figure 89 through Figure 91 illustrate that the expected changes were all observed in the model system outputs. Figure 89 shows the difference in total trips by time-of-day relative to the baseline alternative. Pricing_3 shows declines in travel during the AM and PM peak when freeway tolling is in place, but little change during the midday, which is un-tolled. Pricing_4 and Pricing_5 both show declines in travel during all tolled time periods, with higher tolls resulting in more reductions in travel. Interestingly, all three pricing scenarios show pronounced increases in travel demand during the evenings, suggesting that travelers are rescheduling activities to occur when there are no tolls, as well as fewer scheduling constraints such as are present during the midday. Figure 90 illustrates that the tripmaking by time-of-day affects different purposes differently with the work purpose distributions (blue) relatively unaffected, but the social/recreation distributions (red) shifting noticeably out of the peaks and into the evening. Finally, Figure 91 illustrates that the network-based total delay is higher than the base in all scenarios, as the tolling induces travelers to shift onto more capacity constrained surface facilities. An analysis by facility type indicated that most of this additional delay accumulated on minor arterials. This is likely due to coarseness of the sensitivity test - the specific temporal and spatial extent of congestion on the freeway system did not inform the design of the tolling scheme. While some peak location congestion was likely alleviated as a result of the tolling, in many locations across the broad tolling time periods, increased costs were not offset by travel time reductions due to the low levels of baseline congestion.

Figure 89. Difference in Trips by Time of Day by Freeway Tolling Scenario

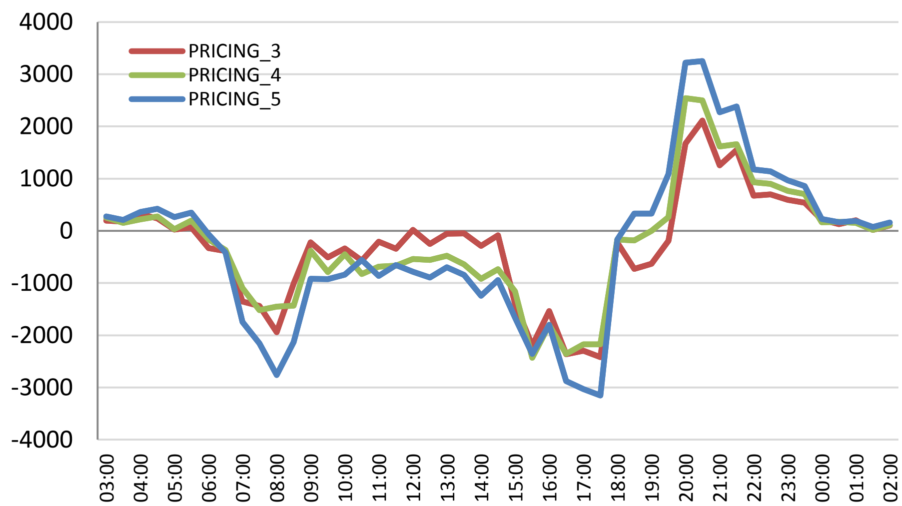


Figure 90. Work and Social/Recreation Trips by Time of Day by Freeway Tolling Scenario

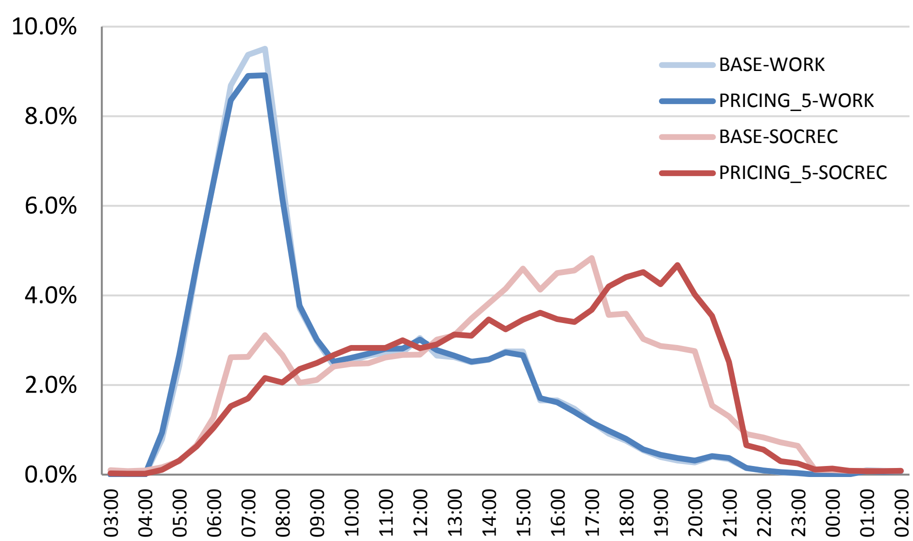
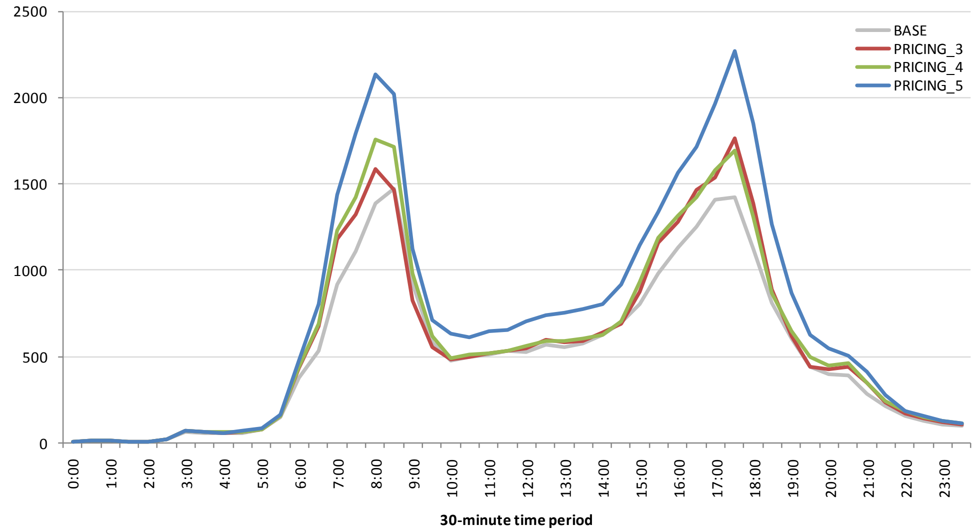


Figure 91. Total Delay (hours) by Freeway Tolling Scenario



4.2.1.2 Auto Operating Costs

The second set of sensitivity test scenarios that were evaluated using the model system involved assessing the effects of changing auto operating costs. Auto operating costs are represented in the model system as an average cost per mile experienced by travelers. These costs are reflected in both the DaySim demand model though the inclusion of these monetary costs in the utility specifications of a number of component models, and in the TRANSIMS model though the inclusion of these costs in the generalized costs used for pathbuilding.

For these sensitivity tests, a set of 3 scenarios were evaluated and compared against the baseline alternative. The baseline alternative assumes a cost of $0.12/mile. A lower cost scenario of $0.06/mile was tested as well as two higher costs scenarios of $0.24/mile and $0.60/mile.

Based on the structure and linkages of the DaySim and TRANSIMS models, one would expect that, in general, increases in auto operating costs might be manifest in overall lower levels of auto ownership, which is a “long term” choice that exists towards the top of the DaySim model system and influences subsequent decision-making. Only marginal effects on activity generation and time-of-day due to higher auto costs are likely, although one might also expect to observe more pronounced effects on overall trip distances (shorter) and on mode choices (more transit).

Table 79 and Table 80 confirm the expected effects on auto ownership and overall activity generation. Table 79 shows that when auto operating costs decline (AOC_X05) the share of households choosing to maintain 0 vehicles also declines, and as the costs increase, the share of 0-vehicle households also increases. Table 80 illustrates changes in regional tour-making by purpose, and demonstrates that lower costs slightly increase tour-making for discretionary purposes while higher costs slightly decrease tour-making by purpose. However, the reductions in tour-making by purpose are not consistent across purposes, with mandatory work and school tours declining while discretionary purposes such as personal business and social recreational purposes are seemingly unaffected. By this measure, personal business and social-recreational trips are less discretionary than meals, escorting passengers and shopping.

Table 79. Household Auto Ownership Shares by Auto Operating Cost Scenario

	BASE	AOC_X05	AOC_X2	AOC_X5
0	5.5%	5.3%	5.7%	6.5%
1	36.7%	36.8%	36.7%	36.3%
2	39.5%	39.6%	39.5%	39.3%
3	13.1%	13.1%	13.0%	12.9%
4+	5.2%	5.2%	5.2%	5.0%

Table 80. Tours by Purpose by Auto Operating Cost Scenario

	BASE	AOC_X05	AOC_X2	AOC_X5	AOC_X05	AOC_X2	AOC_X5
Work	116,928	117,475	115,898	114,321	1.00	0.99	0.99
School	44,011	44,246	43,449	41,906	1.01	0.98	0.96
Escort	41,982	42,022	41,610	41,028	1.00	0.99	0.99
PersBus	45,877	45,756	45,549	45,635	1.00	1.00	1.00
Shop	38,841	39,432	38,210	37,525	1.02	0.97	0.98
Meal	15,908	16,001	16,021	15,794	1.01	1.00	0.99
SocRec	47,181	47,436	46,999	47,076	1.01	0.99	1.00
Total	350,728	352,368	347,736	343,285	1.00	0.99	0.99

Consistent with expectations, Figure 92 indicates that there is little systematic difference in changes in trips by time of day across the three auto operating cost scenarios, while Figure 93 illustrates a slight increase in shorter trips and decrease in longer trips associated with the highest assumed auto operating costs. Figure 94 shows per capita changes in VMT associated with the highest assumed auto operating costs relative to the base, indicating that the most pronounced decreases in VMT are associated with areas located at the periphery of the region, which is reflective both of the overall higher levels of baseline VMT in these areas, and may possibly reflect boundary effects as well, in which residents at the edges of the modeled region may be forced to travel further in order to implement their daily activity patterns. In addition, this figure demonstrates the parcel-level spatial resolution used in DaySim. The pattern of per capita VMT increases and decreases may illustrate the, effect of the Monte Carlo simulation method used in the simulation. A unique feature, as well as limitation, of the spatially and temporally disaggregate model system is that simulation sampling methods used in the conjunction with static user equilibrium network assignment are not as easily employed.

Figure 92. Difference in Trips by Time-of-Day by Auto Operating Cost Scenario

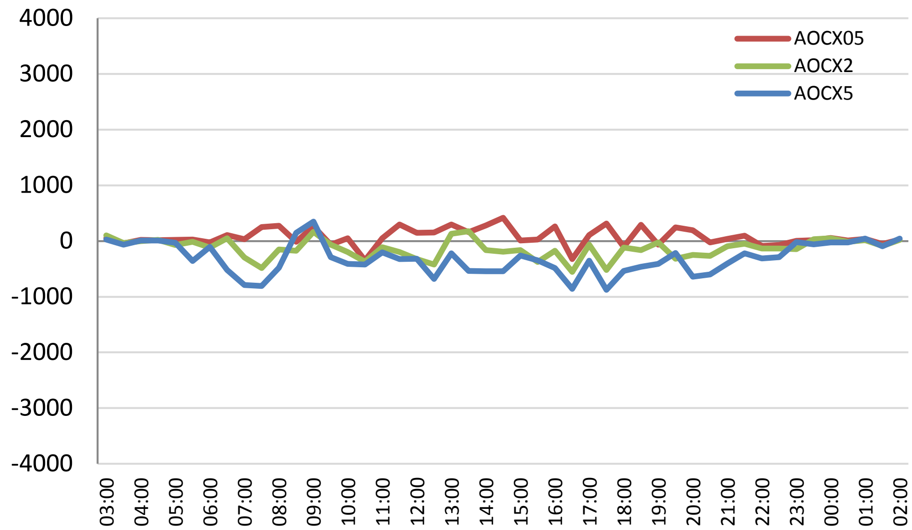


Figure 93. Trip Length Frequency Distribution by Auto Operating Cost Scenario

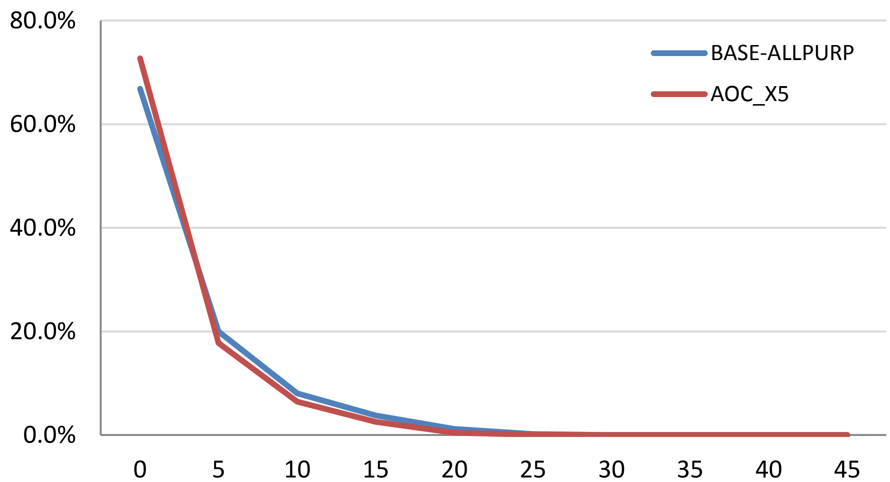
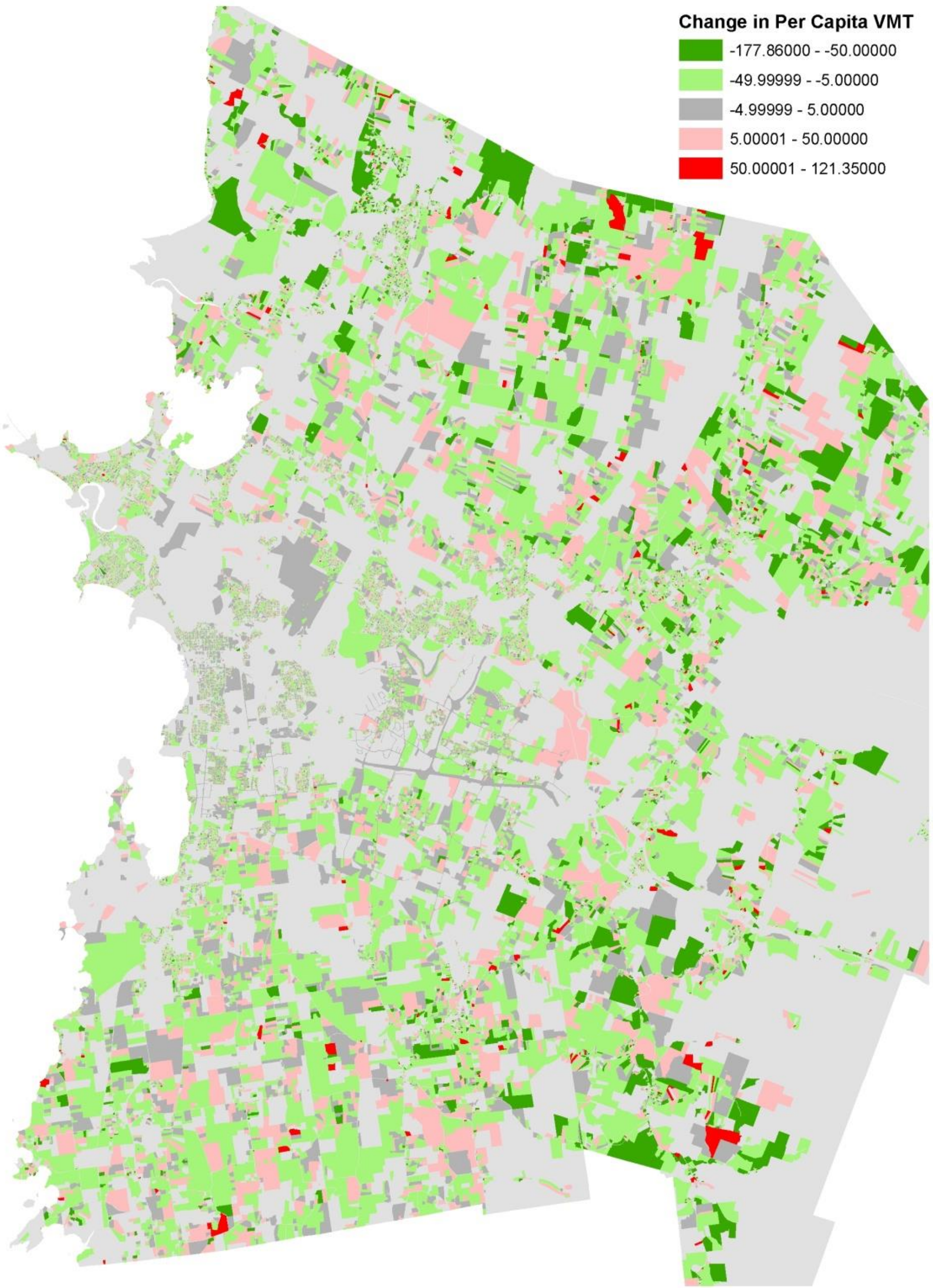


Figure 94. Per Capita Changes in VMT between Baseline and x5 Auto operating Cost Scenario



4.2.2 Travel Demand Management

The third set of sensitivity test scenarios that were evaluated using the model system involved assessing the effects of a travel demand management (TDM) strategy. TDM approaches are intended to change travel behavior to reduce congestion and improve mobility, such as increasing the frequency and numbers of people who work at home, and adjusting work schedules to travel in off-peak, less congested conditions. Because DaySim predicts the daily activity pattern of each individual in the region, it can be used to reflect the effect of workers working fewer days but longer hours. However, this sensitivity is purely "scenario-based." DaySim cannot identify which policies will be most effective at affecting flexible work schedules, though it can estimate the impact on individual travelers' activity patterns and schedules and on the overall transportation system performance assuming that an effective policy is in place. In order to represent this effective policy, model parameters influencing the work tour and trip generation as well as work durations were modified in order to represent a shift to working fewer days but more hours, holding the total aggregate time in work activities constant.

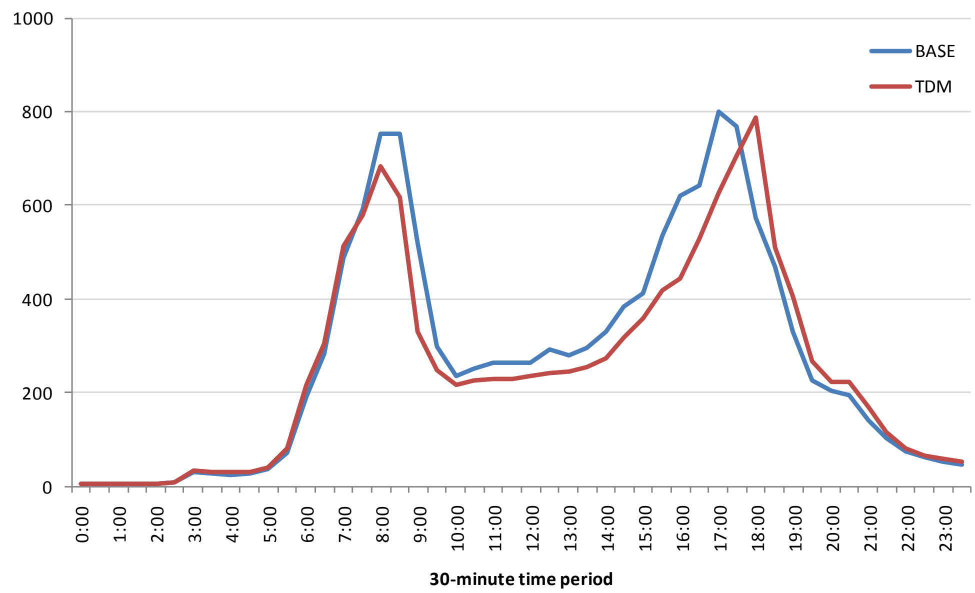
For this sensitivity test, a single scenario was evaluated where workers shifted from a 5-day ~7.5 hour workweek to a 4-day ~9 hour workweek. Based on the structure and linkages of the DaySim and TRANSIMS models, one would expect that, in general, overall levels of activity generation are lower, although the declines in work-related travel may be offset by increases in travel for discretionary purposes. Clearly, shifts in the distribution of travel by time-of-day due to the lengthened workday should be expected. Changes in the destination and mode choices would likely be marginal, though the time-of-day changes should be manifest in volumes by time-of-day on the roadway network.

Note that for this sensitivity test the adjusted work activity duration distribution represents an analyst's qualitative judgment about a potential distribution, which should ideally be informed by more empirical analysis of observed changes in work tour durations. Table 81 demonstrates the impact on the tour patterns of fulltime workers, illustrating that as work tours decline, fulltime workers tend to make more personal business, social/recreation and shop tours. Changes in the travel by time of day are evident in summaries of the DaySim travel demand model outputs, and are also manifest in summaries of network performance by time-of-day. For example, Figure 95 shows a reduction in hours of delay on major arterials associated with implementation of an effective alternative work schedule policy, and this reduction occurs across all types of facilities throughout the region.

Table 81. Fulltime Worker Tours by Purpose by TDM Scenario

	Original	Adjusted	Adj/Orig
Work	94,408	78,472	0.83
School	115	140	1.22
Escort	8,070	9,023	1.12
Pers Bus	13,519	16,848	1.25
Shop	10,531	12,938	1.23
Meal	3,817	3,842	1.01
Soc/Rec	13,076	14,360	1.10
Workbased	27,949	23,211	0.83
Total	171,485	158,834	0.93

Figure 95. Hours of Delay on Major Arterials by TDM Scenario



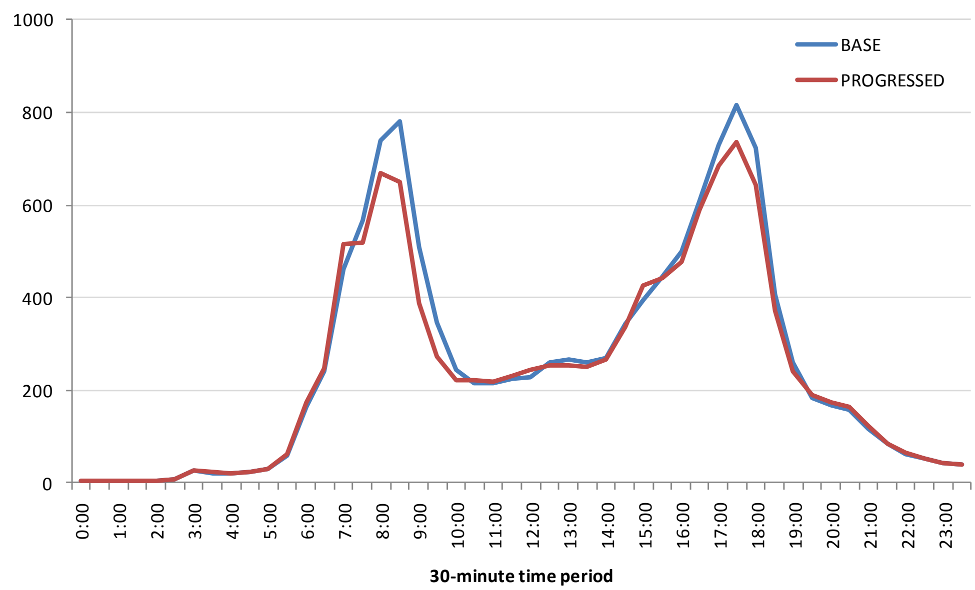
4.2.3 Operations

4.2.3.1 Signal Progression

Operational strategies, also known as transportation system management (TSM) can address a wide range of projects and changes, including bottleneck improvements, corridor improvements and parking strategies. For this project, the sensitivity testing was focused on a signal progression scenario in which signals were coordinated along three primary regional corridors with the goal of reducing bottlenecks and improving the overall traffic flow. The DaySim-TRANSIMS model system provides sensitivity to these improvements, which traditional travel demand forecast models cannot typically represent the due to their linkage with traditional static network assignment methods that lack detailed network operation attributes, as well as their coarse temporal resolution. Due to the limited geographic extent of these improvements we would expect to see very little change in any of the aggregate regional statistics measuring activity generation, time-of-day, or destination and mode choice. However, one would expect to see local level changes that reflect improved speeds along the targeted corridors.

The initial model results showed some reductions in delay by facility type, particularly during the peak periods as show in Figure 96 for major arterials. However, a closer inspection of the speed profiles along the three targeted corridor showed more mixed results, with the signal progression producing better speeds in some corridor directions, and worse speeds in other corridor directions. In implementing these operational tests, a significant amount of iteration was required in order to establish a set of baseline signal timings for the entire region that produced reasonable performance results. Prior to establishing new baseline timings, all the corridor progression changes resulted in the significant simulation problems. This is likely a reflection of the fact that these assumptions are necessarily more detailed in geographic and temporal scope, and ultimately more consequential for the performance of the network simulation. As has been noted by others, the sensitivity of DTA and traffic microsimulation models to these detailed inputs suggest distinct challenges when attempting to incorporate these assumptions in a forecasting mode, especially if at a regional scale. Of all the scenarios evaluated as part of this sensitivity testing, the signal progression scenario required the greatest amount of time, and resulted in the least interpretable results.

Figure 96. Hours of Delay on Major Arterials by Operations Scenario



5.0 CONCLUSIONS

5.1 Model Implementation

5.1.1 Demand Model Data Development

The proposed model system is comprised of three primary components: DaySim, the TRANSIMS Router and Microsimulator, and MOVES. DaySim is a travel demand forecast model that predicts household and person travel choices at a parcel-level on a minute-by-minute basis. The TRANSIMS Router and Microsimulator are dynamic traffic assignment and network simulation software that can perform regional traffic microsimulation on a second-by-second basis. MOVES is the EPA's latest software for estimating emissions and air quality impacts. The integrated model links these model components in an equilibrated model system that provides enhanced policy sensitivities at significantly higher levels of spatial and temporal resolution than found in a traditional regional travel demand forecasting system.

As part of the C10A project, the DaySim-TRANSIMS-MOVES model system was implemented in two regions: Jacksonville, Florida and Burlington, Vermont. The Jacksonville region is comprised of four counties in northeast Florida covering 3,100 square miles. The regional population includes over 525,000 households and 1.25 million people, and generates more than 4 million daily person trips. . The Burlington modeling area is comprised of a single county (Chittenden) of approximately 620 square miles, and is home to 55,000 households in the base year of 2005.

There are a number of basic inputs to the activity-based model (ABM) component, including parcel-level information on the location of employment using nine industrial sectors, housing units, enrollment by school level, parking cost and supply information, access to transit, and numerous "buffer"-based measures of accessibility. A synthetic population representing the detailed socio-demographic attributes of the regions residents, and network performance data (or "skims") are also key ABM inputs.

5.1.1.1 Parcel data

Use of parcel data significantly enhances the sensitivity of the model system to smaller-scale land use and urban form measures and provides better sensitivity to travel modes such as bicycle and walk, which are usually shorter and more greatly influenced by attributes of the surrounding physical environment; however, developing a robust parcel-level base-year data file can be time consuming. Often it is challenging to establish the baseline geographic information on parcel boundaries, and to ensure that data fields are defined and coded consistently across multiple political entities such as counties. Examples of the types of geographic data processing that were required to develop the Jacksonville parcel information included identifying and removing non-travel generating parcels such as highway rights of way and bodies of water, merging together into a single parcel any parcels that are the same parcel but are spread across separate GIS shapes with multiple database rows, and developing a set of common fields across the multiple county databases.

Housing unit data

The model system uses parcel-level housing unit information to allocate the model's synthetic population (described in section 2.3), which is developed at the travel analysis zone (TAZ) level to incorporate sociodemographic controls that are available only at more spatially aggregate levels, to specific parcels. Inconsistencies between parcel-level information, Census block information and the TAZ-level information used by agencies arise frequently and must be rectified. The parcel databases provided by the Jacksonville region counties did not consistently specify the number of housing units on a parcel, so it was necessary to impute this information based on the using a combination of parcel data sources in conjunction with Census-based housing unit information. Multi-family housing presents the largest challenge when imputing housing units, because they are prone to being under-counted in tax assessor databases.

Employment data

Employment data is not usually available at a parcel-level and must instead be developed using address-based business establishment employment information. Unfortunately, there is no single comprehensive and "clean" public source for this information. Instead it is often necessary to rely on private data vendors, but these data often have limitation such as omitted information such as public sector employment, absence of self-employment, and other deficiencies such as the "headquarters effect" in which the employment across all of a firm's locations is allocated to a single headquarters location. In Jacksonville, a datafile of employment locations that had been extensively reviewed and cleaned was provided by the North Florida Transportation Planning Organization NFTPO, the region's MPO, which greatly facilitated the rapid implementation of the model system.

Enrollment data

School location-level information on enrollment is necessary to order to ensure that the proper number of students are being attracted to each school location. School enrollment data by parcel and grade level is typically easier to develop, as most state departments of education can provide enrollment by school address and grade; however, a fair amount of manual effort is required to geocode the enrollment locations as well as to gather enrollment information on trade and other professional development schools.

5.1.1.2 Parking data

Offstreet parking location and pricing information is used in the activity-based models system to influence mode and other choices. Developing parking capacity and cost data can also be challenging if public agencies cannot provide, although this information is used in the model less than the housing, employment, and enrollment data. In Jacksonville, enrollment and parking data were also easily acquired from existing sources.

5.1.1.3 Transit stop data

In addition to using zone-level information on access times to transit, DaySim also incorporates detailed information on the distance to transit by transit sub-mode. This information is based on all transit stop locations in the region (as opposed to a file of only transit stops coded in the model networks). The file of all transit stops was provided by the local transit agency and required virtually no additional processing.

5.1.1.4 Intersection data

Unique measures of urban form that DaySim incorporates are the number of intersections or nodes of different types within ¼ mile and ½ mile buffers. These intersection types include, dead-ends (1 link), T-intersections (3-links), and tradition intersections (4+ links), and help characterize the pattern of urban development. These measures are derived from an "all streets" network that has been updated to include the number of links associated with each intersection. Preparation of this intersection data required some cleaning and manipulation in GIS software.

5.1.1.5 Parcel data preparation tools

A utility program was developed to generate the specific parcel input file to Daysim. This program combines information about parcel-level housing, enrollment, and employment and calculates additional measures, such as buffer variables that describe the amount of housing, enrollment and employment, and intersections by type within user-specified radii of each parcel. The utility also computes the distance from the parcel to the nearest transit stop.. Ultimately, the utility tool greatly simplifies and expedites the preparation of the derived measures described above, and generates the primary DaySim parcel input file. The inputs to this tool are all straightforward and well-documented, and can be ascii text files or other user-specified formats. The sizes of the parcel file can vary significantly based on the regions size, which may affect the data development schedule. If an agency can provide complete and "cleaned" datasets that are used as input to the Daysim parcel data preparation tool then the parcel data development process could be prepared in a matter of weeks. More typically, complete and "cleaned" data is not available and it may take six months or more to develop the required inputs.

5.1.1.6 Future / alternative parcel data

If the model is to be used to support long range forecasting efforts, then it is also necessary to develop future parcel-level assumptions. Ideally, such assumptions would be derived from a land use forecasting model, although rarely are such models available. In the absence of such a tool, most agencies have either applied tools or methods that split parcels and then populated this new parcel geography with updated data, or simply scaled up base year parcel level data to match regional or sub-regional employment and housing controls. For the C10A effort, this latter approach was used to develop "increased demand" scenarios.

It should be noted that not all activity-based models use detailed parcels as a fundamental spatial unit, and in fact the Daysim model can be applied at more aggregate spatial resolutions such as census blocks or travel analysis zones (TAZs). Use of these larger spatial units typically reduces the data preparation burden significantly, though this also limits the sensitivity of the model system.

5.1.1.7 Synthetic population

In addition to the parcel-level inputs to the model system, it is also necessary to develop a "synthetic population" of these regions' residents. This synthetic population is comprised of lists of households and persons that are based on observed or forecast distributions of household-level and person-level socioeconomic attributes and created by sampling detailed Census microdata. These lists function as the basis for all subsequent choice-making simulated in the model system. All base year 2005 data required to develop the synthetic population using the DaySim population generation component were available from the Census (American Community Survey Public Use Microdata Sample (PUMS) and Decennial PUMS), Northeast Florida Regional Planning Model (NERPM) model inputs, and the National Household Travel Survey (NHTS).

The PopGen software, developed by Arizona State University, was used to create the synthetic population, consistent with socio-demographic controls prepared by the project team. The Jacksonville synthetic sample population is comprised of three segments: permanent households and population, seasonal households and population, and group quarters population. These segments were established to reflect the differences in travel patterns associated with these sub-populations, to provide the ability to support seasonal analyses, and to be consistent with the specification of the DaySim model. For permanent and seasonal households, these socioeconomic attributes that were controlled for included:

- The age of the head of household
- Household size
- Household workers
- Household income
- Presence of children

For permanent and seasonal persons, these categories include:

- Gender
- Age

The sociodemographic controls were much simpler for the non-institutionalized group quarters population because of limited census table data for these residents. However, a person-level age distribution was used to properly locate two important GQ subpopulations: college students and retirement center residents.

Finally, a utility program was developed to allocate the synthetic households and persons to individual parcels. Only minor difficulties were encountered when developing and validating the synthetic population, which were easily resolved with changes to the PopGen configuration.

5.1.1.8 Auxiliary demand

"Auxiliary demand" refers to the regional demand that is not forecast by the DaySim model system, but that must be represented in the DaySim-TRANSIMS model system in order to reasonably assess network performance and the impacts of different policies or improvements. In the current project, this auxiliary demand is derived from the existing NFTPO model system. However, it is necessary to add spatial and temporal detail to this demand in order to integrate it with the more detailed demand and supply simulation models.

Types of auxiliary demand

DaySim provides detailed estimates of the long-term and short-term travel choices of Jacksonville residents when traveling within the region, but this travel demand does not fully represent all trips that use the regional transportation networks. Commercial and truck traffic comprise a significant share of all roadway volumes, typically up to 20% or more. In addition, non-residents enter the region through key external gateways to access jobs, shopping or other opportunities, or may simply pass through the region. Similarly, residents may leave the region to satisfy other needs. Special generators may also create demand not explicitly represented by person travel demand models.

5.1.1.9 Auxiliary demand processing

Auxiliary demand is generated through processes exogenous to the DaySim-TRANSIMS model system. The total demand and the spatial distribution, mode, and timing of these trips are fixed within a given forecast or horizon year (but will vary across model run years). Network times and costs influence the routes used, however, so the assignment of this auxiliary demand to network paths is not fixed. It was relatively straightforward to identify and process the data required to perform the temporal and spatial disaggregation of these demand segments. The spatial disaggregation was primarily driven by the parcel-level land use information prepared for input to Daysim, white the temporal disaggregation was informed by a variety of data sources, such as observed traffic volumes at external stations, traffic volumes by vehicle class, and airport departure and arrival schedules. Ideally, the auxiliary demand models would be revised to incorporate sensitivity to the more temporally detailed network performance measures generated by the DaySim-TRANSIMS model system

5.1.1.10 Conclusions

Developing the parcel-level inputs to the ABM was relatively straightforward. The cleaning of the employment data by NFTPO and FDOT significantly reduced the amount of time required to implement the model, although it was still necessary to make relatively crude updates to the employment data in one of the counties. The parcel file required some additional cleaning to establish reasonable totals of housing units and to address inconsistencies in the parcel geography. School enrollment, transit stops, intersection types, and parking data were all relatively easy to assemble from existing data sources. In addition, developing the synthetic population was relatively straightforward given the availability of the data and tools; however, the overall effort still required approximately six months.

Accommodating auxiliary demand within the integrated DaySim-TRANSIMS model system was achieved using readily available static methods from the region's trip-based model; however, revisions to these auxiliary demand components are necessary for a more spatially and temporally consistent integrated demand-supply model system. A drawback of the current implementation is that the auxiliary demand is currently "fixed" for each forecast year. That is, although this demand varies by forecast year, it is not affected by changes in network impedances. Ideally, the auxiliary models would be revised to provide sensitivity to changes in network performance.

5.1.2 Network Model Data Development

In order for the TRANSIMS software to generate the detailed representations of changes in network performance by time of day, it is necessary to develop a representation of the network supply, to identify means of configuring the network assignment process to generate stable simulations with reasonable runtimes, and to establish methods for generating network performance indicators that can be used by the DaySim demand model.

5.1.2.1 Network build tools

As with most medium and large size MPOs, NFTPO maintains a detailed GIS-based "all streets" network that is used as the basis for building the networks used in the region's traditional trip-based travel demand model system. This all streets network was used in conjunction with a number of TRANSIMS network processing tools to synthesize a TRANSIMS network. The TRANSIMS tools provide a quick method of developing a detailed TRANSIMS network without a lot of extra data collection and arduous network coding. The tools reformat, regroup and reconfigure that data into standard TRANSIMS input link and node data files, and then synthesizes the additional information needed for a network simulation such as pocket lanes, lane connectivity, parking lots, activity locations and signal and sign warrants. Using these tools allowed the project team to get the network model running relatively quickly and uses the trip assignment process to identify locations where the synthetic process requires refinement.

5.1.2.2 "All Streets" network

One of the key questions that was of interest to the project team concerned the potential benefits of incorporating more spatial detail into the simulation networks. The networks used in most traditional static traffic assignments are relatively coarse, frequently including only minor and major arterials and freeways. Because static assignment models and networks are insensitive to many operational attributes such as hard capacity constraints, there are rarely concerns about the mismatch between demand and supply. In the context of TRANSIMS, however, consideration was given to the tradeoffs between using a coarser network which would run faster, but may provide the same degree of network sensitivity. As a result, for Jacksonville three different network resolutions were created. The coarsest network developed was the "planning network." This network essentially contains all the same links as found in the MPO's current 4-step regional modeling network, although significant operation detail was included in the TRANSIMS version. The next most detailed version of the network was the "Finegrained Network," which pivoted off the Planning Network but also included additional local through streets to the network. Finally, the most detailed network was the "Allstreets

Network," which was equivalent to the NERPM regional modeling network plus all other existing minor streets such as neighborhood streets and alleys.

5.1.2.3 Network cleaning

After the networks were built, a significant amount of additional time and effort were required in order to correct and debug network coding errors and inconsistencies. At the most basic level, it was necessary to fix network topology issues. For example, in numerous locations freeways were found to be erroneously intersecting with surface streets. These coding errors would not have been an issue had some of these surface streets been dropped when the baseline network was used to create coarser networks for static assignment. In addition to these basic topological corrections, it was also necessary to devote significantly more attention to other network attributes, because the simulation networks are much more sensitive to coding assumptions than traditional static assignment networks. Specifically, discontinuities in the coding of facility types, through lanes, and speeds all significantly affected the network performance.

Overall, speed issues were most common, but lane coding discontinuities proved to be most challenging in debugging the networks. Discontinuities refer to refer to inconsistent coding of link attributes for adjacent links. Because lanes are the primary source of capacity information within the simulation and lane changing is one of the primary reasons for congestion and lost vehicles, lane discontinuity errors are extremely problematic. Lost vehicles refers to vehicles that are unable to complete their assignment paths and get "stuck" in the simulation for more than a user-specified number of minutes. Lost vehicles are highly undesirable because they cause inconsistencies between the demand and supply components. It was also essential to establish on a link-by-link basis whether the lane coding from the master network included parking lanes or turn lanes. Other network coding issues that arose in developing all of the different resolution networks included extremely short links which can cause congestion problems in the simulation, and intersection geometry can impact simulation performance.

5.1.2.4 Intersection controls

In addition to getting the basic network geometric details correct, it is also necessary to get all the intersection control information correct, such as the location, timing and phasing of signals, and the locations of other intersection controls such as stop signs. Regarding signals, for Jacksonville the project team received a dataset of the locations of all signals in the region. Unfortunately, these locations were not linked to the master network data file, and did not containing any information on timing. Information on stop sign locations was available only for a subset of the region.

Ideally, "observed" real world signal timing, phasing and other control information would be available in a readily usable form and coded in the base network. However, developing this real world information to code this into the network is onerous. For this project, TRANSIMS tools were used to synthesize the timing and phasing of signalized intersections. TRANSIMS includes a number of ways of changing the configuration of the network by time of day. In addition to the roadway configurations, traffic controls can also vary by time of day. The signal timing and phasing plans can be adjusted to optimize time-of-day flow conditions. Signal progression tools are also available to coordinate fixed time signals along specified corridors or throughout a grid system. Demand-actuated signals can include multiple detectors and simulate ramp metering behavior. The signal formats also allow changing signal types by time of day.

For unsignalized intersections, it was also necessary to synthesize traffic controls for the base year. It should be noted that even if real world information is available for developing the base networks, the challenge of developing future year intersection controls under increased demand will remain. Use of some tools and optimization strategies presents not only technical challenges, but also poses theoretical and procedural challenges for how to best prepare and analyze future-year networks and how to conduct alternatives analyses.

Finally, it should also be noted that after the initial model network build process, all subsequent alternative networks were derived from the base network, and manipulated using TRANSIMS tools. Recent updates to these tools have significantly reduced the level of burden upon users to simultaneously and consistently update multiple input network files.

5.1.2.5 Conclusions

Developing detailed and usable networks for microsimulation requires a significant level of effort. The TRANSIMS software comes with a wide array of tools to perform many network development tasks, and spatially detailed network data are widely available; however, users should expect to spend on the order of hundreds of hours debugging simulation networks: correcting topological errors, resolving attribute discontinuities, and coding intersection controls. The time-consuming effort involves iteratively evaluating, adjusting and testing the networks by running simulations. In addition, users face numerous challenges when attempting to develop future year or alternative network scenarios, a topic which is discussed in a subsequent section of this document.

5.1.3 Model Integration

The design of the integration scheme focused on a few key questions:

- What information does the DaySim demand model need to provide to the TRANSIMS network supply model?
- What information does TRANSIMS need to provide to DaySim?
- How is the network supply model iteratively executed to achieve an equilibrated, or at least stable, condition?
- How are DaySim and TRANSIMS iterative executed to achieve a stable system solution?

Answering the first two design questions concerning the information exchange is relatively straightforward. Answering the second two questions regarding the execution and interaction of the tools is significantly more complex, forcing users to develop new measures to analyze network and system performance, prompting challenging questions about the nature of equilibration in the context of these complex simulation tools, and highlighting practical issues of runtime. In addition to addressing these implementation and research questions, a key goal of the SHRP2-C10A project is that the integrated demand-supply model is implemented in a dynamic modeling framework that is easily transferable to the local jurisdictions for policy analysis. In support of this goal of transferability, the model system incorporates a system manager that controls the execution of the two primary model system components: DaySim and TRANSIMS.

5.1.3.1 Information exchange

DaySim provides trip and vehicle information to the TRANSIMS Router to perform network assignment. For the Jacksonville region, this process takes approximately an hour of computer processing time. The resulting activity file includes the internal travel demand generated by regional households. This demand is combined with the auxiliary trips to represent the complete travel demand for the region.

Trip lists

In order to transmit these estimates of demand from DaySim to TRANSIMS, minor adjustments were made to the DaySim outputs to generate vehicle-trip records, and to associate the parcel locations with the activity locations used by TRANSIMS and output the records in the format required by the Router. In addition, detailed traveler and purpose specific information on trip value-of-time is included for use by TRANSIMS. Some of the key advantages of using TRANSIMS with Daysim, rather than using a traditional static assignment model, are that trips are kept and simulated in list format rather than aggregating to O-D matrices; less spatial aggregation occurs because "activity locations" used in TRANSIMS are smaller than travel analysis zones (TAZs); and because trip start and end times can be kept in units of individual minutes rather than aggregating to broad time periods.

Network impedances

DaySim's core components use a variety of impedance measures to influence traveler's choices about daily activity patterns, destination choices, mode choices, and time of day choices. These measures are produced by TRANSIMS to provide this information back to DaySim. The TRANSIMS supply side model assigns the DaySim internal demand and auxiliary trips on the TRANSIMS network through assignment iterations designed to achieve dynamic user equilibrium convergence of the individual travel paths. The resulting network performance by time of day is used to generate zone-to-zone travel time, distance, and cost skims for detailed periods for input into the next global iteration. In the current model implementation, this network performance information is created for 22 different time periods (as small as half hours in the AM and PM peaks) and for all zone pairs in the model prior to running the Daysim demand component. TRANSIMS includes tools that can flexibly generate network impedance skims at virtually any temporal or spatial resolution with little additional runtime, because the daily microsimulation of the entire region provides information on changes in network performance by specific times of day. The current solution, which involves iteratively using new paths for all travelers, based on composite travel times across multiple assignment iterations, is effective when the number of time periods and zones is relatively limited. However, it would be ideal to generate network performance indicators using temporal and spatial resolutions that are consistent with the underlying parcel-level spatial resolution and30-minute temporal resolution of Daysim demand model. A number of schemes to achieve this level of detail were hypothesized, such as implementing efficient multi-stage sampling of destinations (and corresponding impedances) at strategic points in the DaySim looping process, and tightly integrating DaySim and TRANSIMS so that DaySim can call TRANSIMS to extract the required measures quickly. Due to runtime considerations, and project resource constraints, none of these methods were ultimately implemented.

5.1.3.2 Component linkages and execution

A critical aspect of integrated model development involved implementing and refining strategies for achieving a condition of model system convergence in the network assignment process, as well as for the overall integrated model system. Convergence is necessary in order to ensure the behavioral integrity of the model system.

Network assignment

At the network assignment level, convergence is achieved through the interaction of the TRANSIMS Router and TRANSIMS Microsimulator. The project team extensively tested a variety of configurations of the Router, Microsimulator and other TRANSIMS tools used in the network assignment process. These included variations on rebuilding paths for all travelers at all iterations, rebuilding paths for both random and targeted subsets of travelers at each iteration, and different methods of calculating link delays and using different levels of temporal resolution when representing changes in network performance by time of day. Interestingly, the methods that seemed to perform best, as measured by traditional link-based relative gap, new trip-based gap, and stability measures such as aggregate vehicle hours and miles travelled, were amongst those that are least accepted by the dynamic traffic assignment (DTA) community. The more commonly practiced network assignment methodologies, such as the use of updated paths for a successively smaller share of travelers across multiple assignment iterations, were often the worst performing. Interestingly, the observed TRANSIMS convergence using these methodologies was relatively consistent with the reported convergence of other dynamic traffic assignment (DTA) tools. This performance is critical, as the project team discovered that the level of convergence can significantly influence the conclusions drawn from alternative analyses.

Network convergence

Another key network assignment convergence issue is how to measure network convergence, and how many iterations are required in order to achieve a converged result. In terms of measuring convergence, the project team developed both typical measures such as network link-based relative gap measures, as well as new measures such as person trip-based gap measures. These measures were calculated across the entire day, as well as by detailed time period. The project team encountered numerous challenges in developing both link based and trip-based gap measures. For example, when calculating link-based relative gap measures, it is necessary to have the total number of vehicles on a given link during a given time period, as well as the travel time experienced by that vehicle during that time periods. Deriving these measures is straightforward in the context of aggregate static assignment methods, but is considerably more complex in the context of the regional network simulation model, where vehicles may travel on a portion of a link during a given time period. Similarly, there were challenges in developing trip-based gap measures. Conceptually, the trip-based relative gap should represent the difference between the ”expected” time for a given trip, based on the most recent detailed network link delays, and the “experienced” time for this trip in the Microsimulator. In practice, this results in negative gaps due to differences in how different TRANSIMS tools calculate travel times. Extensive effort was devoted to conceptualizing and implementing reasonable gap measures.

Regarding the number of network assignment iterations that are required in order to achieve a degree of convergence sufficient to support the use of the model, the project team extensively tested and documented the degree of convergence and stability achieved using different network assignment methodologies. The team demonstrated that running additional network assignment iterations was necessary in order to distinguish differences between policy and investment scenarios. Typically, each model system run was comprised of between three and six system iterations, and within each system iteration forty assignment iterations were performed.

Model system convergence

Model system convergence is achieved when the impedances or level-of-service measurements used as the basis for accessibility measures and as key inputs to the destination and mode choice models are approximately equal to the travel times and costs produced by the final network assignment process and is pursued by iteratively feeding the impedances output at the end of each global iteration back as input in the subsequent global iteration. As with network convergence, the degree of system convergence is significantly influenced by the number of times the overall model system is run.

Schedule consistency

Schedule consistency, which refers to ensuring that the detailed schedules produced by the DaySim model are implemented in the TRANSIMS network model with as much fidelity as possible, was another concern of the project team. Both the demand and supply models incorporate (re)scheduling capabilities. It is an open research question as to how to address re-scheduling when inconsistencies are observed between the demand and supply models. DaySim uses a model to predict the arrival time, departure time and duration at the main destination of each tour using a temporal resolution of half-hours which is subsequently disaggregated to minutes, as well as the departure or arrival time for each individual stop on each tour. TRANSIMS uses this detailed time information when routing trips. In most cases, the travel time experienced by the Router when routing a trip differs somewhat from the travel time assumed by DaySim when it generated the trip.

The project team felt it was important to define how the integrated model should accommodate these types of discrepancies. While it is possible to adjust the trip start time, trip end time and/or activity duration time in a variety of ways, it is important to identify methods that do not lead to biased model system results. In the initial implementation and sensitivity testing of the model system, it was assumed that travelers would preserve their activity durations, and would shift the timing of travel to accommodate differences between the "expected" time use to generate the estimates of travel demand and the "experienced" time in the network simulation of this demand. This assumption had the effect of causing trips to "cascade" into time periods later in the day, affecting the model calibration and consistency with the original demand. Subsequent testing, in which travelers were assumed to preserve the time of travel and to adjust their activity durations demonstrated not only greater schedule consistency between the demand and supply components, but also higher levels of link-based network convergence. Scheduling consistency measures that incorporate information on the differences between expected and experienced departure, arrival, and activity durations are necessary to ensure that the final network assigned demand is consistent with the final calibrated demand model outputs.

5.1.3.3 Conclusions

Configuring Daysim to generate temporally, spatially, and behaviorally-detailed travel demand information for use in TRANSIMS was straightforward. Configuring TRANSIMS to generate the skims for input to DaySim was also straightforward. More sophisticated methods of providing TRANSIMS-based impedances to DaySim, such as implementing efficient multi-stage sampling of destinations (and corresponding impedances) at strategic points in the DaySim looping process, or tightly integrating DaySim and TRANSIMS so that DaySim can call TRANSIMS to extract the required measures quickly could potentially be implemented, though the runtime implications and required resource for development were felt to be prohibitive. It was felt that the current methods provide sufficient spatial and temporal detail, and do not necessitate completely acceptable from a runtime perspective. The network convergence equilibration effort revealed that the most effective convergence strategies were often the least acceptable to the larger DTA community, but were necessary in order to ensure sufficiently converged assignments within reasonable runtimes. Schedule consistency was identified as another measure of the soundness of a model solution. Extensive testing of the model system was necessary to determine the number of network assignment and model system iterations required to ensure that differences between alternative scenario model results were attributable to these policy and investments and not obscured by "noise" in the model system.

5.1.4 Model Enhancements

Two of the key goals of the SHRP2 C10A model system development effort were to provide enhanced representation of travelers' sensitivities to price, and to incorporate findings from other SHRP 2 Capacity projects. A number of enhancements were made to both the DaySim and TRANSIMS model components to achieve these goals.

5.1.4.1 Incorporating findings from SHRP2 C04

For modeling highway pricing and congestion effects, the relationship between auto route choice and the other travel choices is critical. In previous implementations of DaySim, auto route choice has been handled outside of DaySim itself. For example, for single occupancy vehicle drivers DaySim did not indicate whether this person chose to pay to use a tolled facility. If both tolled and non-tolled facilities are available for a traveler's given zone pair, it was up to the network assignment model to determine whether a tolled or non-tolled route was chosen, typically by minimizing generalized cost and excluding any information about tour purpose, traveler income, or other relevant factors. Enhancements were made to DaySim to include "toll" and "no toll" choices nested under the drive modes, and transit submode choices under the transit submodes. When toll cost and/or operating cost are key considerations in choosing a route, each traveler may have different trade-offs between travel time and cost (their value of time). DaySim was enhanced to indicate whether the tolled route is used, incorporating information relevant to value of time such as tour purpose and travel income. This information is then passed to TRANSIMS, which includes a set of ninety assignment classes that reflect both toll/non-toll choices as well as value-of-time class. DaySim also accepts as input a set of user classes skim files that include information about the pre-determined best path for each of three value-of-time classes, two toll/non-toll classes, and for each of the twenty-two time periods. The traveler-specific coefficients used in the model were derived from the findings from the SHRP 2 C04 study to the greatest extent possible, both for the functional forms and the magnitudes (work/nonwork, income, occupancy).

5.1.4.2 Price sensitivity

Parallel enhancements to the TRANSIMS tools were also implemented. The original version of the TRANSIMS Router only had the ability to vary values of time used in the generalized cost formulas at the household level, and thus could only reflect household income. The revised Router has the capability to incorporate the trip-specific values of time that Daysim generates, and in the current implementation, approximately 45 value-of-time classes are used in the pathbuilding and assignment. The feature can be used in conjunction with the Router's ability to include or exclude individual links, which is configured to mirror Daysim's toll/notoll pathtype choices.

5.1.4.3 Network microsimulator

Another significant modification to the TRANSIMS tools involved revisions to the Microsimulator. These changes provided important new capabilities, such as the ability to restart trips that get "lost" during the simulation, incorporation of improved speed profiles, production of better estimates of network performance at finer resolutions, and support for multi-threading across multiple processors, which reduces runtimes.

5.1.4.4 Conclusions

The enhancements made to the model system were necessary to improve the model system's sensitivity and to fulfill the goals of the SHRP2 C10A project. Updates to the DaySim model system were relatively straightforward to implement, although these updates were not fully completed until a new DaySim software architecture was implemented, which took significantly longer than expected. The updates to the TRANSIMS model components were much more extensive and involved much more time to implement, and many of these enhancements were under development during the C10A project. While these enhancements were necessary to fulfill project goals, they undoubtedly also resulted in schedule delays.

5.1.5 Model Application

The integrated model controlled through the use of the TRANSIMS Studio software and is comprised includes four software components:

- The TRANSIMS Studio user interface and application management software
- The TRANSIMS and DaySim modeling software,
- Python scripts that define the modeling process, and
- A folder structure housing network and other input data

All of these components are available free of charge from websites that distribute this open-source software.

5.1.5.1 Hardware requirements

The model system can be run on computers using either the Windows or Linux operating systems. Runtimes for the model system are influenced by a number of factors including:

- The computing resources available to run the simulation (such as the number of processors, the amount of memory available, and the speed of storage drive input/output);
- The degree of convergence required for a given model application;
- The size of the synthetic population used as the basis for all choice-making in the model; and
- The level of detail of segmentation employed in the model system (such as the number of time periods).

The project team successfully implemented the model system hardware running on Windows and Linux operating systems. Both the Linux and Windows implementations of the model system typically were configured to use between 24 and 32 processing cores, although this can be flexibly set. Both DaySim and TRANSIMS are multithreaded applications, and TRANSIMS also includes partitioning capabilities. The project team tested a number of alternative configurations of the model system with respect to temporal detail and market segmentation of the skims. As few as four time periods and as many as twenty-two time periods were used in the model system, while some configurations employed no value of time segmentation while other configurations included up to fifty value of time segments. The simplest of these schemes could be run with only 2GB of RAM, while the most elaborate of the schemes required approximately 20GB of RAM. Data storage and access influences runtimes, so in addition to having sufficient RAM, the model system (specifically, TRANSIMS) also benefits greatly from the use of fast hard drives for storage – ideally solid state drives. The project tested the model system using both the TRACC Linux computing cluster located at Argonne National Laboratory, as well as using local Windows-based servers.

5.1.5.2 Application modes

In addition to testing different iterative feedback schemes, the project team also investigated different "application modes" of the model system. These application modes were developed to address the fact that the runtimes for a fully integrated and well-converged model run with repeated full regional simulations of Jacksonville could literally take up to four weeks of time. Three application modes were conceptualized. The "planning" application mode involves the iterative exchange of data between DaySim and the TRANSIMS Router, but not the TRANSIMS Microsimulator. This mode can be used when the analysis needs are expected to result in regional-scale changes in overall level of travel demand, or changes in regional travelers' destination, mode, or time of day choices, but which are not expected to be significantly impacted by local-scale traffic dynamics. The "operations" application mode can be used when the analysis requires an assessment of the impacts of a policy or strategy on local traffic dynamics, and where these improvements are not expected to result in changes in overall level of travel demand, or in destination, mode or time of day choices. The "planning+operations" application mode represents the fully integrated DaySim and TRANSIMS model system that has been described in this document. In this application mode, the TRANSIMS Router and Microsimulator are used to perform a regional traffic microsimulation as part of every model system global iteration, and microsimulation-based network impedance measures are fed back and used as input to DaySim. The advantage of this application mode is that it provides the full range of sensitivities to both changes in regional demand and to local and regional traffic dynamics. However, these extensive sensitivities come with the significant disadvantage of extremely long model system runtimes.

5.1.5.3 Runtimes

The project team extensively tested various iterative schemes to evaluate trade-offs between network assignment and model system convergence and overall model system runtimes. Generally, greater degrees of convergence are required for more spatially, temporally, or behaviorally detailed analysis. However, the level of convergence required for any given analysis is context specific, and the user is afforded many potential means with which to configure the model system and investigate performance and runtime tradeoffs. Keys and parameters which affect DaySim and TRANSIMS runtimes are easily user configurable via two primary control files. The table below illustrates the runtimes for the Jacksonville model system given the different application modes.

5.1.5.4 Staff resources

As with any model system, the level of staff expertise required to interact with the model system will vary depending on the specific nature of how the model is to be used. However, it is safe to say that a higher degree of knowledge and patience is required with interacting with the new integrated model system than is required for using a traditional trip-based model system. Both the DaySim activity-based demand model component and the TRANSIMS network supply model component are more complex than the demand and supply components of a traditional model. Generating the input data for these model components and further modifying these inputs to reflect alternative scenarios is generally straightforward using both typically available data as well as the automated tools that have been developed to support model implementation.

5.1.5.5 Challenges

The challenges in interacting with the model are primarily associated with debugging the model system. As has been repeatedly mentioned, the network simulation model is very sensitive to small scale network coding and parameter assumptions, and the network simulation is subject to frequent failures as input assumptions are being refined. It is critical that staff have the ability to understand and mine which data generated by the model system can illuminate the source of simulation problems, but also to make informed decisions about how to modify model inputs to achieve the proper model sensitivity. It is essential that model users have a basic understanding of Python in order to understand the overall model system flow, as well as robust data manipulation, statistical analysis, and GIS skills. It is not essential that model users know C# or C++, the development platforms used for DaySim and TRANSIMS, respectively.

Although the model system is relatively easy to configure and interact with, transitioning to the using a new model system such as the DaySim-TRANSIMS model system represents a paradigm shift. Modeling staff must have a broader set of skills, and a willingness to investigate and experiment. The types of analysis that can be performed with the new model system are fundamentally different and more expansive than can be performed with a traditional model system, and the application and interpretation of model outputs must be thoughtfully considered. Use of the fully integrated model system would be most valuable when the proposed policies or strategies are expected to result in regional-scale changes in overall level of travel demand, changes in regional travelers' destination, mode, or time of day choices, and which expected to be significantly impacted by local-scale traffic dynamics. These scenarios may include pricing alternatives, capacity enhancements, travel demand management policies, or operation improvements. Finally, it should be noted that the hardware requirements of the integrated model system are clearly greater than a traditional travel demand model

5.1.5.6 Conclusions

The model system software can be flexibly deployed on hardware running either Windows or Linux, and the implementation can be scaled or configured to reflect available hardware resources. The availability of additional processors, more memory, and fast data storage help reduce runtimes of the model system, which can be extremely long for a region the size of Jacksonville. In order to avoid these long runtimes it also is possible to use the model in different application modes, where either the regional network simulation is not used, or where the demand is held fixed and thus there is no need to run multiple global model system iterations. A higher degree of staff knowledge and patience is required when interacting with the new integrated model system than is required when using a traditional trip-based model system. Although many DaySim and TRANSIMS tools exist to assist in data preparation and coding, the model system is highly sensitive to alternative configurations of the model system and to small scale coding issues, and it can take anywhere from an hour to many weeks to generate plausible alternative scenarios.

5.2 Model System Calibration / Validation

5.2.1 DaySim

5.2.1.1 Model transfer

Often, the implementation of a new activity-based model system in a region for the first time involves the estimation of region-specific coefficients for use in the individual models that comprise the overall model system. However, one of the goals of the C10A project was to use existing tools and data to implement the model system. Rather than estimate new models for C10A for the Jacksonville region, which would involve the preparation of estimation datasets, testing of various model specifications for some or all of the individual model components, and the coding of these models for application, instead the activity-based model components and associated coefficients were simply transferred from the existing model DaySim ABM implementation in Sacramento. This approach radically reduced the amount of time required to get the model system up and running.

Because the Jacksonville DaySim implementation was "transferred" from Sacramento, and the model coefficients and alternative-specific constants were initially estimated and calibrated for the Sacramento region, it was necessary to recalibrate the core model components to reflect Jacksonville region-specific travel patterns. Calibration and validation of the entire model system was a highly iterative process that involved making changes to individual model components to better match observed data and evaluating the impacts of these changes on other model components and on overall model system performance. One of the advantages of disaggregate nature of activity-based microsimulation models such as DaySim is that they support more flexibility and realistic calibration adjustments than is possible with aggregate trip-based models.

5.2.1.2 NHTS

Prior to the calibration of the core behavioral components, it was necessary to prepare observed datasets against which to compare the model outputs. The primary observed data source for the calibration of the core DaySim component models was the National Household Travel Survey (NHTS) collected in 2008-2009. For some model components such as household vehicle availability model and the work tour destination model, the NHTS was supplemented with data from the 2005-2009 American Community Survey (ACS). Observed calibration targets for virtually all the components of the activity-based DaySim model system can be prepared using the NHTS or other household survey data. Because transit was not a focus of the C10A effort, detailed transit information such as an on-board survey was not used.

Although additional NHTS "add on" survey data was collected in the Jacksonville region, the overall number of households, persons, tours and trips was relatively small (only 800 households). Since DaySim models travel behavior for a "typical weekday," it was necessary to remove weekend days from the dataset, further reducing the sample size. Although the NHTS contains all the data items required for activity-based model system development, such a small regional sample is insufficient to completely estimate the coefficients in the DaySim component models. However, in the absence of any other datasets contain the information required for activity-based model development, it was deemed acceptable to use the NHTS to derive calibration targets. In addition to the small sample, a number of other issues with the NHTS also affected the calibration process. These issues included the absence of any person, tour or trip information for children under five years of age, missing travel information for some members of the household, inconsistent expansion factors, and unreasonable information such as extremely high shares of individuals choosing to work at home. In order to address these issues in calibration, persons for who complete daily travel information was not available were dropped, all observed and estimated calibration summaries excluded children under five, and revised expansion factors were developed.

5.2.1.3 Calibration process

The results from each individual model component of the DaySim model were reviewed and adjustments made to model coefficients and constants. Although all model calibration adjustments have a simultaneous impact on the model predictions, the calibration effort followed a sequential process from the top to the bottom of the DaySim model hierarchy, because adjustments to upper-level models such as the day pattern model will tend to impact lower-level model predictions such as trip mode choice more than the reverse.

Overall, recalibration of the DaySim model system did not require extensive efforts. Many of the models showed reasonable consistency between the initial estimated and observed results without any adjustments. For example, the initial estimated and observed arrival, departure and duration results aligned closely. It should be noted that the goal of the recalibration effort was to achieve a generally good calibration and validation across a broad range of model statistics, rather than the highly tuned calibration often desired for regional demand models that are to be used extensively in application. The greatest amount of time was spent calibrating the day pattern models, which predict the number and purpose of tours and intermediate stops made by each individual. It is often necessary to adjust these models after the initial calibration to ensure that sufficient travel demand is being generated to match observed roadway and transit volumes. Recalibration of the tour destination models was challenging because for a few of the tour purposes there were few observed tour records. For these purposes, no adjustments to the calibration were made.

5.2.1.4 Calibration challenges

One challenge of the calibration process was that model development involved the use of multiple sets of regional skims at different temporal resolutions and using different network simulation methods at different points in the project. For example, an initial calibration was performed using skims for four broad time periods. Subsequently, this calibration was revisited when Microsimulator-based skims for twenty-two time periods became available, although the additional time period detail did not significantly affect the calibration. The calibration was further revised when Router-based skims from the fully integrated model system were developed. It should be noted that there were noticeable differences between the Router-based and Microsimulator-based skims, with the former generally being more congested. The Router-based skims are based on volume delay functions applied at fine time periods such as 15-minute intervals, while the Microsimulator-based skims are based on the composite speeds of vehicles in the simulation, at this same temporal resolution. These differences, and their implications for model calibration and application warrant further exploration.

5.2.2 TRANSIMS

In concert with the calibration of the DaySim demand model components, it was also necessary to calibrate and validate the TRANSIMS network supply model components. The demand and supply components interact to produce the overall model system calibration and validation. One unique aspect of the TRANSIMS network validation process was that the validation tests were performed using all three network resolutions described earlier in this report. However, the final reported validation results reflected the use of the "planning" network, which was adopted due to faster runtimes.

5.2.2.1 Observed data

The observed network validation dataset was primary derived from temporally detailed (typically, 15 minute intervals) data sources, such as ITS detectors located on major freeways, and telemetered and portable traffic monitoring stations. One of the defining aspects of the C10A project is that both the demand and supply models operate at fine-grained temporal resolutions. It is essential that the observed data used to calibrate and validate the model is also temporally detailed, and as a result it was necessary to develop an entirely new observed count validation dataset. A key challenge in developing the observed count data was cleaning the available datasets for spatial and temporal consistency. The datasets were also review to eliminate "atypical" counts such as those that occurred on extreme weather days, or during construction periods. Due to the seasonal nature of the Florida population, an analysis of variation by time-of-year was also performed, which indicated that volumes during the peak spring period in Jacksonville were approximately 7% higher than during the off-peak summer season. In addition, it was necessary to associate the observed count locations with specific directional links in the networks. Ultimately, approximately 1,000 directional links were tagged with observed counts.

5.2.2.2 Calibration process

Estimated roadway volumes were also compared to observed volumes for four broad time periods, as well as at a daily level. It should be noted that, unlike a traditional static assignment model, the TRANSIMS assignment process does not assign demand to the network by broad time period. Rather, the entire day demand of individual trips are loaded onto the network using minute-level departure time information provided by DaySim. However, time period summaries were still helpful in assessing the performance of the assignment model and informing adjustments to be made to both the DaySim demand and TRANSIMS supply models. Overall, the highway assignment validation by time period looked reasonably good, although the evening time period performed worst, perhaps reflective of the "cascade effect" mentioned earlier.

5.2.3 Conclusions

Transferring the DaySim activity-based demand component from Sacramento to Jacksonville radically reduced the amount of time required to implement the activity-based demand model component of the model system. Additional calibration and validation of some of the subcomponents of the model, such as the daily activity pattern component of DaySim, or the refinement of TRANSIMS networks was necessary in order to improve model performance. However, a number of the models required little, if any, recalibration. Use of the NHTS as the primary observed data source for developing demand model calibration targets was somewhat challenging given the limited weekday sample size in the Jacksonville region and other data completeness issues. Ultimately, more time was spent refining and validating the roadway networks than refining the calibration of the DaySim demand model components. This reflects the fact that network microsimulation models are significantly more sensitive to network coding assumptions and it simply requires more time to identify and resolve these issues.

5.3 Model Sensitivity Testing

Travel demand forecasting model systems are only able to test the effects of policies and assumptions which have been explicitly included when designing and implementing the model system, and are not intrinsically sensitive to the increasingly broad range of transportation policies and improvements of interest to decision-makers. While most regional models are sensitive to large-scale assumptions about land use and demographics, few are sensitive to more detailed assumptions about pricing policies, or to traffic or travel demand management strategies. Even where models have the capability to address these types of policies, they are typically not sufficiently sensitive to the dynamic interplay between travel behavior and network conditions by time-of-day, and are unable to reasonably represent the effects of road pricing, travel demand management, and other policies. Sensitivity testing of model systems involves the evaluation of the effects of changes in model inputs on model outputs.

A key motivating force behind the SHRP2 C10A project is the need to address transportation policies that are being considered in metropolitan planning organizations around the U.S. These policies are not adequately addressed by the current state of the practice travel forecasting models and so the integrated modeling tool developed for this project seeks to improve how these policies are addressed. In order to assess the increased sensitivity of the integrated model system, a set of tests were designed, implemented and evaluated. These tests were designed to illustrate the unique capabilities of the model system and included pricing, travel demand management, and operations.

5.3.1 Pricing

5.3.1.1 Freeway tolls

Two types of pricing tests were evaluated as part of this effort. In the first, a number of scenarios were defined in which freeway tolls varied by time of day. In the second, a number of scenarios were defined in which auto operating costs were modified from the “baseline” condition. For these sensitivity tests, a set of three scenarios were evaluated and compared to the baseline alternative. In the baseline alternative, no costs were assessed at any time. In the first scenario, a fixed $0.25/ mile charge was assessed for anyone using the freeways during the peak periods. In the second scenario, these fixed, peak charges were maintained, while a fixed $0.10/freeway mile charge was added in the midday. Finally, in the third scenario the fixed, peak period charges were increased to $1.00/freeway mile and the fixed, midday charges were increased to $0.50/freeway mile. These values were chosen because a set of similar tolling scenarios were tested for a value pricing project in the Seattle region.

The expected responses to these policies – that travel would decline during tolled periods and on tolled facilities, and that there would be differences in these changes by activity purpose - were all observed in the model system outputs. Declines in travel during the AM and PM peak when freeway tolling is in place, but little change during the midday, which is un-tolled. In the second and third scenarios, there were declines in travel, whether measured as total trips, or vehicle miles or hours traveled, during all tolled time periods, with higher tolls resulting in more reductions in travel. Interestingly, in all three pricing scenarios there were pronounced increases in travel demand during the evenings, suggesting that travelers are rescheduling activities to occur when there are no tolls, as well as fewer scheduling constraints such as are present during the midday. It was also observed in these tests that total trips by time-of-day affects different purposes. For example, work-related travel was relatively unaffected, but the social/recreation-related travel shifted noticeably out of the peaks and into the evening. Finally, the network-based total delay was higher than the base in all scenarios, as the tolling induces travelers to shift onto more capacity-constrained surface facilities. An analysis by facility type indicated that most of this additional delay accumulated on minor arterials. This was likely due to coarseness of the sensitivity test; the tolling scheme was not designed with the specific spatial and temporal extent of network congestion in mind. While some peak location congestion was likely alleviated as a result of the tolling, in many locations across the broad tolling time periods, increased costs were not offset by travel time reductions due to the low levels of baseline congestion.

5.3.1.2 Auto operating costs

For the second set of pricing sensitivity tests, a set of 3 auto operating cost scenarios were evaluated and compared against the baseline alternative. The baseline alternative assumed a cost of $0.12/mile. A lower cost scenario of $0.06/mile was tested as well as two higher costs scenarios of $0.24/mile and $0.60/mile. These tests confirmed that when auto operating costs decline the share of households choosing to maintain zero vehicles also declines, and as the costs increase, the share of zero-vehicle households also increases. However, these changes were relatively modest – a 5-fold increase in the assumer per mile operating cost resulted in only a 1% increase in the number of households choosing to have zero vehicles. In addition, the tests revealed changes in regional tour frequency by purpose, demonstrating that lower costs slightly increased tour-making for discretionary purposes while higher costs slightly decreased tour frequency by purpose. In addition, these reductions in tour-making by purpose were not consistent across purposes, with mandatory work and school tours declining while discretionary purposes such as personal business and social recreational purposes are seemingly unaffected. The declines in work and school tourmaking are partially attributable to more workers and students choosing their home as their usual work or school location. The tests revealed little systematic difference in changes in trips by time of day across the three scenarios, although slight increases in shorter trips resulted from the highest assumed auto operating costs. These shifts did not result in significant changes in network performance or congestion.

5.3.2 Travel Demand Management

TDM approaches incorporate a wide range of strategies aimed at changing travel behavior to reduce congestion and improve mobility. The sensitivity testing for C10A focused on assessing the impacts of a flexible work schedule in which all workers worked fewer days but longer hours on those days. The overall time spent in work activities was held fixed. Because DaySim predicts the daily activity pattern of each individual in the region, it can be used to reflect such a scenario. However, this sensitivity was purely "scenario-based." DaySim cannot identify which policies would be most effective at affecting flexible work schedules, though it can be used to estimate the impact on individual travelers' activity patterns and schedules and on the overall transportation system performance assuming that an effective policy is in place. The model results were consistent with expectations based on the structure and linkages of the DaySim and TRANSIMS model. In general, overall levels of activity generation were lower, although the declines in work-related travel were offset by increases in travel for discretionary purposes.

The model produced shifts in the distribution of travel by time-of-day due to the lengthened workday, although as expected, changes in the destination and mode choices were relatively small. This test did reveal noticeable changes in network performance, with reduced congestion across all facility types throughout most of the day, with a slight increase in congestion in the evening, which reflects both the later return times from work, as well as increased participation in discretionary activities in the evening.

5.3.3 Operations

The sensitivity testing was focused on a scenario in which signals were coordinated using TRANSIMS tools along three primary regional corridors, with the goal of reducing bottlenecks and improving the overall traffic flow. The DaySim-TRANSIMS model system provides sensitivity to these improvements, which traditional travel demand forecast models cannot typically represent the due to their linkage with traditional static network assignment methods that lack detailed network operation attributes and have coarse temporal resolution. The initial model results showed some reductions in delay by facility type, particularly during the peak periods. However, closer inspection of the speed profiles along the three targeted corridor showed more mixed results, with the signal progression producing better speeds in some corridor directions, and worse speeds in other corridor directions. In implementing the operational tests, approximately twenty iterations were required to establish a set of baseline signal timings for the entire region that produced reasonable performance results. Prior to establishing new baseline timings, all the corridor progression changes resulted in the significant simulation problems. This is likely a reflection of the fact that these assumptions are necessarily more detailed in geographic and temporal scope and ultimately more consequential for the performance of the network simulation. As has been noted by others, the sensitivity of DTA and traffic microsimulation models to these detailed inputs suggest distinct challenges when attempting to incorporate these assumptions in a forecasting mode, especially if at a regional scale. Of all the scenarios evaluated as part of this sensitivity testing, the signal progression scenario required the greatest amount of time, and resulted in the least interpretable results. Whereas implementing the other sensitivity test scenarios took approximately one day or less to configure, developing the signal progression test scenario took over one week to configure.

5.3.4 Disaggregate framework

Because both the demand and the supply components of the model system are fully disaggregate, it is possible to trace the impacts of policies and investments on individual travelers from long term choices such as usual work locations all the way down to the specific paths taken by each individual traveler on a second-by-second basis. For example, if capacity is reduced on a key facility, it is possible to determine which specific travelers are affected by this change, and how they are affected. Although disaggregate model results are not reported, this disaggregate framework provides tremendous flexibility for aggregating model results for specific travel markets or communities of concerns. In addition, the disaggregate framework is useful for debugging, calibration and refining model sensitivity.

5.3.5 Simulation variation

A concern frequently expressed about detailed demand and supply microsimulation models is that of random simulation variation. On the demand side, simulation variation can result from the Monte Carlo simulation methodology used to convert the probabilities in the model system to discrete choices. However, the DaySim software has been implemented in such a way as to reduce simulation variation by holding the random number seeds and sequences that are used in the model system fixed across iterations. As a result, the changes between alternatives reflect only the changes in the probabilities for individual choices, rather than the simulation method itself. On the supply side, simulation variation proved to be somewhat more complicated due to the dynamic nature of the network microsimulation. As described earlier, the network assignment methodology involves the iterative simulation of demand and feedback of network performance measures. Because the network simulation model is very sensitive to even small scale changes, significant and sometimes problematic variation in network performance can result from one assignment iteration to the next. In order to address this variation, the project team tested a number of assignment configurations and selected one that produced the most stable results. This configuration produced relative gaps and other stability measures that, while achieving a level of convergence comparable to a static assignment model, produced results that allowed the project team to distinguish between alternatives. Also as described earlier, the project extensively investigated how many assignment and model system iterations were required to produce outputs in which the differences between alternatives were truly attributable to the alternatives and not to "noise" within the model system.

5.3.6 Extracting, managing, and interpreting results

Extracting, managing, and interpreting these results from the model system is not difficult. Much of the network performance data that is of primary interest is available in link-based text files that can be easily processed. Similarly, the travel demand outputs of the model system are also easily interpretable, mirroring the structure of a typical household travel survey. Of course, due to the detailed nature of the model system, the sizes of the files can be quite large. In addition to supporting much more targeted analyses, another advantage of the detailed data produced by the model system is that it can be used to develop visualizations. A number of tools have been developed for TRANSIMS that provide visualizations of second-by-second vehicle movements on the regional networks, capture network performance "heat plots" of congestion, and provide other visualizations as well. These visualizations can be effective not only for conveying model results to decision-makers, but are also essential tools used to debug the model system performance.

5.3.7 Conclusions

The new model system is more sensitive to a wider range of policies than a traditional travel demand model system, and this sensitivity is further enhanced by the detailed representation of temporal dimension, as well as the increased behavioral and spatial detail. In addition, the model system produces a wider range of statistics of interest to decision-makers. Extracting, managing, and interpreting these results was not difficult; however, the level of effort required to effectively test different types of improvements varied widely, from as little as an hour to as more than a week. It was relatively easy to use the model to evaluate the pricing and TDM scenarios, requiring straightforward adjustments to network coding or to model coefficients. Using the model to evaluate the operational scenarios required significantly more effort due to the sensitivity of the network simulation to different signal coordination and timing assumptions. This level of effort would undoubtedly increase if more extensive changes to operational assumptions were required. In addition, even with the additional effort the results produced by the model system did not seem as intuitive as the results of the other scenario tests.

Random simulation variation did not compromise the ability to use the model system, provided a sufficient degree of convergence was achieved both within the network assignment and for the model system overall.