

# Maryland SHRP2 C10 Implementation Assistance - MITAMS: Maryland Integrated Analysis Modeling System

## Final Report

**C10 Implementation Assistance Award Recipient:** Maryland State Highway Administration

**Prime Contractor:** University of Maryland (UMD)

**Subcontractors:** Cambridge Systematics (CS) and Arizona State University (ASU)

### Principal Investigator

**Name:** Lei Zhang

**Title:** Herbert Rabin Distinguished Professor  
Director of National Transportation Center

**Organization:** Department of Civil and  
Environmental Engineering,

University of Maryland – College Park

**Phone:** 301-405-2881

**Email:** lei@umd.edu

### SHA Project Manager

**Name:** Subrat Mahapatra

**Title:** Transportation Engineer Manager

**Organization:** Maryland State  
Highway Administration

**Phone:** 410-545-5649

**Email:** smahapatra@sha.state.md.us

### FHWA Liaison and Headquarters

**Name:** Brian Gardner

**Title:** Team Leader

**Organization:** FHWA- Office of Planning

**Phone:** 202-366-4061

**Email:** brian.gardner@dot.gov

### FHWA Division

**Name:** Kwame Arhin

**Title:** Planning and Program Manager

**Organization:** FHWA-DelMar  
Maryland Office

**Phone:** 410-779-7158

**Email:** kwame.arhin @dot.gov

**Date:** February 20, 2017

# PROJECT KEY MEMBERS

**Maryland State Highway Administration:**

Subrat Mahapatra, Mark Radovic

**Baltimore Metropolitan Council:**

Charles Baber, Yijing Lu

**University of Maryland:**

Lei Zhang, Carlos Carrion, Chenfeng Xiong, Sepehr Ghader, Liang Tang, Di Yang

**Cambridge Systematics:**

Tom Rossi, Martin Milkovits

**Arizona State University:**

Xuesong Zhou, Ram Pendyala

# TABLE OF CONTENTS

<b>1. Introduction to MITAMS...</b>	<b>4</b>
<b>2. Data Assembly...</b>	<b>6</b>
2.1. Networks...	7
2.2. Network Performance Measures...	9
2.3. Land Use...	9
2.4. Travel Demand Models...	11
2.5. Trip Tables...	12
2.6. Traffic Control Data...	13
2.7. Travel Behavior Data...	13
2.8. Multi-resolution Network...	20
2.9. DTALite with MSTM...	24
<b>3. Integrated SILK AgBM-DTALite...</b>	<b>35</b>
3.1. SILK AgBM...	37
3.2. SILK AgBM: Model Calibration and Validation...	52
3.3. SILK AgBM: Sensitivity Testing...	62
3.4. DTALite Traffic Simulation Engine...	66
3.5. Integrated SILK AgBM-DTALite...	68
3.6. Calibration and Validation of SILK AgBM-DTALite...	77
3.7. SILK AgBM-DTALite Demonstration...	87
3.8. Applications of SILK AgBM-DTALite...	92
<b>4. Integrated InSITE-DTALite...</b>	<b>107</b>
4.1. BMC InSITE ABM...	107
4.2. BMC InSITE ABM – Validation and Sensitivity Testing...	111
4.3. DTALite with InSITE ABM...	138
4.4. Integrated InSITE-DTALite – Sequential Integration...	140
4.5. Convergence Paradigm in Sequential Integration...	148
4.6. Integrated InSITE-DTALite: Validation...	151
4.7. Integrated InSITE-DTALite: Application...	155
<b>5. Moving towards Tighter Integration of Dynamic Activity-Travel Demand and Network Assignment Models...</b>	<b>157</b>
5.1. Introduction...	157
5.2. Levels of an Integrated Transport Modeling Framework...	161
5.3. Case Study...	168
5.4. Conclusions...	172
<b>6. Outreach and Technology Transfer...</b>	<b>174</b>
6.1. Planned Workshops...	174
6.2. Supporting Documents...	179
<b>7. Conclusions and Future Research...</b>	<b>181</b>
<b>8. References...</b>	<b>184</b>

# 1. Introduction to MITAMS

The Maryland State Highway Administration (SHA) and the Baltimore Metropolitan Council (BMC) recognize the value of an integrated, multi-modal, safe, reliable, and efficient transportation system to support and sustain economic growth in the region, state and nation. Home to 5.8 million people, the State of Maryland is in the center of the eastern seaboard of the United States. Within the state's boundaries is the heavily congested Baltimore-Washington metropolitan region. Over time, the dense land use, economic activity, over-saturated conditions and limited system expansion capabilities have presented unique mobility, reliability, environmental and economic vitality challenges. SHA has coordinated efforts with Maryland Department of Transportation (MDOT) and other modal agencies, metropolitan planning organizations (MPOs), Federal Highway Association (FHWA), and other agencies and stakeholders to develop transportation solutions that address current and future needs. SHA has been working very closely with the BMC to develop and advance complimentary tools and applications that support regional and statewide transportation decision-making.

Advanced modeling methods and products—such as an integrated advanced travel demand model and a fine-grained, time-sensitive network that can operate at statewide, metropolitan, and subarea/corridor levels—are required by SHA and BMC to meet agency objectives and address various key challenges. With disaggregated data and behavior models, the interactions between supply, demand, policy and external factors will be better understood. SHA and BMC perceive this FHWA SHRP 2 C10 Implementation Assistance project as a great opportunity towards developing a comprehensive data and model improvement program, referred to as the **Maryland Integrated Travel Analysis Modeling System (MITAMS)**.

The **MITAMS** project is composed of integrated, advanced travel demand models with fine-grained, time-sensitive traffic network models to support agency goals in the areas of planning, integrated planning and operations, and transportation systems management and operations (TSM&O). MITAMS operates at statewide, metropolitan, and subarea/corridor levels. This project leverages previous efforts, such as the Maryland Statewide Transportation Model (MSTM) and the BMC InSITE Activity-based Model (ABM).

The Maryland team has developed two integrated models: **BMC InSITE-DTALite** and **SILK AgBM-DTALite**. The first is the result of integrating a dynamic traffic assignment (DTA) tool based on an existing DTALite model that covers the entire State of Maryland, and a state-of-the-art activity-based model (ABM), InSITE, (developed by Cambridge Systematics for the BMC metropolitan area). This integrated model covers most of the urbanized areas in Maryland and will support long-range planning, transportation conformity analysis, investment programming, project prioritization, equity analysis, and corridor/project planning at both SHA and Baltimore-Washington MPOs. The second model is an agent-based microsimulation travel demand model (AgBM), named SILK (for its emphasis on Search Information, Learning, and Knowledge in the travel decision-making process). SILK is funded by the FHWA Exploratory Advanced Research Program and developed by the University of Maryland for TSM&O, active traffic management, integrated corridor management, and pricing analysis along selected corridors in Maryland. This subarea/corridor AgBM-DTA model will be integrated into the Reliability Roadmap and TSM&O

Strategic Plan for sustainable future applications at SHA, MPOs, and MPO member jurisdictions. This report is divided into four main chapters: (1) Data Assembly; (2) Integrated SILK AgBM-DTALite Model; (3) Integrated InSITE-DTALite Model; and (4) Outreach and Technology Transfer.

## 2. Data Assembly

This chapter describes the data items available for the MITAMS project in the data hub. The data items support the development of the integrated model systems: InSITE-DTALite and AgBM-DTALite. In addition, the data items support the development of a multi-resolution transportation network for multi-scale DTA modeling. Both concepts are discussed in this section.

The data items are categorized as follows:

- **Networks**
  - HERE network
  - MSTM planning network
  - MSTM geodatabase network
  - MSTM loaded network
  - 2012 BMC highway network
  
- **Network performance measures**
  - INRIX traffic dataset
  - Traffic count data
  
- **Land use**
  - Land use data at the SMZ, TAZ, and census block levels
  - 2012 BMC zonal land use data
  - MSTM land use data
  
- **Travel demand models**
  - Maryland Statewide Transportation Model (MSTM)
  
- **Trip tables**
  - MSTM trip tables
  
- **Traffic control data**
  - SHA traffic control data
  
- **Travel Behavior Survey Data**
  - 2007/2008 TPB-BMC Household Travel Survey (HHTS)
  - Transit On-Board Survey
  - Regional Air Passenger Travel Survey
  - I-495 Joint SP/RP Departure Time and Lane Choice Survey
  - Mode Search and Switch Memory Recall Survey
  - Maryland Inter-County Connector GPS Survey
  - Multidimensional Travel Behavior Pilot Survey

Each category is described in subsequent sections.

## 2.1 NETWORKS

### 2.1.1 HERE Network

The University of Maryland (UMD) acquired the HERE network in GIS format from NOKIA. This high definition network provides very high resolution and accurate GIS features of the D.C.-Baltimore region street network. The main attributes of the network include but are not limited to:

- Number of lanes
- Speed limit, in miles per hour
- Travel direction
- Distance, in miles
- Presence of tolls
- Presence of bridge or tunnel
- Presence of ramp
- Functional class

### 2.1.2 MSTM Planning Network

The Maryland State Highway Administration (SHA) maintains the MSTM planning network in GIS format, which is used for the Maryland Statewide Transportation Model (MSTM). The main attributes of this network include but are not limited to:

- Distance, in miles
- Free flow speed, in miles per hour
- Functional class
- Initial congested speed, in miles per hour
- Maximum daily lane capacity divided by 50 (service level “E”)
- Off-peak toll, in cents
- Peak toll, in cents
- Posted speed limit, in miles per hour
- Route name

Both the HERE network and the MSTM network were prepared and imported to DTALite using NEXTA. NEXTA (Network EXplorer for Traffic Analysis) is the tool used to import the GIS shape files and attribute tables into the DTALite dynamic traffic assignment simulator. These steps are necessary for the development of multi-resolution transportation networks for multi-scale DTALite. The research team has access to transit line data from the MSTM planning network, but are still working to process and add additional transit data (e.g., Google transit feeds).

### 2.1.3 MSTM Geodatabase Network

The Maryland State Highway Administration (SHA) maintains this network. It was obtained by joining multiple planning networks from the following sources: Baltimore Metropolitan Council (BMC), Metropolitan Washington Council of Governments (MWCOG), Delaware Department of Transportation (DelDOT) Peninsula Model, and the National Highway Planning Networks

(NHPN). This network is the current routable network for the Maryland Statewide Transportation Model (MSTM); it is evolving as MSTM is validated. The geodatabase provides a platform to display other spatial datasets in a shapefile format as well.

The network includes all necessary attributes for demand modeling, including number of lanes, functional classification, posted speed and link prohibitions. The network has national spatial coverage with increased resolution in Maryland. The network attributes cover the following time periods: 2012, 2015, 2020, 2025, 2030, 2035, and 2040.

#### **2.1.4 MSTM Loaded Network**

The Maryland State Highway Administration (SHA) maintains this network. It is an output of the Maryland Statewide Transportation Model (MSTM) and similar to the MSTM geodatabase network. The format of the network is Cube network file.

The network contains traffic volumes on highway links for each of the four time periods: a.m. peak period – 6 a.m. to 9 a.m.; MD off-peak period – 9 a.m. to 3 p.m.; p.m. peak period – 3 p.m. to 6 p.m.; and NT off-peak period – 6 p.m. to 6 a.m. of the following day. It also contains daily vehicular volumes. The network accounts for traffic volumes for several vehicle types: autos; commercial vehicles; long-distance autos; short-distance single unit trucks; short-distance multi-unit trucks; and long-distance trucks. The network has national spatial coverage with increased resolution in Maryland. The network attributes cover the following time periods: 2012, 2015, 2020, 2025, 2030, 2035, and 2040.

#### **2.1.5 2012 BMC Highway Network**

The Baltimore Metropolitan Council (BMC) maintains the BMC highway network. The network is an input to the InSITE activity-based travel demand model developed by Cambridge Systematics (CS). The BMC highway network covers the BMC region, Maryland counties in the Metropolitan Washington Council of Governments (MWCOG), and the District of Columbia. The format of the BMC highway network is Cube network file. The data is for the year 2012.

The BMC highway network includes but is not limited to the following attributes:

- Functional type
- Roadway type
- Area type
- Managed lanes
- Toll code
- HOV limit
- Speed limit
- Capacity
- Free flow speed
- Number of lanes
- Truck restriction

The skim data obtained using the highway network as an input to the BMC trip-based model was used for estimation of the InSITE travel demand model system.

## 2.2 NETWORK PERFORMANCE MEASURES

### 2.2.1 INRIX Traffic Dataset

The University of Maryland (UMD) has access to the INRIX traffic dataset. This data set contains real-time and historical travel time data for over 100 million vehicles in more than 32 countries. This data is obtained from different sources, such as traffic sensors on the network, local transportation authorities, delivery vans, trucks, taxis and users of the INRIX traffic app. This data can provide by-the-minute, 24-hour travel time and speed information for certain road segments.

The HERE network for the MITAMS study site (see Figure 1 in the appendix) was integrated with the INRIX travel time data using 15-minute intervals. This INRIX dataset will be used to compute reliability measures. In addition, the possibility of using this dataset as an input to estimate a logit-based route choice model for DTALite will be explored.

### 2.2.2 Traffic Count Data

The University of Memphis and the Maryland State Highway Administration (SHA) provided this data for the project. The data is in both spreadsheet and GIS format. The GIS format contains the station locations at the MITAMS study site and the attributes contained in the spreadsheet format, which include but are not limited to:

- **A:** A-node (source node)
- **B:** B-node (terminal node)
- **Station:** SHA count station ID
- **Direction:** directional code
  - 0 = North
  - 1= East
  - 2= South
  - 3=West
- **Station\_Dir:** Text string of Station and Direction joined by an underscore
- **A\_B:** Text string of A-node & B-node joined by an underscore (for joining centerline data)
- **2012AAWDT:** Annual Average Weekday Traffic count (for 24 hours) in the year 2012

The traffic count data will be used for the calibration of the integrated model systems and the validation of the integrated model systems.

## 2.3 LAND USE

### 2.3.1 Land Use Data at the SMZ, TAZ, and Census Block Levels

The University of Memphis and the Maryland State Highway Administration (SHA) provided this data for the project. Formatted in GIS, it contains a database of relevant land use information at three levels of resolution: statewide modeling zone (SMZ), traffic analysis zone (TAZ), and census block. The SMZs are the basis for the MSTM transportation assignment and input land use. They nest within counties and are aggregations of existing metropolitan planning organization (MPO)

TAZs. The TAZs are the basis for the MSTM transportation model and land use at the urban level. The census blocks provide land use datasets with finer detail than TAZs and SMZs. For the MITAMS study site, there are 20 SMZs, 48 TAZs, and 688 census blocks. The land use data at SMZ, TAZ and census block levels include but are not limited to the following attributes:

- **JUR\_NAME:** Name of the zone's jurisdiction
- **STFIPS/STATE:** State FIPS code (State of MD= 24)
- **STCOFIPS/CNTY2010:** County FIPS code (State of MD = 24001-24510)
- **COUNTYFP10:** Last 3 digits of STCOFIPS, referring to the county the TAZ belongs to
- **CENT/SMZ\_N:** Unique identifier for each statewide model zone (SMZ)
- **BLOCK:** Unique identifier for each census block
- **TAZ10:** Unique identifier for each traffic analysis zone (TAZ)
- **SQ\_MI:** Area of the zone (or block) in square miles
- **ACRES:** Area of the zone (or block) in acres
- **Emp30Den:** Employment density (projected) in the year 2030
- **HH10:** Number of households in the year 2010
- **RE10:** Number of employment in the retail sector in the year 2010
- **OFF10:** Number of employment in the office sector in the year 2010
- **IND10:** Number of employment in the industrial sector in the year 2010
- **OFF10:** Number of employment in other sectors in the year 2010
- **TOT10:** Total number of employment in the year 2010
- **HH40:** Number of households in the zone in the year 2040
- **RE10:** Number of employment in the retail sector in the year 2040
- **OFF10:** Number of employment in the office sector in the year 2040
- **IND10:** Number of employment in the industrial sector in the year 2040
- **OFF10:** Number of employment in other sectors in the year 2040
- **TOT10:** Total number of employment in the year 2040

The land use data are used as input to the different models in integrated model systems of the MITAMS project.

### 2.3.2 2012 BMC Zonal Land Use Data

The Baltimore Metropolitan Council (BMC) maintains the BMC zonal land use data. It is an input to the InSITE activity-based travel demand model developed by Cambridge Systematics (CS). The BMC zonal land use data covers the BMC region, Maryland counties in the Metropolitan Washington Council of Governments (MWCOCG) and the District of Columbia. The data contains 2934 traffic analysis zones (TAZ). The format of the BMC zonal land use data is a text file. The data is for the year 2012.

The BMC zonal land use data includes but is not limited to the following attributes:

- **TAZ**           Traffic analysis zone (2934 TAZs)
- **RPD**           Regional planning district
- **Tpop**          Total population (including group quarters)
- **IGQ**           Institutional group quarter population

- **NIGQ** Non-institutional group quarter population
- **HH** Number of households
- **MedInc** TAZ median income
- **Workers** Number of workers
- **Emp\_Ret** Retail employment
- **Emp\_Off** Office employment
- **Emp\_Ind** Industrial employment
- **Emp\_Edu** Educational employment
- **Emp\_Hea** Health employment
- **Emp\_Foo** Food employment
- **Emp\_oth** Other employment
- **Acres** TAZ area in acres
- **TKZones** Truck zone (0-12)
- **ENROLL** School enrollment (K-12)
- **ZeroAuto** Zero auto household (NA for ABM)
- **Jur** Jurisdictions

Land use data were used as inputs for model estimation of the InSITE travel demand model system.

### **2.3.3 MSTM Land Use Data**

This data was derived from various sources: Metropolitan Washington Council of Governments (MWCOC) cooperative forecasts; Baltimore Metropolitan Council (BMC) cooperative forecasts; Wilmapci land use projections; Delaware Department of Transportation (DelDOT) Peninsula Model; Maryland Department of Planning projections; and a national econometric model. The spatial coverage includes the Maryland Statewide Transportation Model (MSTM) modeling region, which covers the entire state and at least one county beyond. The unit of the data is at the statewide model zone (SMZ) level. The data is available for the years 2012, 2015, 2020, 2025, 2030, 2035, and 2040. The format of the data is CSV file.

The MSTM land use data includes but is not limited to the following attributes:

- Population
- Households
- Office employment
- Retail employment;
- Industrial employment
- Other employment
- Total employment
- School enrollment

This data is the socioeconomic input data for the Maryland Statewide Transportation Model (MSTM).

## **2.4 TRAVEL DEMAND MODELS**

### **2.4.1 Maryland Statewide Transportation Model (MSTM)**

The Maryland State Highway Administration (SHA) maintains the Maryland Statewide Transportation Model (MSTM), and uses it to forecast and analyze performance measures of the transportation system at national, statewide and urban levels. MSTM provides comprehensive demand modeling and forecasting capabilities in Maryland and beyond that include but are not limited to: freight movement; transportation/land use changes; long-range planning; and corridor status. MSTM contains the following items:

- **Nationwide traffic model:** nodes, links, and travel demand
- **Maryland state boundary shape file**
- **Sensor counts and speed data:**
  - Zond\_id (sensor\_id)
  - Direction
  - Latitude
  - Longitude
  - Measurement\_start (time)
  - Speed
  - Volume
  - Occupancy
- **TMC Speed data:**
  - Tmc\_code
  - Measurement\_stamp
  - Speed
  - Travel\_time
  - Direction
  - Start\_latitude
  - Start\_longitude
  - End\_latitude
  - End\_longitude

MSTM spatial coverage includes the nationwide traffic network and travel demand, and the Maryland statewide sensor data. MSTM temporal coverage includes 24-hour travel demand and 24-hour sensor data for several links. The MITAMS project may leverage MSTM to provide external travel demand trips (e.g., freight trips) that are outside of the demand models in the integrated model system as another input to the DTALite. In addition, the sensor and TMC data may be used for demand calibration and validation of the integrated model systems.

## **2.5 TRIP TABLES**

### **2.5.1 MSTM Trip Tables**

The Maryland State Highway Administration (SHA) maintains this data. It is an output of the Maryland Statewide Transportation Model (MSTM). The tables are available by time period: a.m. peak period – 6 a.m. to 9 a.m.; MD off-peak period – 9 a.m. to 3 p.m.; p.m. peak period – 3 p.m. to 6 p.m.; and NT off-peak period – 6 p.m. to 6 a.m. of the following day. The tables consist of two matrices, one for passenger travel demand (i.e., autos) and one for freight travel demand (i.e., trucks). The spatial coverage of the network is at the national level, with increased resolution within Maryland. The network attributes cover the following time periods: 2012, 2015, 2020, 2025, 2030, 2035, and 2040.

## 2.6 TRAFFIC CONTROL DATA

### 2.6.1 SHA Traffic Control Data

The Maryland State Highway Administration (SHA) maintains this data. It originates from the SHA’s Office of Traffic and Safety. This traffic control data contains the signal timings and signal design plans for each individual traffic controller. This data may be modified as needed to meet the changing traffic conditions. The spatial coverage of the data is the entire state of Maryland. It should be noted that Montgomery County maintains its own signals. The data format is SYNCHRO binary. This data is used for operational analysis and is a required input for calibrating base year conditions.

## 2.7 TRAVEL BEHAVIOR SURVEY DATA

### 2.7.1 2007/2008 TPB-BMC Household Travel Survey (HHTS)

The Transportation Planning Board (TPB) conducted this survey in both the Baltimore and Washington regions from February 2007 to March 2008, using the same survey designs. The survey comprised a combined total of nearly 14,000 households (about 31,000 persons). The survey data are organized into three relational databases described as the household file, person file and trip file. The data is also geocoded at the traffic analysis zone (TAZ) level. The data are formatted as a database.

The data includes but are not limited to the following attributes:

- **Household ID number:** Survey ID field
- **Person ID number** Survey ID field
- **Number of persons in household**
- **Number of vehicles in household**
- **Total household income level** Categorical household income
- **Gender** Male/Female
- **Age** Years
- **Employment status** 1=employed full-time, 2=employed part-time, 0=not employed
- **Student status** 1=enrolled full-time, 2=enrolled

- **Type of school enrollment** part-time, 0=not enrolled  
1=preschool, 2=K-12, 3=post-HS, 0=not enrolled
- **Relationship to respondent** 1=head, spouse, partner, 2=other  
HH member, 3=visitor

The travel survey data are one of the main inputs used for the model estimation of the InSITE travel demand model system.

### **2.7.2 BMC Transit On-Board Survey**

BMC collected transit riders' data through a survey in 2007. This survey was conducted on board transit vehicles, including local, MARC, Metrorail, light rail, and commuter services in the region. This survey contains data on the respondent's current transit trips, including: start and end location; trip start time; time spent waiting for the transit vehicle; and access and egress modes. This survey also collects socioeconomic characteristics of the respondent such as gender, age, and vehicle availability. The usable sample size for the survey is about 13,000 questionnaires. This travel survey data was used as an input in the development of the InSITE travel demand model system.

### **2.7.3 Regional Air Passenger Survey**

The Washington-Baltimore Regional Air Passenger Surveys have been conducted annually since 1981, and recently, on a biannual basis. The 2011 survey collected responses of about 21,000 air passengers at the following airports: Ronald Reagan Washington National; Baltimore/Washington International Thurgood Marshall; and Washington Dulles. The questionnaires collect information such as: trip purpose; trip origin; trip destination; mode of access; household income; time of day; and characteristics of the air passengers. This travel survey data was used as part of a conventional four-step travel demand model that was integrated with InSITE.

### **2.7.4 I-495 Joint SP/RP Departure Time and Lane Choice Survey**

The University of Maryland conducted this joint stated preference (SP) and revealed preference (RP) departure time and mode/lane choice survey. It consisted of two waves of data collection, gathering a total number of 151 effective samples. The first-wave survey was conducted March 21-25, 2011, and the second-wave survey, supplemented by carefully designed departure time search survey questions, was conducted May 23-27, 2011.

The survey questionnaire consisted of two parts. The first part was a list of RP questions about the current travel behavior of respondents and their socio-economic status. The second part contained an SP experiment, which used the answers given to the RP questions to determine the values of attributes presented to respondents.

In the second wave, a series of carefully designed memory-recall questions were employed to gather behavior process data related to the search for alternative departure times. Each respondent was asked to recall the order of alternative departure times they had considered and tried, as well as the travel conditions corresponding to those departure time choices.

The survey used a simple random sampling method. The population of recruited respondents were car drivers traveling on the Capital Beltway in the Washington, D.C. metropolitan area during

weekday morning and afternoon peak periods (6:00 a.m.-10:00 a.m., and 3:00 p.m.-7:00 p.m.). The research team reached a total of 2,200 drivers during the first-wave data collection period. From these, 173 drivers responded to the survey questionnaire, which resulted in a response rate of 7.9 percent. Within the 173 responded surveys, 80 of the respondents completed the survey, resulting in the effective sample size of 80 observations. In the second wave, 1,800 drivers were intercepted; 131 responded to the web-based survey, resulting in a response rate of 7.3 percent. This time, 71 respondents completed the survey. Therefore, the final sample size of the two consecutive survey waves totaled 151 observations.

The RP questions collected the driver's socio-economic information as well as the information about their most recent trip using the Capital Beltway. The variables are listed in the following table.

**Table 2. 1 Variable List**

<b>Variable</b>	<b>Description</b>
<i>Most Recent Trip Variables</i>	
TRIP_PURPOSE	1=Commute, 2=Shopping, 3=Social, 4=Other
TRAVEL_TIME	Travel time (min.)
DEPART_TIME	Departure time (total number of minutes starting from midnight)
DISTANCE	Total mileage for the trip (mile)
FUEL_COST	Fuel cost (\$)
PARK_COST	Park cost (\$)
TOLL_COST	Toll cost (\$)
ENTER_LOCATION	The exit where the driver enters the Beltway
EXIT_LOCATION	The exit where the driver exits the Beltway
SHORTEST_TRIP	The shortest experienced travel time for the same travel (min.)
LONGEST_TRIP	The longest experienced travel time for the same travel (min.)
TT_MIN_495	The shortest experienced travel time on the Beltway for the same travel (min.)
TT_MAX_495	The longest experienced travel time on the Beltway for the same travel (min.)
<i>Socio-economic Variables</i>	
GENDER	1=Male, 0=Female
AGE	
INCOME	1=less than 50,000, 2=50,000-99,999, 3=100,000-149,999, 4=more than 150,000
EDUCATION	1=less than high school, 2=high school graduate, 3=some college, 4=associate degree, 5=bachelor's, 6=graduate
OCCUPATION	1=work for private company, 2=work for government, 3=school staff or faculty, 4=student, 5=self-employed, 6=retired, 7=unemployed, 8=other

Variable	Description
WORKERS	Number of workers in the household
CARS_PER_HOUSEHOLD	Number of cars in the household
VEHICLE_SIZE	1=small, 2=medium, 3=large
VEHICLE_AGE	1=less than 1 year, 2=1-5 years, 3=5-10 years, 4=over 10 years
WORK_ZIP	The ZIP code of the work location
PREFER_DEPART	Preferred departure time (total number of minutes starting from midnight)
ARRIVE_TIME	Actual arrival time (total number of minutes starting from midnight)
PREFER_ARRIVE	Preferred arrival time (total number of minutes starting from midnight)
WORK_START_TIME	Work start time (total number of minutes starting from midnight)
FLEX_AVAIL	1=work schedule is flexible, 2=otherwise

The SP portion of the survey was designed to explore the decisions that commuters make in response to different traffic conditions and pricing schemes. It presented each respondent with two stated choice experiments: (1) toll lane usage, and (2) departure time choice. Each experiment presented seven scenarios of randomly-generated stated choices on the joint alternatives of departure time and lane choice. A typical scenario displayed to the respondents is illustrated in Figure 2.1.

## Capital Beltway -- HOT Lanes

A survey about travel on the Maryland section of the Capitol Beltway (I-495).

**Question 26.**

The following travel options are available for your trip along the Capitol Beltway.  
Your trip is from Exit 36 to Exit 25.

	Normal Travel Lane	SOV Lane (No Passengers)	HOV Lane (Passengers)
Departure Time	8:40 AM	7:40 AM	7:40 AM
Travel Time	45 - 75 mins	30 - 40 mins	30 - 40 mins
Fuel Cost	\$3.90	\$3.30	\$3.30
Toll Cost	\$0.00	\$3.84	\$0.00

Which travel option would you prefer for your trip?

I Will Use the Normal Travel Lanes.  
 I Will Use the SOV Lane (Single-Occupant Vehicle)  
 I Will Use the HOV Lane (High-Occupancy Vehicle)  
 I Will Not Use the Beltway (I will use an alternate route)

Press **BACK** to go to the last question...

**Figure 2.1 Departure time questionnaire**

This survey was used in the estimation of the models in the agent-based microsimulation travel demand model (AgBM), named SILK (for its emphasis on Search Information, Learning, and Knowledge in the travel decision-making process).

### **2.7.5 Mode Search and Switch Memory Recall Survey**

The average commuter usually tries several travel modes before figuring out which is the best choice. One can understand a person's mode search and switch behavior by observing their mode choice dynamics. To collect such longitudinal data and empirically demonstrate the proposed model, a pilot survey was conducted to collect samples from students at the University of Maryland. A total of 360 students from the university were solicited for participation in the data collection. Among them, 146 respondents completed the survey. This survey was conducted in September 2012 and covered approximately three years, from fall 2009 to summer 2012. The observed three-year horizon was divided into six time periods, representing semesters in each academic year. This semi-annual precision was chosen because it was accurate enough to record long-term commute mode search and decisions for the target subjects; at the same time, this time distinction helped maintain a manageable memory-recall burden.

In the survey, each respondent was asked to complete a questionnaire regarding household/personal socio-economic and demographic characteristics. This was followed by a series of memory-recall questions, employed to reconstruct which travel modes respondents employed and/or considered for their commuting trips. First, each respondent was asked to recall the time period when she/he arrived at the University of Maryland and, for each time period, which mode she/he primarily used to commute. The associated level-of-service information was also gathered. Travel cost data for drive-alone and carpool modes were gauged based on trip distance, vehicle fuel efficiency, number of passengers and annual parking costs. Then, the respondent was asked to recall the most-considered alternative travel mode during that time period. This recall process was repeated for each of the time periods to recollect:

- **The primary mode of transportation for commuting to the university:** This was identified as the habitual travel mode for that time period.
- **The alternative mode of transportation:** This was identified as the alternate travel mode most often considered during that time period. If the subject indicated that no alternative mode was considered, the alternate travel mode would be recorded the same as the habitual mode for the brevity of the analysis.

The following variables were collected from this survey:

- Primary mode for each time period
- Alternative travel mode
- Shortest door-to-door travel time (min)
- Longest door-to-door travel time (min)
- Transit in-vehicle travel time (min)
- Transit out-of-vehicle travel time (min)
- Travel distance (mile)
- Travel cost (\$)

- Bus/Metro fare (\$)
- Other costs (\$)
- Number of transfers

This survey was used in the estimation of the models in the agent-based microsimulation travel demand model (AgBM), named SILK (for its emphasis on Search Information, Learning, and Knowledge in the travel decision-making process).

### **2.7.6 Maryland Inter-County Connector GPS Survey**

The Maryland Inter-County Connector GPS Survey was conducted from November 16th, 2011 to February 6th 2012. The dataset includes 264 individuals and their travel patterns, recorded in one-minute intervals for the 70-day survey period. After data cleaning, the GPS dataset was translated to a trip format, with each trip starting from the origin latitude/longitude, ending with the trip destination latitude/longitude, and including each point in between.

The GPS Survey participants were selected as a sample of Prince George's and Montgomery Counties in Maryland. They responded to a postcard campaign soliciting participants in a study that records their travel patterns. The postcard led them to a survey website where they completed information about themselves (including demographic information and personal driving habits) for consideration in the GPS portion of the survey. The total number of individuals to sign up for the participation survey was 800 individuals. Of those 800 individuals, 300 were selected in order to ensure an accurate representation of the area (Prince George's and Montgomery County). The demographic variables that were considered include household income, age, education, gender, race, employment status and marital status. Driving characteristics, such as miles driven on an average day and whether they have a valid driver's license, were also considered in the selection process.

The sample size for the GPS dataset is 264 individuals out of the 300 that were initially selected. The discrepancy comes from dropouts during the survey, GPS calibration errors, GPS failure and user errors. The dataset includes about 1,500,000 minutes of data over the survey span, averaging roughly 95 hours of continuous data collection per person. This averages out to one-and-a-half hours of travel information per person per day. Depending on the individual, however, the amount of data collected varies widely. While all individuals live in the survey area (Prince George's and Montgomery County), everything, including long distance data were collected for the GPS survey; therefore, the data occurred in a total of twenty-one states (from Florida to the South, Maine to the North, and as far as Kansas to the West).

The GPS data collection method was straightforward. The survey participants were shipped the GPS logger with instructions on how to plug it in and turn on the device. Each device was previously calibrated by the research team in order to record data at the correct interval and give a unique identification number to match the GPS data to the previous survey information (demographics and driving behavior). The participants were told to keep the device in their car and leave it on for the duration of the survey. During two points during the GPS survey, participants were asked to log onto the survey website to complete a more traditional travel survey. This online travel survey was used to match locations and purposes to those recorded by the GPS device, as well as to serve as a backup that could be used to validate the data during the data

cleaning process (pointing possible errors). There were 500 participants who completed a one-day online travel that could be used alongside the GPS survey.

Each GPS data point in the survey includes these variables:

- **Latitude/Longitude:** Basic location data
- **Date:** Data for when the GPS device was recording
- **Time:** Time for data point (accurate to the second)
- **Heading:** The heading of the device (direction it is pointing)
- **Altitude:** Elevation above sea level (ft.)
- **Speed:** Miles per hour the device is moving
- **HDOP:** Horizontal dilution of precision
- **Number of Satellites:** The number of satellites tracking the position

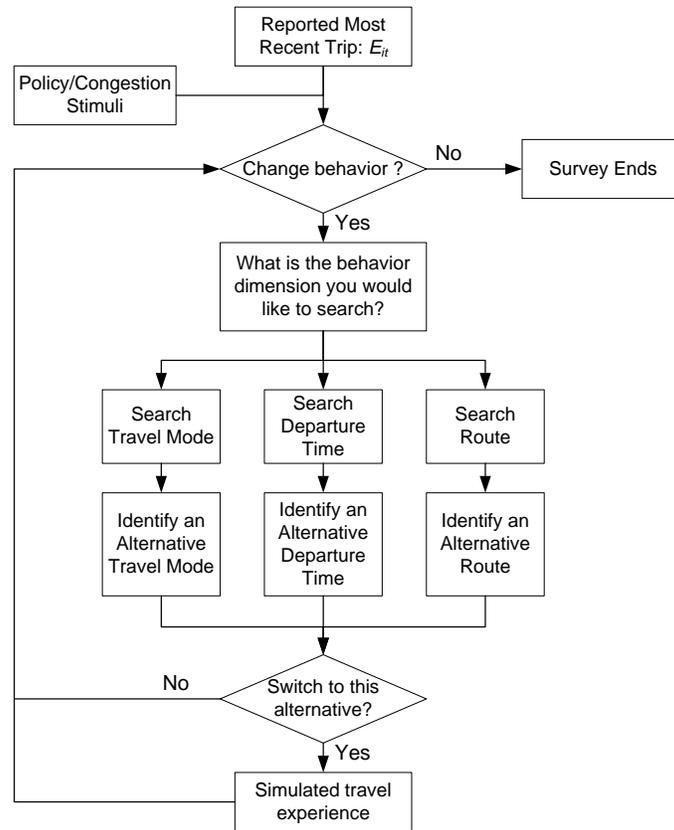
Added/Calculated variables for each data point:

- **Point of Interest Type:** GPS data is merged with the OpenStreet map POI.
- **Soak Time:** Time from the end of destination point to the next origin point
- **Location Identifier:** Identification number for origin/destination to match other points with a similar location (within 300 ft.). This way, the user's home location will always have the same Location Identifier number.
- **Distance:** Distance from one point to the next point in a trip.

This survey was used as the estimation of the models in the agent-based microsimulation travel demand model (AgBM), named SILK (for its emphasis on Search Information, Learning, and Knowledge in the travel decision-making process).

### **2.7.7 Multidimensional Travel Behavior Pilot Survey**

The University of Maryland conducted a stated adaptation experiment administered online to explore possible substitutions to the longitudinal information that is typically missing. This survey was conducted in the fall of 2013. This survey method is particularly useful when seeking answers from respondents on several what-if questions such as, "what would you do if you were faced with specific constraints/conditions." It helps capture respondents' multi-faceted behavioral responses. Furthermore, it has the capability to infer in the procedural decision-making process, which embeds the behavioral foundation of the proposed theory and models, since respondents will naturally exhibit satisficing behavior if playing the scenarios repeatedly for enough iterations. The procedure is reported in Figure 2.2.



**Figure 2.2 The stated preference experiment flowchart**

Beginning with the most recent self-reported trip, exogenous policy/congestion changes are assumed in each scenario to alter the travel condition for that trip. It is further assumed that each agent will adapt to those changes by searching new modes, departure times, and/or routes. The dimensions wherein the behavior adjustment occurs are asked explicitly in the survey for each subject. The subject then is asked to present the alternative she/he would identify and search along that dimension (this data infers the determination of search rules). Once a search has been recorded by a subject, the program will feed a corresponding travel condition simulated in the back-end for the subject to consider and make a switching decision between the alternative and the habitual one (this data infers the decision rules). Another round of behavior adjustment (which could be in the same dimension or in another dimension) will occur unless the subject states satisfaction about the travel experience. Iteratively repeating this process, a complete behavioral adjustment sequence of each subject can be observed. The pilot study included 110 University of Maryland staff and students. They performed adaptations under schemes such as overall congestion increase and road-pricing scenarios.

This survey was used in the estimation of the models in the agent-based microsimulation travel demand model (AgBM), SILK (for its emphasis on Search Information, Learning, and Knowledge in the travel decision-making process).

## 2.8 MULTI-RESOLUTION NETWORK

To take advantage of different networks and related data sources, the Maryland team has devised a united network based on a common true shape base network that integrates all other related networks and data sources. This integrated network would be able to support various modeling purposes, from statewide/regional planning models to subarea/corridor level operational models. If such an integrated network is built, different agencies can collect and maintain modeling-related data in the same network, which makes data sharing and transferring much easier. The major challenge of building such an integrated network is determining how to merge two networks with all the useful attributes. An intuitive solution would be to overlay the two networks in a geographic information system (GIS) software like ArcGIS and utilize some functionalities, like spatial join to connect the networks based on spatial relationship. However, this approach would generate a lot of errors if the two networks were not perfectly overlaid; researchers would have to visually compare the networks to find the errors and manually correct them. Such methods would require a lot of time and effort to identify and correct errors, especially for a large area. In addition, it would be hard to perform quality control and identify whether all the errors were rectified.

To solve these issues, this paper proposes a network conflation method that can merge two networks together with all the useful attributes; usually a planning network and a true shape network. Midpoint rectangle buffers are created for each link in the true shape network. Instead of directly joining the two networks, the rectangle buffers are merged into the planning network. The buffer joint count during the spatial join process can be used to quickly identify and locate the errors. The buffer joint count can also be used to quantitatively measure the percentage of errors, which can be used for quality control purposes.

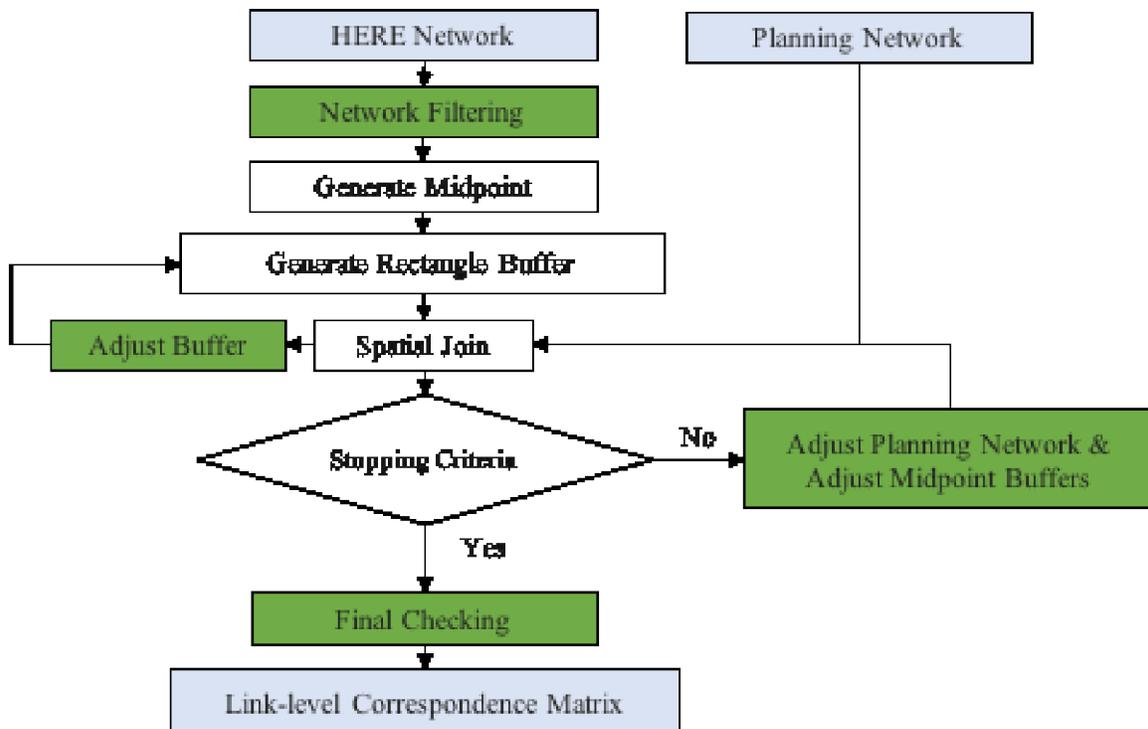


Figure 2.3 Work flow of the proposed network conflation method

The primary goals of the proposed network conflation method are to (1) fully take advantage of the built-in GIS functionalities in order to reduce manual work and facilitate the process, and (2) to develop a measurement that can be used to quantitatively measure the percentage of errors, which can then be used for quality control purposes. To achieve these two goals, the proposed method first generates rectangle buffers in the midpoint of each link in the HERE network. Instead of directly joining the HERE network with the planning network, the generated midpoint buffers are used to spatially join with the planning network. Spatial join refers to a GIS operation that transfers attributes from one layer to another based on their spatial relationship. When the rectangle buffers spatially join with a network, all the links that intersect with the rectangle buffers will be joined to the buffers. The number of planning network links joined to each buffer can be used to judge whether the link-to-buffer join is correct. Several link-to-buffer join cases may be found: one-to-one join, zero-to-one join and multiple-to-one join. If the buffer is joined with exactly one planning network link, which means the joint count equals one for that buffer, the join is considered correct. If the buffer joint count is not one, which means either no link is joined with the buffer or multiple links are joined with the buffer, an error occurs. The buffer joint count can be used to quickly identify and locate buffer spatial join errors for later correction. The buffer joint count can also be used for quality control. The percentage of buffers with a joint count not equal to one can be considered as the error rate of the network conflation process. As the errors are continuously corrected, the error rate will become smaller. When the joint count equals one for all the buffers, the network conflation work is considered complete. Figure 2.3 shows the work flow of the proposed network conflation method. The final output of the network conflation would be a link-level correspondence matrix that can be used to merge the networks afterwards. The stopping criteria refers to the condition where the joint count equals one for all buffers. Several important steps of the method are further explained below:

### **Step 1. Network filtering for HERE network**

As a true shape network, HERE network covers both major highways and local connectors, while the planning network usually only covers the freeways and major arterials. In this step, only HERE network links that are covered by the planning network are retained, while the other links are masked temporarily. The purpose of this step is to avoid generating unnecessary buffers.

### **Step 2. Generate midpoint rectangle buffers on HERE network**

For each link in the filtered HERE network, a rectangle buffer perpendicular to the link is generated at the midpoint. The buffer width is a fixed value. The height of the rectangle buffers needs to be optimized with the planning network during the spatial join process to achieve the lowest error rate, which is the percentage of buffers with joint count not equal to one. To achieve the optimized height of the rectangle buffers, different buffer heights are tested and the one with the lowest error rate is used. After obtaining the optimal value for the buffer size, the midpoint rectangle buffers are generated in the filtered HERE network.

### **Step 3. Spatial join the buffers with the planning network**

After generating the midpoint rectangle buffers in the HERE network, the buffers are spatially joined with the planning network and the percentage of buffers with joint count not equal to one is calculated. If the joint count equals one for all the buffers, jump to Step 5; otherwise, go to Step 4.

#### Step 4. Identify and correct the errors of the spatial join

Identify all the buffers of which the joint count is not equal to one and correct the errors. Various methods can be used for error correction, such as adjusting the planning network, adjusting the rectangle buffers, etc. Since the goal is to get the link-level correspondence matrix between the two networks, it is fine to perform these adjustments if the link ID remains the same. After correcting the errors, go back to Step 3 and repeat the steps until the error rate equals zero.

#### Step 5. Final check

In this step, check the networks again to make sure there are no additional errors. After that, generate the link-level correspondence matrix, which can be used to integrate the planning level network and the HERE network.

### 2.8.1 Maryland Case Study

(a) Planning network



(b) HERE network



Figure 2.4 Planning network and HERE network for Maryland study area

The proposed network conflation method was implemented in the Maryland study area. The planning network and HERE network in the study area are shown in Figure 2.4. After testing different values, 50 meters was selected as the height for the midpoint rectangle buffers. Several rounds of manual adjustment were conducted to correct the errors in the buffer spatial join process. Table 2.2 shows the buffer joint count information during each round of spatial join and adjustment. As shown in the table, the error rates reduced significantly in the first several rounds, which indicates that the marginal benefit reduces as the process is repeated. After eight rounds of adjustment, the joint count becomes one for all buffers. The link-level correspondence matrix was generated after the final check.

Once the link-level correspondence matrix is created, the HERE network is connected to the planning network; all data stored in the planning network can be transferred to the HERE network.

Table 2.2 Summary of the Buffer Joint Count and Error Rate for each Round of Buffer Spatial Join

Joint Count	0	1	2	3	Sum	Error Rate
Round 1	24.9%	64.3%	9.8%	1.1%	100.0%	35.70%
Round 2	16.2%	72.6%	10.1%	1.1%	100.0%	27.40%
Round 3	12.7%	76.0%	10.3%	1.0%	100.0%	24.00%
Round 4	11.0%	79.5%	8.6%	0.9%	100.0%	20.50%
Round 5	10.2%	80.5%	8.3%	0.9%	100.0%	19.50%
Round 6	1.7%	96.1%	2.0%	0.1%	100.0%	3.90%
Round 7	0.1%	99.3%	0.6%	0.0%	100.0%	0.70%
Round 8	0.0%	100.0%	0.0%	0.0%	100.0%	0.00%

## 2.9 DTALite Model with MSTM

The Maryland Statewide Transportation Model (MSTM) intends to be a multi-layer travel demand model representing nationwide, regional, and statewide travel. Once completed, MSTM will be a powerful tool in analyzing transportation movements within Maryland and immediate surrounding areas. The boundaries of the network have been established to include all of Maryland, Delaware and the District of Columbia, as well as adjacent portions of Virginia, Pennsylvania and West Virginia.

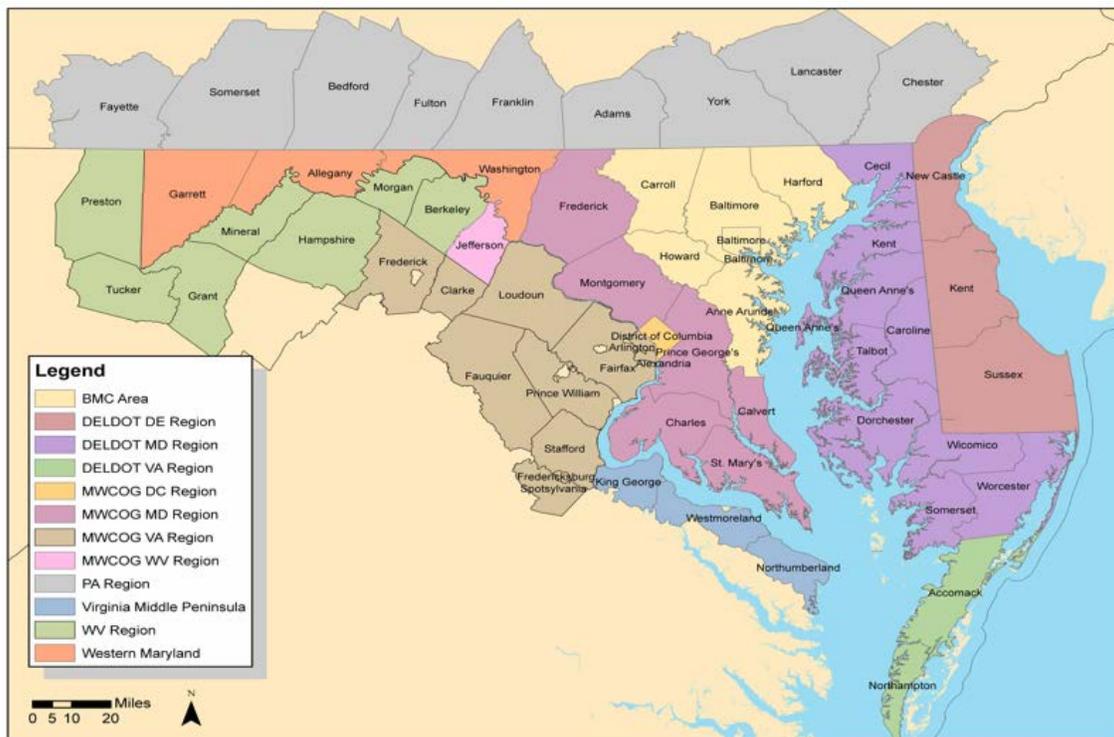


Figure 2.5 Maryland statewide modeling region

The general process follows the steps listed below. Each step is described in more detail in the following section.

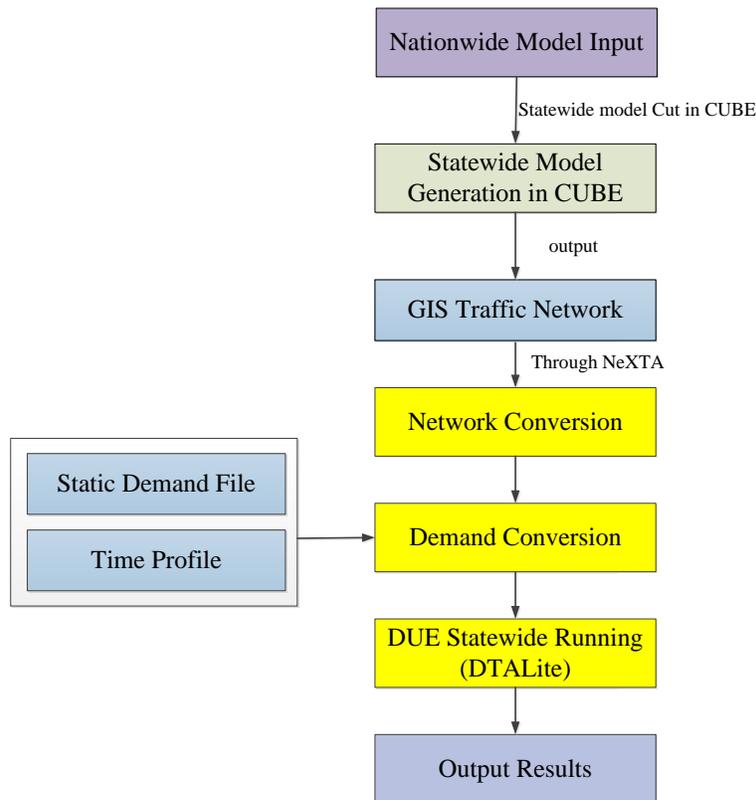
**Step 1:** A nationwide model is generated in Cube.

**Step 2:** The MTSM model is extracted from the nationwide model in Cube, which includes GIS traffic network shape files.

**Step 3:** All basic network data are imported into NeXTA and converted to the required DTALite input format.

**Step 4:** The static demand file and time profile provides the required dynamic travel demand input, which is converted for processing in the dynamic traffic assignment model.

**Step 5:** Through an iterative process in DTALite, simulation results are provided as output, including information related to time-dependent link volume, density, speed, etc.



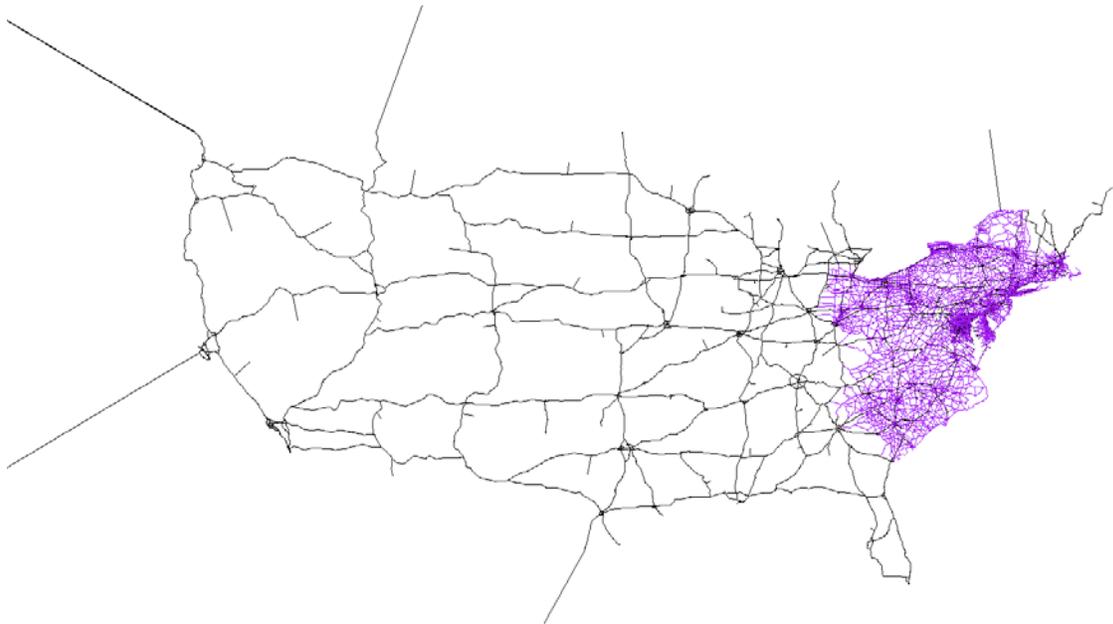
**Figure 2.6** Flow chart showing required steps to build the MSTM in DTALite

### **2.9.1 Import the Traffic Network**

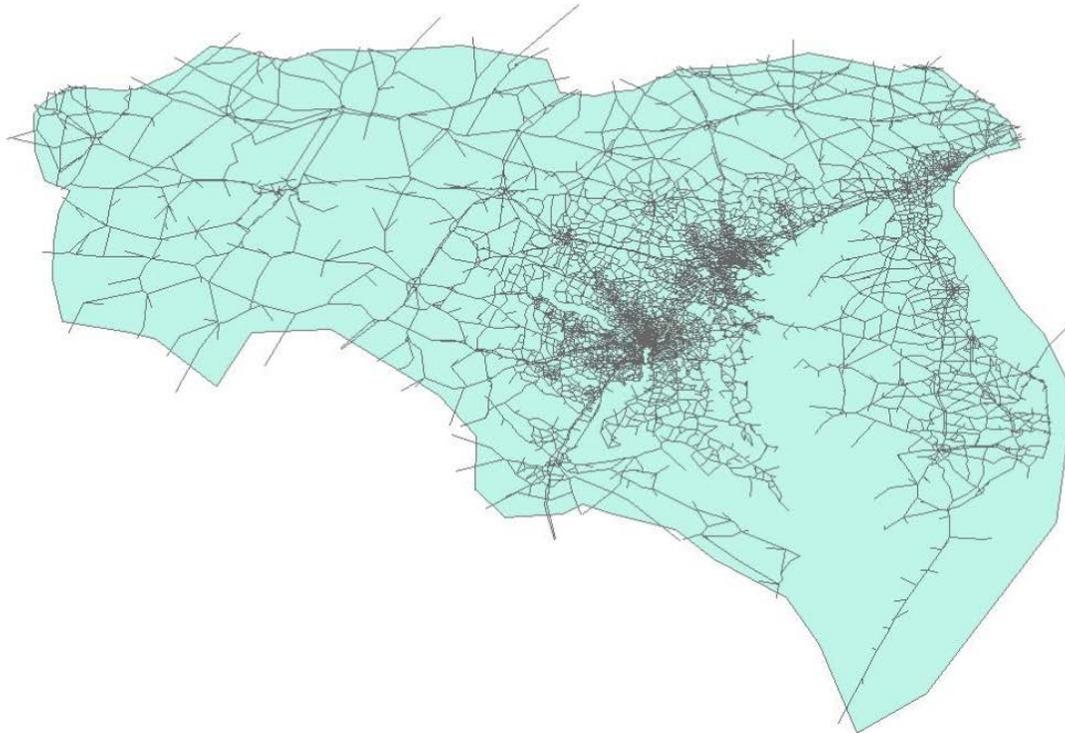
DTA modeling efforts require high-level representation of transportation system components, leading to a more challenging model development process. Preparing networks with the required level of detail is one of the most cumbersome steps for DTA planning applications that require a specific data format. To assist in this process, Network Explorer for Traffic Analysis (NeXTA)

software can be applied as a user-friendly interface and data hub. NeXTA has the ability to convert traffic network information from different planning packages into the required data format for DTALite.

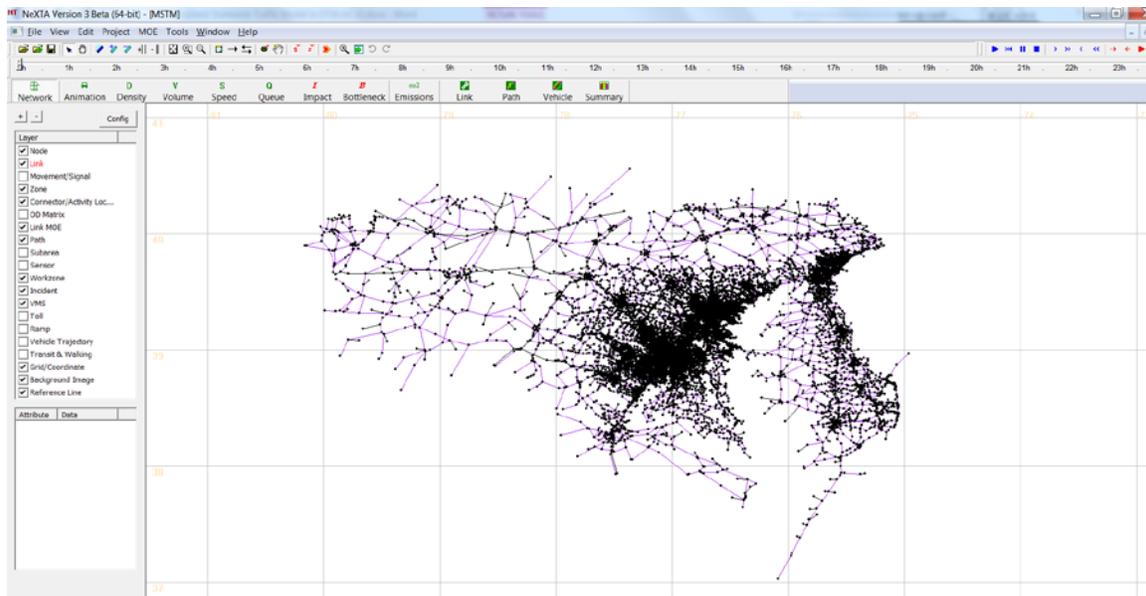
To create the MSTM network, a nationwide traffic model was constructed in CUBE and then the subarea of interest for the statewide model was extracted (Figure 2.8). Based on the GIS network shape files, NeXTA was used to extract the necessary traffic data (such as lane capacity, speed limit, link length, number of lanes, from node, to node, link id, etc.) and store them in the following CSV formatted files: *input\_node.csv*, *input\_link.csv*, *input\_zone.csv*, and *input\_activity\_location.csv*. Figure 2.9 presents the converted MSTM traffic network in NeXTA, which consists of 21,742 nodes, 54,323 links, and 1,696 zones.



**Figure 2.7 Nationwide traffic network for MSTM**



**Figure 2.8 MTSM Statewide model extracted from Cube**



**Figure 2.9. Statewide model imported into NeXTA**

### **2.9.2 Prepare Time-dependent Origin-Destination Travel Demand**

The MSTM Cube model included 216 statewide origin-destination (OD) trip tables disaggregated by purpose, income, mode and time period, and an additional 20 nationwide OD tables for long distance truck and auto trips. These trip tables were converted to DTALite format using Cube Voyager scripts and were subsequently consolidated with time profiles corresponding to each trip

table within a demand metadata file. Time profiles were generated from the 2007 Baltimore-Washington Region Household Travel Survey data after applying a moving average smoothing technique. Time profile data is stored in the following file: *input\_demand\_meta\_data.csv*.

### 2.9.3 Build Toll Settings and Movement Prohibition

The MSTM model incorporates truck restrictions and toll collections on certain links at specific times. It also restricts entry of SOV and truck demand on HOV links. All restrictions and tolling can be accomplished in the MTSM model by setting time-dependent toll values in the *scenario\_link\_based\_toll.csv* file, which is shown in Figure 2.10. Links with toll values have the word “toll” added. Different restrictions require different settings in the *scenario\_link\_based\_toll.csv* file. Key field names in this file are listed in Table 2.3.

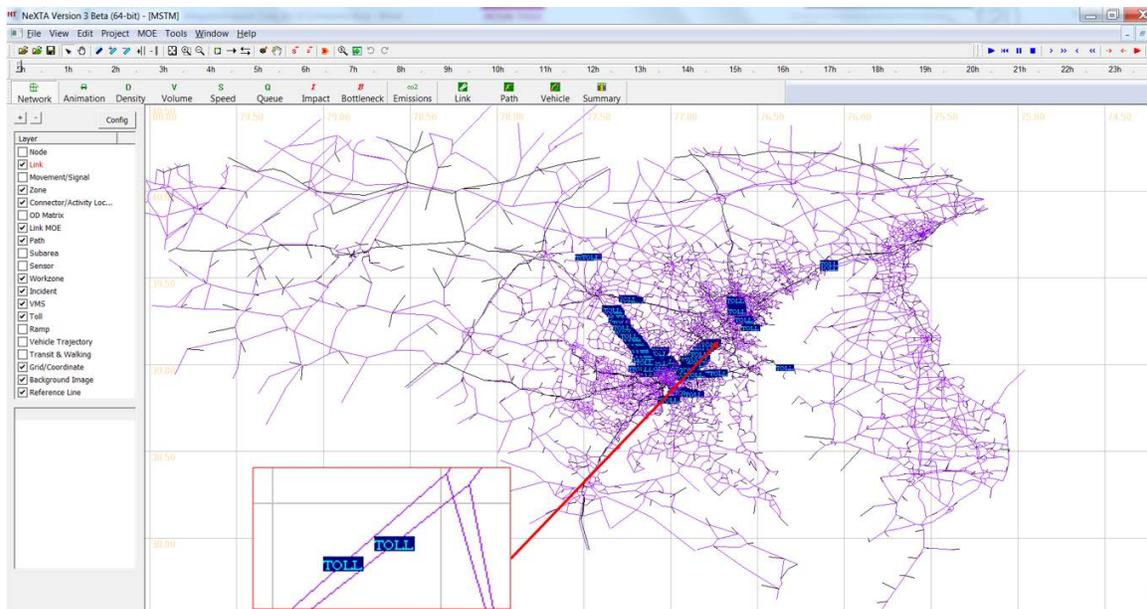


Figure 2.10 Link toll visualized in NeXTA.

Table 2.3 Sample of Settings in Scenario\_link\_based\_toll.csv

Link	Start Time (min)	End Time (min)	Charge for LOV (\$)	Charge for HOV (\$)	Charge for Truck (\$)	Charge for Intermodal (\$)
[10855, 15795]	390	570	9999	0	9999	0

For truck restrictions, a large toll value in “Charge for Truck” is added on links based on the restriction start and end times. Similarly, for HOV links, a large toll value is added in “Charge for LOV” and “Charge for Truck”. For normal toll collection periods, actual toll values appear in the corresponding columns.

In addition, transit links and walking links are included in the MSTM model. To avoid vehicles assigned on those links, users are referred to the “demand\_type\_code” field in the *input\_link.csv* file. For example, the demand type for LOV, HOV and Truck is 1, 2 and 3, respectively. Therefore, the “demand\_type\_code” for transit links and walking links takes a value other than 1, 2 or 3.

Turning movement restrictions are identified in the *AMS\_movement.csv* file in DTALite, as shown in Table 2.4. A turning movement is defined by *up\_node\_id*, *node\_id* and *dest\_node\_id*; the *prohibited\_flag* will be activated when this column has a value of “1”.

**Table 2.4 Sample of Movement Prohibition Settings in AMS\_movement.csv**

<b>node_id</b>	<b>up_node_id</b>	<b>dest_node_id</b>	<b>prohibited_flag</b>
2391	2396	2397	1

### **2.9.4 Select Traffic Flow Models**

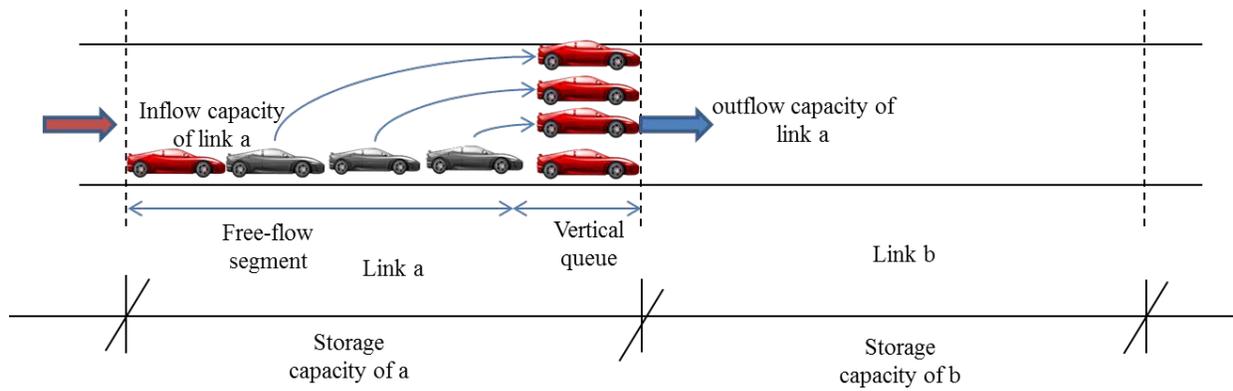
Traffic flow models available in DTALite include BPR function, point queue model, spatial queue model and Newell’s kinematic wave model. These models are reviewed in Table 2.5.

**Table 2.5 Unified Capacity-Oriented Modeling Framework (from simplest to complex)**

<b>Traffic Flow Model</b>	<b>Link Temporal Flow Capacity</b>	<b>Spatial Storage Capacity, <math>K_{jam}</math></b>	<b>Congestion Propagation through Backwave Speed, <math>W</math></b>
VDF (BPR)	Considered through demand/delay function, but allow $DV/C > 1$	No	No
Point-queue	Yes (inflow capacity is infinite, and outflow capacity is equal to link capacity)	No	No
Spatial-queue	Yes (inflow capacity is infinite, and outflow capacity is equal to minimization of link capacity and storage capacity of downstream link)	Yes	Yes
Newell’s Kinematic Wave Model	Yes (inflow capacity is equal to outflow capacity under free-flow traffic state for each link)	Yes	Yes, shock wave propagation through $K_{jam}$ and backward wave $w$ (applied on freeways only)

As part of this project, two traffic flow models have been evaluated: BPR function-based model and point queue model. With a BPR function-based model, the assigned traffic volume on a link is permitted to exceed its capacity, so there is no queue spillback and no propagation captured using this model. Therefore, for DTA models, the use of the BPR function is only preferred in the case of very large models because of their complexity; a general result can be obtained for further analysis and for application of more complicated traffic flow models. With a point queue model,

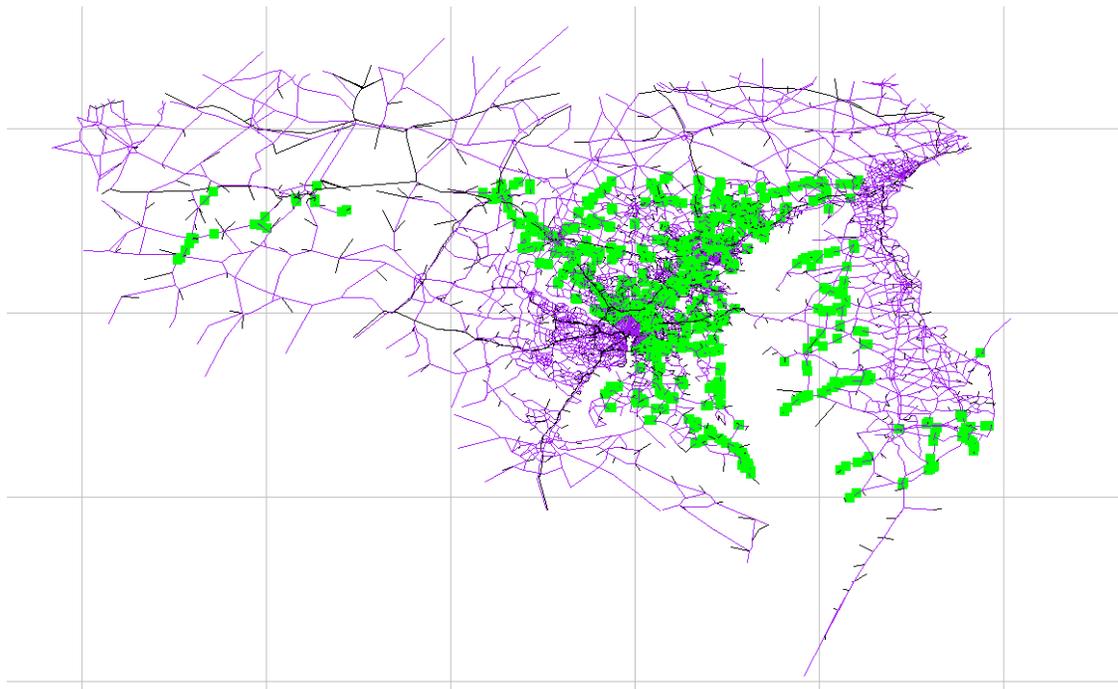
the inflow capacity on a link is infinite and the outflow capacity becomes the only constraint to capture the effect of traffic congestion at major bottlenecks. Figure 2.11 represents a point queue represented as a vertical stack queue of vehicles, where some of the vehicles are mapped or “rotated” from the physical link (shaded) to the vertical stack queue.



**Figure 2.11 Point queue model represented as a vertical stack queue.**

### **2.9.5 Collect Sensor Data**

Traffic count data are also required for model calibration and validation. These hourly data were obtained from the MSHA (Maryland State Highway Administration) in 2009 and cover a 24-hour period; there are hourly link count data for 1,247 links within the Maryland network. They are stored in the *sensor\_count.csv* file. As shown in Figure 2.12, the green small squares represent the location of link traffic count sensors within the network.



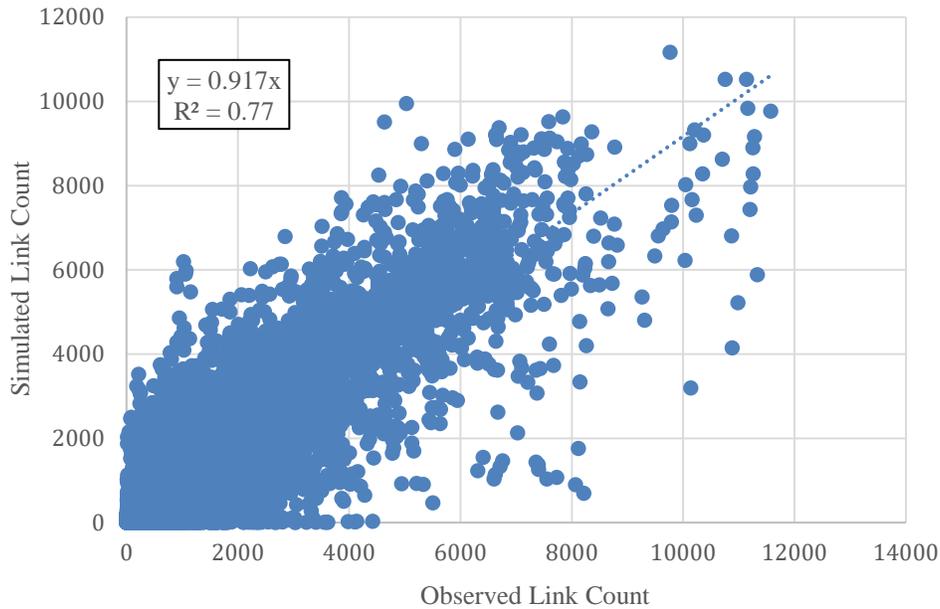
**Figure 2.12 Sensor location visualized in NeXTA.**

### 2.9.6 Validation of Simulation Results

MSTM model is run using both BPR function and point queue model in DTALite in order to validate simulation results with observed data.

#### 2.9.6.1 BPR Function Simulation Results

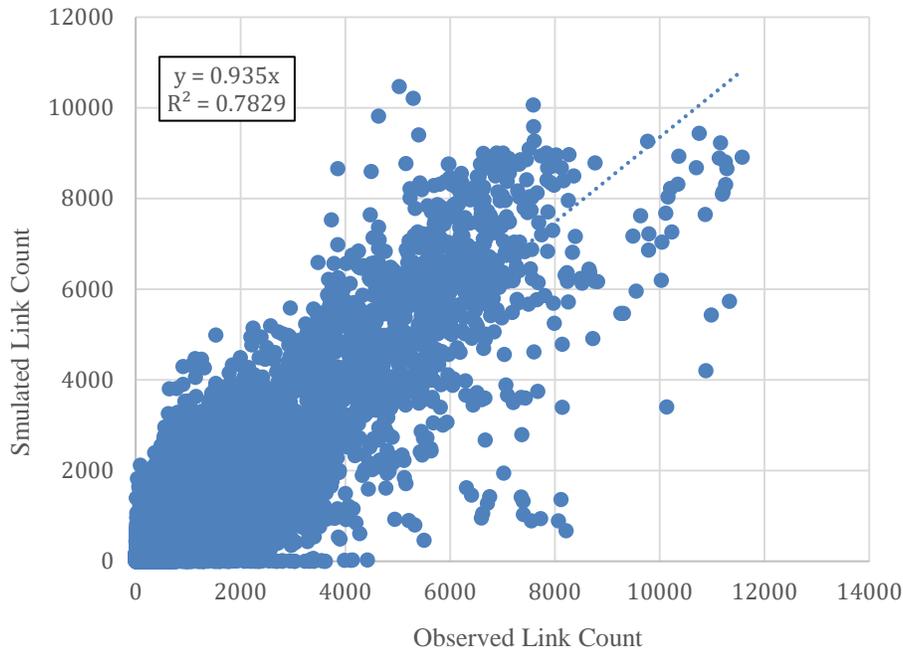
Figure 2.13 shows the comparison of simulated hourly link counts and observed hourly link counts for 1,247 links within the Maryland statewide traffic network. Its  $R^2$  is equal to 0.77. In this test, seven iterations were performed using a workstation with 192 GB memory and 40 threads running at 2.80 GHz. Each iteration took approximately 2 hours and 50 minutes.



**Figure 2.13 Simulated link count (BPR function) vs. observed link count.**

#### 2.9.6.2 Point Queue Model Simulation Results

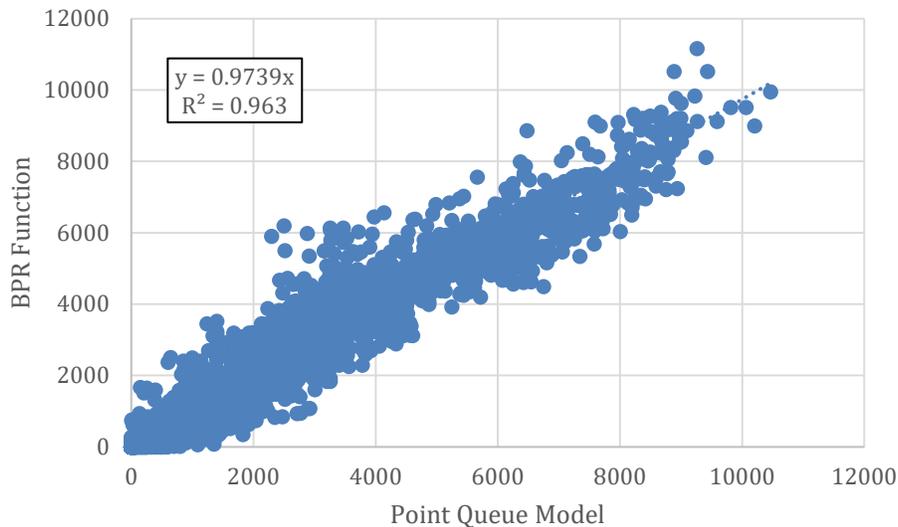
Figure 2.14 shows the comparison of simulated hourly link counts and observed hourly link counts for 1,247 links within Maryland statewide traffic network using the point queue model. Its  $R^2$  is equal to 0.78, which is an improvement compared with the result of BPR function. This is because the point queue model captures traffic interactions better than the BPR function between time intervals. In this test, twenty iterations were performed using a workstation with 192 GB memory and 40 threads running at 2.80 GHz. Each iteration took about 2 hours and 20 minutes.



**Figure 1 Simulated link count (point queue model) vs. observed link count.**

**2.9.6.3 Comparison of BPR and Point Queue Model Simulation Results**

Figure 2.15 shows the comparison of simulation results of BPR function and point queue model. The  $R^2$  is 0.96, which indicates that the simulation result of BPR is very close to the results of the point queue model. From a computational time perspective, the BPR function is an acceptable model for a large-scale traffic network such as MTSM, even though the point queue model produces more accurate results.



**Figure 2.15 Simulation result comparison of BPR function and point queue model.**

**2.9.7 Origin-Destination Matrix Estimation**

The general origin-destination (OD) demand estimation problem looks to find an estimate of the OD demand matrix by effectively utilizing traffic flow observations and other available information. The objective in DTALite is to minimize (1) the deviation between observed and estimated traffic states and (2) the deviation between aggregated path flows and target OD flows, subject to the dynamic user equilibrium (DUE) constraint, which is represented by a gap-function-based formulation. The general process is described in the following sections.

### **2.9.7.1 Prepare Sensor Data**

As mentioned in Section 2.5, 24-hour, hourly link counts data is one key input for OD demand estimation. One of difficulties for ODME is the consistency of data format between collection format and required simulator input. An example for the format in DTALite is presented in Table 2.6. The “count\_sensor\_id” in the *sensor\_count.csv* and *input\_link.csv* files must be consistent for proper mapping during ODME. It also requires input for the start and end times of data observation.

**Table 2.6 Sample of Link Count Data in sensor\_count.csv**

from_node	to_node	count_sensor_id	start_time_in_min	end_time_in_min	link_count
21358	21404	21358->21404	0	60	18

### **2.9.7.2 Input Settings for ODME in DTALite**

The attributes for ODME are the required input into the *input\_scenario\_settings.csv* file. The required information is explained in Table 2.7.

**TABLE 2.7 Related Attributes in the File input\_scenario\_settings.csv for ODME**

Data Field	Value	Notes
number_of_assignment_days	500	Total iterations (value should be greater than ODME start iteration)
traffic_flow_model	1	This parameter defines a specific traffic flow model used in both assignment and ODME of DTALite. The number 1 indicates a point queue model in this example. Selection of the spatial and Newell’s KW model is also feasible.
signal_representation_model	0	This parameter defines a specific signal control for DTALite.
traffic_assignment_method	3	Assignment method “3” is dedicated to ODME.
ODME_start_iteration	20	Defines the first iterative assignment period for ODME to converge to the user equilibrium state and further generate a sufficient number of paths for path flow adjustment.
ODME_max_percentage_deviation_wrt_hist_demand	30	The maximum percentage of demand deviation from base-line dynamic demand.

ODME_step_size	0.05	Moving size of each step in the path flow adjustment algorithm.
calibration_data_start_time_in_min	0	These parameters specify the time window for ODME to use sensor data. Note that, if the sensor data is only available between 990 and 1050 mins, calibration will be performed only for that time period.
calibration_data_end_time_in_min	1440	

### **2.9.7.3 Run the ODME model and Check Simulation Results**

After ensuring all sensor data are correctly prepared, DTALite (64 bit version) runs the simulation, which can make use of more than 2 GB RAM (random-access memory). After the simulation is complete, the ODME result can be reviewed in the *output\_summary.csv* file. There is one field in the file “Age UE gap” that helps judge the user equilibrium (UE) condition before ODME and another field are “ODME: r\_squared”. It is used as a measure of the deviation between simulated results and observed data.

### 3. Integrated SILK AgBM-DTALite

The objective of this research is to produce an application-ready, integrated transportation model that can predict both the changes in travel behavioral adjustments and the dynamics in network traffic conditions for a future year or hypothetical scenario. Existing practices that link travel demand models and traffic models often follow a sequential approach. Travel demand patterns are predicted and then treated as fixed input in traffic assignment or dynamic traffic assignment (DTA) models. In addition to a sensitivity to long-term land use and socio-demographic changes, day-to-day travel demand patterns should also be responsive to traffic conditions. Excessive travel times could lead to route changes, departure time adjustment, or even transportation modal shifts. This reversible linkage is necessary for an accurate representation of travel behavior, as many traveling agents try to improve their travel experience through behavioral changes every day. In return, behavioral representation is reflected in traffic dynamics. To achieve this objective, a developed agent-based travel behavior model (AgBM) that emphasizes modeling individual searching, information learning and knowledge, called SILK (SILK, see Xiong et al. 2015), is employed as the travel behavior-modeling tool. Traffic dynamics are simulated using DTALite, a lightweight dynamic traffic assignment and simulation engine (Zhou and Taylor, 2014). These two tools are integrated together to incorporate day-to-day behavioral adjustments into transportation and traffic simulation processes.

Microscopic traffic simulation models exhibit offer strong advantages in capturing detailed traffic dynamics and have been approved in practice as a valuable tool for evaluating corridor capacity expansion and traffic operation improvements. Their applications have recently been extended to address a broader range of transportation-related issues, including congestion management, multimodal corridor improvements, evacuation planning, land use and economic development. However, analyzing these issues requires models that can consider travel demand responses to traffic management strategies such as peak spreading, modal shifts, and traffic diversions at the corridor and regional levels. These travel demand dynamics can be readily addressed by activity/agent-based travel demand models (Bhat and Koppelman 1999; Bowman and Ben-Akiva 2001; Salvini and Miller 2005; Vovsha et al., 2008). On the other hand, agents in demand models require traffic conditions and travel experience as inputs for behavioral adjustments. An integration of agent-based travel demand models with microscopic traffic simulation models can provide researchers with a powerful tool to simulate the complex transportation system and provide answers to planning and policy questions.

Several research efforts have been dedicated to this field. Dia (2002) presents an agent-based approach to model dynamic driver behavior under the influence of real-time traffic information. For each form of information provided to drivers (e.g., quantitative delay, predictive and prescriptive delay), a number of multinomial logit models are developed to determine the factors that affect the propensity of the drivers to adjust their travel patterns and to determine the values of these factors. Based on these driver behavioral models, and in order to evaluate the impacts of providing drivers with travel information, an agent-based traffic simulation tool is developed, which applies the belief, desire, and intention (BDI) agent architecture. The feasibility of this approach is demonstrated through a case study on the same corridor where the travel behavior survey was conducted. Another study that applies the BDI concept is Rossetti et al. (2002). They propose an extension to an existing microscopic simulation model called Dynamic Route Assignment Combining user Learning and micro-simulation (DRACULA). In this extension, the

behavior of agents is represented in terms of mental attitudes that allow them to make decisions about route choice and departure time.

Hao et al. (2010) focus on integration of an activity-based travel demand model, TASHA, with a dynamic agent-based traffic simulation model, MATSim. In this study, an iterative process is applied for the integration and a series of data conversions are proposed to make this process possible. The modeling framework is implemented in the greater Toronto area. Flötteröd et al. (2011) is another study that links the demand models to the agent-based traffic simulation. This study concentrates on the calibration of demand models in the context of dynamic traffic assignment. The calibration simultaneously adjusts the route choice, departure time choice and mode choice (car versus no car) of individual travelers by employing a Bayesian framework.

One existing integration effort that is drawing increased research attention is the use of computational process models to represent travel behavior. These belong to a group of new methods that depart from rationality assumptions; they implement learning, adaptations, information acquisition, and decision-making efficiently by taking advantage of computer power. These models are microsimulations relying on heuristic arguments and imitation of human behavior. A large number of real-world or benchmark problems can be analyzed by applying these models to simulate numerical results in different set-ups. Examples on the rapidly growing list of research in this area include Pendyala et al., (1998), Arentze and Timmermans (2004), Wahba and Shalaby (2011), and Auld and Mohammadian (2010). On one hand, these models introduce more complex learning, adaptation and behavioral rules, instead of utility maximization. On the other hand, multi-agent simulation cannot prove—but only suggest—a certain feature of travel pattern, and still assumes sequential decision process when integrating traffic assignment or DTA models. Therefore, it requires additional theories to conceptualize the more rigorous behavioral foundation and better explain behavior adjustments along multiple choice dimensions (Arentze et al., 2004; Pinjari et al., 2011).

SILK-AgBM is an agent-based travel behavioral model system that emphasizes the role of Searching, Information, Learning and Knowledge (SILK). It can be efficiently interfaced with a dynamic traffic assignment model. SILK AgBM does not rely on rules extracted from data mining to explain behavior, but instead theorizes multidimensional knowledge updating, search start/stopping criteria, and search/decision heuristics. These behavioral modules are formulated and empirically modeled and integrated in a unified and coherent system. Procedural and dynamic behavioral rules are estimated to represent rich behavior in travel mode, departure time, route, and en-route diversion choices. They can be enhanced to potentially accommodate real-time decision-making on the part of the traveler and the driver's response to network conditions/congestion and real-time traveler information (in addition to a host of other dynamic mobility applications and active traffic and demand management strategies).

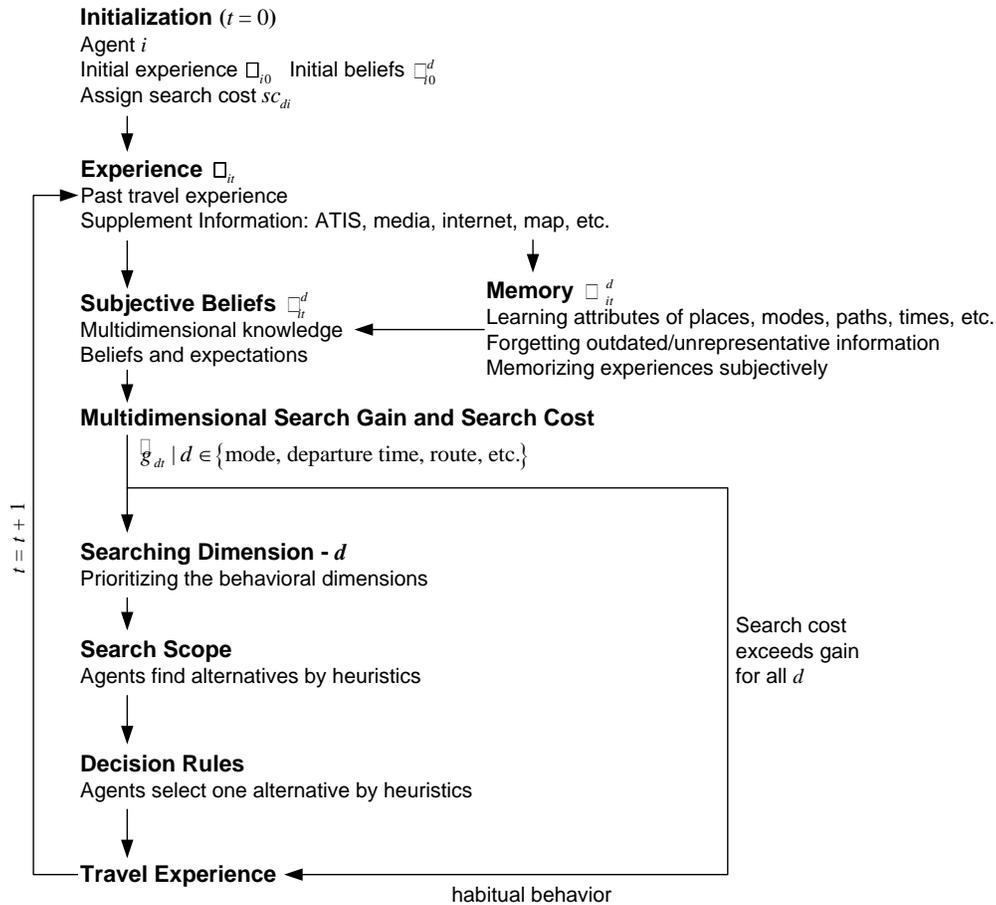
In this chapter, SILK-AgBM is integrated with a dynamic traffic simulation tool called DTALite. Among a number DTA simulators and software packages, DTALite is an open-source mesoscopic (i.e., a resolution level that lies between macroscopic and microscopic) DTA simulation package. Coupled with a graphic user interface named the Network eXplorer for Traffic Analysis (NeXTA), DTALite provides a useful research and educational tool for researchers, students and practitioners to model traffic network and to extend research capabilities to cover various applications in traffic management and transportation planning. Its open-source nature and computational efficiency

ensured by the parallel computing power are among the reasons DTALite was selected in this model integration effort.

### 3.1 SILK AgBM Model

The theory starts with the definition of artificially intelligent agents and their characteristics. Each agent  $i$  is treated differently with socio-demographic attributes, personal experience, knowledge, and subjective beliefs. At any given time, an agent has a certain level of knowledge about places, activities and transport networks in an urban area. This spatial/temporal knowledge can be employed to solve various spatial/temporal decision tasks, such as choosing destination, departure time or routes. This problem-solving process consists of several procedural steps in the true behavioral sense. First, each agent  $i$  at a given time period  $t$  possesses experiences, denoted as  $\mathbf{E}_{it}$ . Agents acquire  $\mathbf{E}_{it}$  through past searches or through information sources such as Internet, media, advanced traffic information system (ATIS), etc.  $\mathbf{E}_{it}$  is time-variant as the agent searches and accumulates a-priori experiences in the urban transportation network day-by-day. Travel experiences with similar payoffs that occur routinely may reinforce the agent's memory, while the travel experiences that are not representative may be easily forgotten (Arentze and Timmermans, 2003). Moreover, agents are assumed to be able to search information about one behavioral adjustment dimension at a time; for example, agents may search for an alternative route or search for an alternative travel mode. Thus, each past experience can be mapped into one single dimension  $d$  and form a multidimensional memory space  $\mathbf{M}^d$ .

The memory space keeps updating, alters the aspiration level, and changes subjective beliefs  $\mathbf{P}_{it}^d$ . Therefore, an agent determines the expected gain  $g_{dt}$  from a search for alternatives in each behavioral dimension  $d$  based on his/her subjective beliefs. Information acquisition and other mental efforts are explicitly modeled as perceived search cost  $sc_{dt}$  when agents are searching for alternatives for each behavioral dimension. These search cost variables are recognized in this theory as inconveniences and risks associated with each behavior adjustment dimension. It is the interplay of these subjective search gains and costs that jointly determines when a search for alternatives in dimension  $d$  is initiated or stopped in time period  $t$ . Although the subjective search gain is defined by an individual's beliefs and therefore can be quantitatively derived, it is much more difficult to theoretically determine the magnitude of a perceived search cost that should be individually different. Once the multidimensional behavioral adjustment evidences can be observed, the perceived search cost and its relations with other variables can be empirically derived.



**Figure 3.1 Conceptualization of multidimensional travel decision-making theory**

If an agent decides not to search in a dimension, habitual behavior in that dimension is executed. Otherwise, the agent will employ a set of search rules to search from her/his knowledge and identify a new alternative. After identifying an alternative, she/he needs to determine whether or not to switch to that alternative. The decision rules constitute a mapping from spatial/temporal knowledge (especially experienced travel conditions corresponding to different alternatives) to a binary decision: switch to the alternative or retain habit. Both the search rules and the decision rules should be empirically estimated from observed search processes.

### **3.1.1 Modeling Imperfect Knowledge**

Search, learning and knowledge play a crucial role in decision-making. A rational person will choose the best alternative from a set of feasible alternatives. The term “rationality” would also require that this rational person holds the knowledge that is derived from coherent inferences. In contrast, more realistic models are intended to allow modelers to construct agents who systematically do not possess perfect knowledge and who do not make correct inferences, but rather, biased ones.

An agent explores decision opportunities by searching her/his feasible environment and learns knowledge about various payoffs related to the search and decisions. Here the spatial/temporal knowledge is generalized as multidimensional vectors with each vector corresponding to a

particular dimension. Assume that each agent  $i$  at any given time period  $t$  possesses a list of past experiences,  $\mathbf{E}_{it}$ . Each experience is characterized by a generalized cost:

$$C_E = \sum_n \lambda_n \psi_n \quad (1)$$

Where  $n$  denotes the index of different related attributes such as travel time, cost, schedule delays, mode comfort, etc.;  $\psi$  denotes the vector of attributes; and  $\lambda$  denotes the coefficient to translate values into monetary costs (e.g., value of time). This generalized cost is adopted to measure the outcome of each experience and to set an anchoring point for the search model. Assuming that in each behavioral dimension  $d$ , an individual's perceptual capabilities allow the separation of generalized cost into a number of categories. If  $C_E$  that falls into the generalized-cost category  $j$  has been observed  $m_j$  times in prior experiences, the memory this individual has about the generalized cost in dimension  $d$  is fully described by a vector  $\mathbf{M}^d = (m_1, \dots, m_j, \dots, m_J)$ . Individuals update memory space through learning and forgetting processes. Bayesian learning relies on the premise of some prior knowledge. When new information from various sources becomes available, learning occurs and obeys the Bayes' rule. Forgetting relies on the cognitive weighting of each past experience, which can be measured as a function of the recentness and representativeness of the experience (Arentze and Timmermans, 2003). Once the weight is lower than a certain threshold parameter, the experience will be eliminated from  $\mathbf{E}_{it}$ .

Bayesian learning theory relies on the premise of some prior memory ( $\tilde{\mathbf{M}}$ ). When new evidence ( $e$ ) from various information sources is available, learning occurs and follows Bayes theorem. In other words, the posterior memory is updated using conditional probabilities:  $P(\tilde{\mathbf{M}} | e) = P(e | \tilde{\mathbf{M}}) \cdot P(\tilde{\mathbf{M}}) / P(e)$  (this equation can also be expressed as *posterior = likelihood · prior / evidence*). When a new alternative in this dimension is experienced and the associated generalized cost falls into category  $j$ , the updated memory becomes  $\mathbf{M}^d = (m_1, \dots, m_j+1, \dots, m_J)$ . Let the vector  $\mathbf{P}^d = (p_1, \dots, p_j, \dots, p_J)$  represent an individual's subjective beliefs, where  $p_j$  is the subjective probability that an additional search in dimension  $d$  would lead to an alternative with  $j$ th level of generalized cost. In order to quantitatively link  $\mathbf{M}^d$  and  $\mathbf{P}^d$ , we assume that individuals' prior beliefs and memory follow a Dirichlet distribution, which is a  $J$ -parameter distribution. Therefore, the posterior beliefs will also be Dirichlet distributed, since the Dirichlet is the conjugate prior to the multinomial distribution (Rothschild, 1974). The probability density function is defined as:

$$P = \frac{\Gamma(N)}{\prod_{j=1}^J \Gamma(m_j)} \prod_{j=1}^J p_j^{m_j-1} \quad (2)$$

where  $N$  denotes the total number of  $\mathbf{M}^d$  observations and Gamma function  $\Gamma(m_j) = (m_j - 1)!$ . According to the law of large numbers, as sample size  $N$  grows, this assumption asymptotically converges to:

$$E(p_j) = \frac{m_j}{N} \quad (3)$$

Bayesian learning is capable of describing updates of spatial knowledge about the attributes of spatial objects, and relations between spatial objectives when repeated observations are available.

Travel time on a roadway section, waiting time at a transit station, level of congestion for a specific trip during a peak hour, attractiveness of a housing unit in a neighborhood, distance between an origin and a destination, proximity of a shopping center to the route from work to home, etc.

### 3.1.2 Modeling Multidimensional Search

An individual, based on her/his experience  $\mathbf{E}_{it}$  and subjective beliefs  $\mathbf{P}_{it}^d$ , forms expectations on potential gain (search gain) from behavioral adjustments along each dimension. The decision to search for a new alternative is based on the interplay of subjective search gain and perceived search cost. Let an agent's generalized cost on the currently used alternative be  $C$ . The subjective search gain ( $g_{dt}$ ) is based on subjective beliefs,  $\mathbf{P}_{it}^d$ , and defined as the expected improvement regarding generalized cost savings per trip from an additional search:

$$g_{dt} = \sum_{j(\forall C_j < C)} p_j \cdot (C - C_j) \quad (4)$$

Where  $C$  is the minimum of all experienced generalized costs because individuals can select from all tried alternatives in dimension  $d$  and pick the one with the lowest costs  $C_{\min}^d$ . We assume all individuals start with a preferred travel pattern. It can be the stabilized travel pattern with an initial generalized cost  $C_0$ . Once a policy/congestion stimulus emerges, travel condition deteriorates. Let us further assume that individuals have the initial beliefs that search and switching to another alternative will lead to a travel condition as good as their original travel condition  $C_0$ , until they search and experience otherwise. As the search process proceeds, the subjective probability of finding an alternative with  $C_0$  after  $N$  searches is  $1 / (N+1)$ . Therefore, Eq. (4) can be further simplified as:

$$g_{dt} = \frac{C_{\min}^d - C_0}{N + 1} \quad (5)$$

While  $C_0$  remains universal among all dimensions,  $C_{\min}^d$ , the currently best travel option(s) in dimension  $d$ , can differ in each dimension  $d$  since the search processes in different dimensions vary and result in diverse outcomes. The subjective search gain  $g_{dt}$  evolves and reflects how much value each search can gain based on subjective beliefs. Once  $g_{dt}$  is less than or equal to zero, it indicates that search along dimension  $d$  is no longer worthwhile and the search process will not initiate. A positive  $g_{dt}$  will asymptotically decrease to zero as the number of searches increases and as a better alternative is found ( $C_{\min}^d$  getting increasingly closer to  $C_0$ ).

Furthermore, the theory formulates satisficing behavior that even with positive gains, individuals may stop a search once the gain is lower than the perceived search cost. The search and information acquisition are no longer free, as this theory recognizes the inconveniences and risks associated with each behavior adjustment dimension. This impedance is conceptualized as a search cost for each agent and each dimension. Search cost can be perceived and inferred once individuals' searching sequences can be reconstructed using empirical observations collected from survey. The empirical data provides evidence about agents' search and decision processes. Each individual follows her/his own path along the three dimensions in reaching the final behavior decisions. When it is observed that an individual ends her/his search in dimension  $d$  and has searched  $N$  times along

that dimension for the time being, it infers that the individual satisfies after  $N$  rounds of search in  $d$ . The search cost must be lower than  $g_{d,t-1}$  so that the  $N$ th search is meaningful and rewarding.

Meanwhile, the search cost must be higher than  $g_{dt}$  so that the  $(N+1)$ th search does not occur. Let us denote individual  $i$ 's search cost along dimension  $d$  as  $sc_{di}$ , which is viewed as an innate personal characteristic for individual  $i$ . It can be estimated by using the lower and upper bounds:

$$sc_{di} \leq g_{d,t-1} = \frac{C_{\min,t-1}^d - C_0}{N} \quad (6-1)$$

$$sc_{di} \geq g_{dt} = \frac{C_{\min,t}^d - C_0}{N+1} \quad (6-2)$$

$$\overline{sc}_{di} = \frac{1}{2}(g_{d,t-1} + g_{dt}) \quad (6-3)$$

Note that for each individual, only one of the multidimensional recognized search costs can be perceived from the empirical data. Therefore, a subsequent regression analysis for all survey subjects and all dimensions needs to be estimated in order to empirically model search cost. We specify the search cost model in dimension  $d$  as:

$$sc_{di} = \beta_0 + \beta_1 C_0 + \beta_2 \text{gender} + \beta_3 \text{fixedsch} + \beta_4 \text{purpose} + \beta_5 \text{income1} + \beta_6 \text{income2} + \beta_7 \text{income3} + \beta_8 \text{distance} + \beta_9 \text{peak} + \beta_{10} \text{veh} + \varepsilon_i \quad (7)$$

where  $C_0$  is the generalized cost for the originally reported travel experience and distance measures the mileage that the subject travels. Dummy variables include gender (equals to 1 if the subject is female); fixedsch (equals to 1 if the subject has fixed travel schedule); purpose (equals to 1 if the trip purpose is work/school); peak (equals to 1 if the travel is in peak-hour periods); and veh (equals to 1 if household number of vehicles is greater than 2). Different household annual income levels are considered in the model (income1: less than \$50,000; income2: \$50,000 - 100,000; income3: \$100,000 - \$150,000; income4: \$150,000 and above). In our model,  $C_0$  is identified as an instrumental variable (IV) in order to better incorporate the sufficiently high correlation between  $C_0$  and other independent variables. We employ the generalized method of moments (GMM) and a two-stage least-squares (2SLS) estimator. Denoting the IV as  $\mathbf{z}$  and the independent variables as  $\mathbf{x}$ , we can estimate parameters  $\boldsymbol{\beta}$  from the population moment conditions:

$$E[\mathbf{z}(sc_{di} - \mathbf{x}\boldsymbol{\beta})] = 0 \quad (8)$$

The estimation result is reported in Table 3.1. The search cost is positively related to the initially experienced generalized cost of the travel. Lower-income agents have higher search costs along the mode dimension. Female agents are more reluctant to search departure times and routes than alternative modes. Fixed schedule and traveling during peak-hour increase the search cost for all dimensions. Travel distance has a negative impact on search cost; the longer the travel distance, the more likely she/he will search for alternatives. The coefficients for trip purpose indicate that commuting agents have more incentive to search for alternative modes and departure times. By estimating and applying search cost models, one can make personal/household characteristics endogenous in the search process and model a diversified and behaviorally rich multidimensional search. It helps explain why some travelers may adjust routes first while others may adjust departure time first in response to the same stimulus. This feature can potentially provide a rich

level of details, especially for policy/social equity analysis, whence measuring the impacts/benefits by different socio-economic strata of society is of interest.

**Table 3.1 Multidimensional Perceived Search Cost Models (Generalized Method of Moments and Instrumental Variable)**

<b>Models:</b>	<b>Search Cost (<i>d</i>: mode)</b>	<b>Search Cost (<i>d</i>: departure time)</b>	<b>Search Cost (<i>d</i>: route)</b>
<b>Variables</b>	Coefficients (std. err.)	Coefficients (std. err.)	Coefficients (std. err.)
<b>Generalized cost <math>C_0</math></b>	0.023 (0.010)***	0.008 (0.001)***	0.001 (0.000)***
<b>Gender (female)</b>	0.014 (0.088)	0.162 (0.071)**	0.098 (0.046)***
<b>Fixed schedule</b>	0.118 (0.065)***	0.194 (0.080)**	0.115 (0.045)***
<b>Purpose (work/school)</b>	-0.101 (0.062)*	-0.091 (0.056)*	0.098 (0.048)**
<b>Annual income (&lt; \$50k)</b>	0.188 (0.106)*	-0.272 (0.201)	-0.299 (0.060)***
<b>Annual income (\$50k-\$100k)</b>	0.085 (0.41)**	-0.285 (0.203)	-0.207 (0.066)***
<b>Annual income (\$100k-\$150k)</b>	-0.007 (0.007)	-0.542 (0.234)**	-0.089 (0.086)
<b>Travel distance (10 mi)</b>	-0.020 (0.003)***	-0.008 (0.001)***	-0.006 (0.000)***
<b>Peak-hour travel</b>	0.161 (0.094)*	0.112 (0.062*)	0.010 (0.041)
<b>Number of Cars (&gt; 2)</b>	-0.088 (0.021)***	0.298 (0.092)***	-0.035 (0.053)
<b>Constant</b>	1.341 (0.148)***	0.402 (0.225)*	0.384 (0.068)***

\*, \*\*, \*\*\* - significant at 90%, 95%, 99% confidence level

It is hypothesized that agents will search the most rewarding dimension with the highest search gain/cost ratio. Consecutive unrewarding searches along a particular behavioral adjustment dimension (e.g., route) will lead to diminishing subjective search gain for that dimension and, at a later point, cause the search to shift to a different behavior dimension (e.g., departure time). Once the ratios for all dimensions drop down below one, the multidimensional search process ceases. Since  $g_{dt}$  is monotonically decreasing and converges to zero, the search is guaranteed to reach stability. The interplay of these search gains and costs along all feasible behavioral dimensions defines the bounded rationality embedded in the theory. It collectively determines the prospects for profitable searches over finite horizon and guarantees a convergence of behavioral changes. It quantitatively theorizes when individuals start seeking behavioral changes, how they initially change behavior, how they switch behavioral adjustment dimensions, and when they stop the search.

### **3.1.3 Modeling Search and Decision Rules**

An agent will keep the status quo and repeat her/his habitual behavior once she/he decides not to search in any dimension. Once an agent determines a dimension to search, a search process is invoked to find useful alternatives to meet travel demand. A spatial/temporal search is not random and can be biased (Humphreys and Whitelaw 1979). For instance, if a person currently departs at 8 AM and is not satisfied with the resulting travel and/or schedule delay, the person is more likely to try departing at 7:30 AM or 8:30 AM than at 7 AM or 9 AM (i.e., an anchor effect). Different knowledge-extracting technologies can be applied to mine individuals' search rules and decision rules. Here, we adopt production rules for shorter-term departure time search and route search. For longer-term travel mode search, the process of identifying an alternative mode is theorized as a hidden Markov process.

Production rules provide a roadmap from a variety of arguments (i.e., travel time, cost, socio-economic characteristics) to the behavioral outcomes (i.e., searching alternatives and decisions).

The mapping arguments are denoted as  $\text{arg}_i$  and the behavior response alternatives as  $A_j$ . One typical heuristic rule set can be described as follows:

$$\text{Rule 1: if } \{\text{arg}_1 \in S_1^1\} \cap \{\text{arg}_2 \in S_2^1\} \cap \dots \cap \{\text{arg}_m \in S_m^1\}, \text{ then choose } A_1 \quad (9\text{-a})$$

$$\text{Rule 2: if } \{\text{arg}_1 \in S_{12}\} \cap \{\text{arg}_2 \in S_{22}\} \cap \dots \cap \{\text{arg}_m \in S_{m2}\}, \text{ then choose } A_2 \quad (9\text{-b})$$

⋮

$$\text{Rule } n: \text{ if } \{\text{arg}_1 \in S_{1n}\} \cap \{\text{arg}_2 \in S_{2n}\} \cap \dots \cap \{\text{arg}_m \in S_{mn}\}, \text{ then choose } A_n \quad (9\text{-c})$$

Without loss of generality, assume the model considers  $m$  mapping arguments and  $n$  heuristic rules. Then,  $\text{arg}_i$  represents the  $i$  the argument in the mapping;  $S_i^j$  represents a subspace of the entire space domain of the argument  $i$  (denoted as  $S_i$ ) for the Rule  $j$  ( $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ ). This entire space domain is split in a way that the rule set is capable of responding to all combinations of  $\text{arg}_i$  ( $i = 1, 2, \dots, m$ ). At the same time, each combination of arguments has only one corresponding heuristic rule. This condition can be formulated as:

$$\bigcup_j S_i^j = S_i \quad (10)$$

$$S_i^a \cap S_i^b = \emptyset, \forall a, b = 1, 2, \dots, n, \text{ and } a \neq b \quad (11)$$

Subjects' actual departure time-changing behaviors under uncertainty are observed from the departure time survey, where a specific travel time duration with unexpected delay is given to each respondent in each stated-preference (SP) scenario. This allows the research team to extract decision rules under uncertainty with machine-learning algorithms. Different machine-learning algorithms are applied to derive these heuristic rules. The team has tested four proven algorithms, including C4.5 (Quinlan, 1986), PRISM (Cendrowska, 1987), RIPPER (Cohen, 1995), and PART (Frank et al. 1998). The C4.5 algorithm seeks an attribute to split on; it maximizes the separation between classes in each step of developing a rule set, based on information gain ratios. The PRISM (Programming in Statistical Modeling) algorithm optimizes the best combination of attribute values based on error ratios. The RIPPER (Repeated Incremental Pruning to Produce Error Reduction) algorithm employs a greedy approach with an incremental rule pruning method to generate and revise the rule set. PART (partial decision tree) algorithm combines the advantages of C4.5 and RIPPER to generate a concise rule set without loss of classification accuracy. The empirically derived departure time search (DTS) rule set consists of nine rules, presented below. PART estimation is selected based on predictive accuracy of the derived search rules on the validation dataset.

Search 60+ min earlier, if [ASDL > 70]	DTS Rule 1
Search 30-60 min earlier, if [45 < ASDL <= 70]	DTS Rule 2
Search 0-30 min earlier, if [ASDL > 0 AND Delay > 0]	DTS Rule 3
Search 0-30 min later, if [0 < ASDL <= 30 AND Delay > 40%]	DTS Rule 4

OR[ASDL <= 10 AND ASDE <= 40 AND Delay <= 50% AND TT <= 65]	DTS Rule 5
Search 30-60 min later, if [ASDL = 0]	DTS Rule 6
Search 60+ min later, if [ASDE > 75]	DTS Rule 7
OR[ASDE > 45 AND Delay > 10%]	DTS Rule 8
Otherwise, search 0-30 min earlier.	DTS Rule 9

DTS Rule 1 states that individuals will consider shifting their departure times earlier by more than one hour if their experienced arrival schedule delay is over 70 minutes. All other rules can be similarly interpreted. These rules collectively replicate the heuristics individuals use to identify alternative departure times based on their current experiences and knowledge. As knowledge is updated during the search process, the same rule set can generate different alternatives for the same individual.

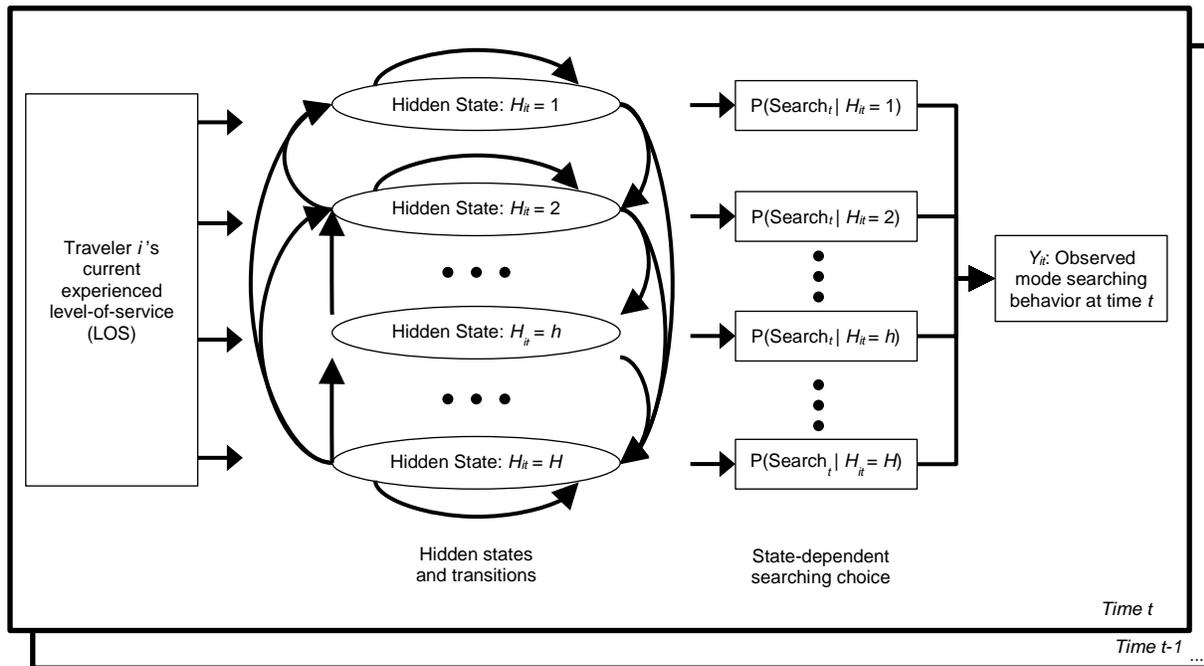
Similar to the search rules of an alternative departure time, if-then rules are selected to represent route searching heuristics. RIPPER is chosen for its better predictive performance on the dataset. The final route search rule set derived from the survey data consists of 16 rules for selecting access points and 13 rules for selecting egress points on the basic network. The two rule sets are quite similar, and therefore only the rules for selecting access points are shown in disjunctive normal forms, where “ $\Delta$ ” denotes changes or percentage changes (Route B attributes – Route A attributes); “[ ]” constitutes a complete antecedent condition of the if-then rules; “Time” is the total travel time; “Btime” is travel time on the basic network; and “Transfer” is the number of transfers between different levels of roads. Route Search (RS) Rules are listed below:

Choose Route A as the alternative route for consideration, if [ $\Delta$ Time = (0.21 ~ infinity)]	RS Rule 1
Or [ $\Delta$ Time = (0.13 ~ 0.21) And $\Delta$ Btime = (-infinity ~ -0.57)]	RS Rule 2
And $\Delta$ Btime = (-0.57 ~ 0.19) And $\Delta$ Transfer = 0 or 1]	RS Rule 3
And Time = (30 minutes ~ infinity)]	RS Rule 4
And $\Delta$ Btime = (0.19 ~ infinity) And $\Delta$ Transfer = 1]	RS Rule 5
Or [ $\Delta$ Time = (0.04 ~ 0.13) And $\Delta$ Btime = (-infinity ~ -0.57)]	RS Rule 6
And $\Delta$ Btime = (-0.57 ~ 0.19) And ( $\Delta$ Transfer = 0 or 1)]	RS Rule 7
And $\Delta$ Btime = (-0.57 ~ -0.19) And ( Time = (15 ~ 30 minutes)]	RS Rule 8
And $\Delta$ Btime = (0.19 ~ 0.57) And ( $\Delta$ Transfer = 1)]	RS Rule 9
Or [ $\Delta$ Time = (-0.04 ~ 0.04) And $\Delta$ Btime = (-infinity ~ -0.57)]	RS Rule 10
And $\Delta$ Btime = (-0.57 ~ 0.19) And $\Delta$ Transfer = 1]	RS Rule 11
And $\Delta$ Btime = (-0.57 ~ -0.19) And ( $\Delta$ Transfer = 0)]	RS Rule 12
Or [ $\Delta$ Time = (-0.21 ~ -0.04)	

And $\Delta B_{time} = (-\infty \sim -0.57]$	RS Rule 13
And $\Delta B_{time} = (-0.57 \sim 0.19)$ And $(\Delta Transfer = 1)$	RS Rule 14
And $\Delta B_{time} = (-0.57 \sim -0.19)$ And $(\Delta Transfer = 0)$	RS Rule 15
Otherwise, choose Route B as the alternative route for consideration.	RS Rule 16

For instance, RS Rule 1 suggests that drivers will identify a specific route in a round of searches if its travel time is significantly lower (21 percent) than other routes. The execution of this rule alone can exclude a large percentage of feasible alternative routes from further considerations. As the travel time difference becomes less apparent (RS Rule 2 to 15), other factors related to the simplicity of routes—such as the percentage of travel on the basic network and number of transfers between different sub-networks—also play an important role. Collectively, these rules replicate the heuristics individuals use to identify alternative routes based on their existing spatial knowledge. As knowledge changes (e.g., a congested section in the network is learned), the same rule set can generate different routes (e.g., a new route bypassing the congested section). Repeated executions of these rules in each round of route search produce one alternative route for the subsequent decision step. Finally, when used to predict behavior, these rules can be executed as deterministic or probabilistic rules (based on their accuracy on the estimation/validation datasets).

Mode search is presented by a hidden Markov method (the main idea is illustrated in Figure 3.2), which emphasizes the dynamic linkages between time periods.



**Figure 3.2 A hidden Markov model of travel mode searching dynamics**

As displayed in Figure 3.2, two major components are highlighted in this model:

- Hidden states and transitions: The transition is triggered by the evolving travel experience. Starting from an initial state distribution (i.e., at time 1, the probability density function  $\Pr(H_{i1})$  that traveler  $i$  is in state  $H_{i1}$ ), a sequence of Markov chains is employed to express

the likelihood that the level-of-service (LOS) experiences of the habitual mode in the previous periods are strong enough to transition the traveler to another hidden state  $\Pr(H_{it} | H_{it-1})$ . For example, successively experiencing longer waiting time when using transit may cause the traveler to switch to an auto-loving state.

- State-dependent decision rules: Given the hidden state that a traveler  $i$  is in, the probability that she/he will identify mode  $y_{it}$  as the alternative in the mode searching stage at time  $t$  is determined by  $\Pr(Y_{it} = y_{it} | H_{it})$ .  $Y_{it}$  is the mode searching decision made by traveler  $i$  at time  $t$ .

An individual's decision probabilities are correlated through the underlying path of the hidden states  $(y_{i1}, y_{i2}, \dots, y_{iN})$  because of the Markovian properties of the model. Therefore, the joint likelihood function is given as:

$$\begin{aligned}
 L(H_{it}) &= \Pr(Y_{i1} = y_{i1}, \dots, Y_{iN} = y_{iN}) \\
 &= \sum_{H_{i1}} \sum_{H_{i2}} \dots \sum_{H_{iN}} \left[ \Pr(H_{i1}) \prod_{t=2}^N \Pr(H_{it} | H_{it-1}) \right] \cdot \prod_{t=1}^N \Pr(Y_{it} = y_{it} | H_{it})
 \end{aligned} \tag{12}$$

More details of this hidden Markov search model regarding model variables and estimation procedure can be found in Xiong and Zhang (2015).

After each search round, a new alternative is identified. Agents either change behavior to use the new alternative or stay with their habitual behavior. This is determined by a set of decision rules. Even though during the multidimensional search process many alternatives may be visited, the final decision is assumed to be the outcome of a series of switching decisions. Production rules derived by various machine learning algorithms (Quinlan 1986; Cendrowska 1987; Cohen 1995) are selected here to represent decision rules. Departing from random utility maximization, this assumption about the search-decision procedure relaxes the unrealistic assumption of human information processing and computational capabilities, and incorporates individual-based historical dependencies. It also improves the computational efficiency of agent-based simulation since the execution of production rules only requires minimum computational resources. These search and decision rules are empirically derived for each behavioral dimension.

For departure time choice (DTC), the travel time uncertainty (*RANGE*) is specified here as the 95 percent confidence interval of the travel time duration. Other explanatory variables in the decision rules include: preferred arrival time (*PAT*); departure time (*DT*); preferred departure time (*PDT*); travel time (*TIME*); monetary cost (*COST*); household income (*INCOME*); and gender (*GENDER*). The variable *flex* is a dummy variable that is equal to 1 if the trip maker's preferred arrival schedule is flexible, and 0 otherwise.  $\Delta$  denotes percentage changes of the alternative departure time attributes from the attributes of current departure time choice.

Switch to the alternative departure time choice (DTC), if

$[\Delta RANGE \leq -16.7\% \text{ and } \Delta TIME \leq -15.4\%]$	DTC Rule 1
$[\Delta TIME \leq -25\% \text{ and } \Delta RANGE \geq 0\%]$	DTC Rule 2
$[\Delta RANGE \geq 0\% \text{ and } \Delta COST \leq -35.2\% \text{ and } flex = 1]$	DTC Rule 3
$[\Delta RANGE \leq 0\% \text{ and } INCOME < 150K \text{ and } -8.3\% < \Delta COST \leq 0\% \text{ and } \Delta TIME \leq 10\%]$	DTC Rule 4

[ $\Delta RANGE \leq 0\%$ and $INCOME < 150K$ and $\Delta COST \leq -8.3\%$ and $\Delta ASDL \leq 35\%$ ]	DTC Rule 5
[-16.7% $\leq \Delta RANGE \leq 0\%$ and $INCOME \leq 50K$ ]	DTC Rule 6
[ $\Delta ASDL \leq -38\%$ and $\Delta RANGE \geq 0\%$ and $\Delta TIME \geq 17\%$ ]	DTC Rule 7
[-66.7% $\leq \Delta RANGE \leq -16.7\%$ and $-4.2\% \leq \Delta COST \leq 10\%$ ]	DTC Rule 8
[ $INCOME \leq 50K$ and $GENDER = female$ and $\Delta RANGE \leq -70\%$ ]	DTC Rule 9
[ $INCOME \leq 50K$ and $flex = 1$ and $-22.7\% \leq \Delta TIME \leq 16.6\%$ and $\Delta COST \leq 20\%$ ]	DTC Rule 10
[ $INCOME \leq 100K$ and $GENDER = female$ and $\Delta TIME \leq 8.3\%$ and $\Delta RANGE \leq -44.4\%$ ]	DTC Rule 11
[-21% $\leq \Delta TIME \leq -10\%$ and $\Delta ASDL \geq 33\%$ and $\Delta RANGE \leq -40\%$ ]	DTC Rule 12
Otherwise, continue to use the current departure time.	DTC Rule 13

The route changing rules for commute trips derived from machine learning algorithms are present below, where  $\Delta$  denotes changes or percentage changes (new route attributes – current route attributes), and absolute values are attributes of the currently used route.

#### Route choice change (RCC) rules, commute trips

Change route, if

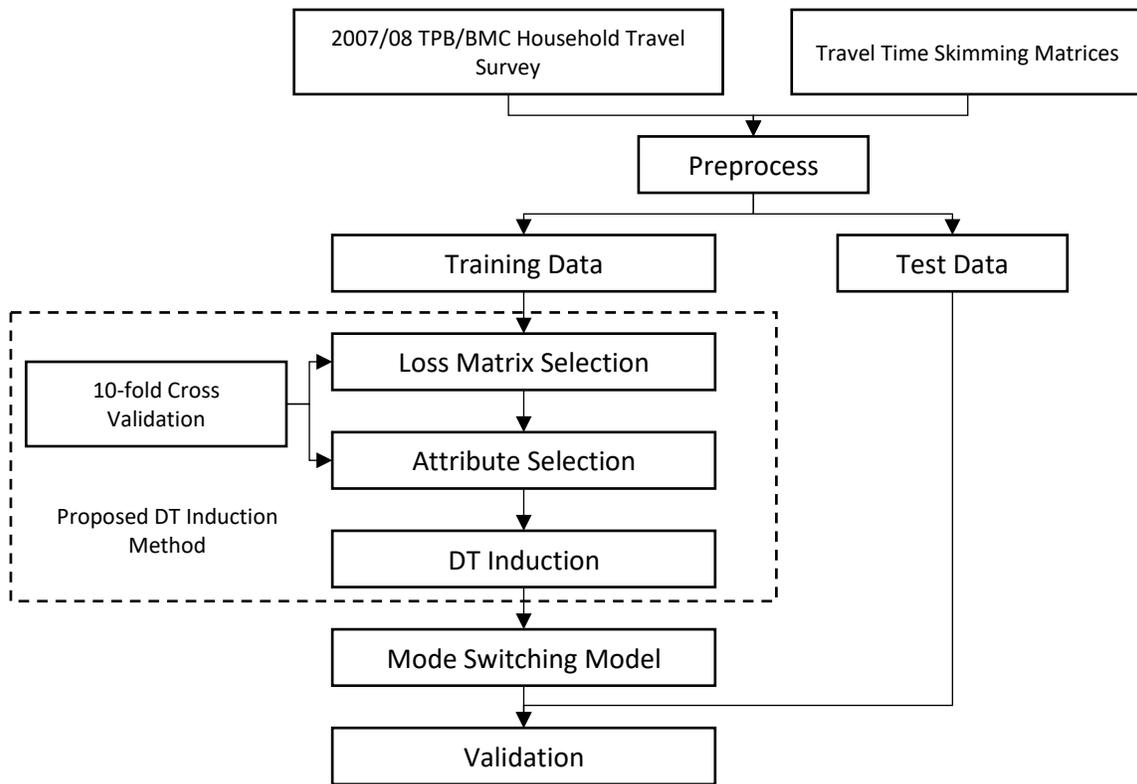
[ $\Delta Time \leq -39\%$ ]	RCC Rule 1
Or [ $\Delta Time \leq -11\%$ and $\Delta Pleasure \geq -1$ ]	RCC Rule 2
Or [ $\Delta Familiarity \geq 3$ and $Commute\ time \leq 20\ min$ ]	RCC Rule 3
Or [ $\Delta Time \leq 6\%$ and $\Delta Pleasure \geq 3$ ]	RCC Rule 4
Or [ $\Delta Time \leq 15\%$ and $\Delta Familiarity \geq 2$ and $\Delta Delay \geq -40\%$ ]	RCC Rule 5
Or [ $Familiarity = 1$ and $\Delta Time \leq 51\%$ and $Commute\ time \leq 20\ min$ and $Income = 1$ ]	RCC Rule 6
Or [ $Delay \geq 4\ min$ and $\Delta Stops \leq 0$ and $Commute\ distance \leq 8\ miles$ ]	RCC Rule 7
Or [ $\Delta Pleasure \geq 2$ and $\Delta Familiarity \geq 0$ and $Commute\ time \leq 16\ min$ ]	RCC Rule 8
Otherwise, continue to use the current route.	RCC Rule 9

A perception threshold exists in route changing behavior. For instance, RCC Rule 1 implies drivers will change routes as long as travel time can be reduced by more than 39 percent. Travel time reduction less than 11percent is insignificant. Variable “familiarity” is present in several rules, which is evidence of historical dependencies in route choice. A comparison shows that the predictive performance of the route changing rule set is superior to a normative logit model using measures such as hit ratios at the individual level and route flows at the aggregate level (see Zhang 2006a, b for details).

Decision tree (DT) is used to explore the underlying rules of travelers’ travel mode decision between the habitual mode and the newly identified alternative mode. This section proposes a DT induction method which can efficiently handle the class imbalance issues. A longitudinal travel survey is a good method for exploring the rules of people’s model switching behavior, by observing people’s mode shift behavior and using that data to build mode switching model. However, this type of data is relatively hard to collect. In this study, Household Travel Survey data (which is much easier to obtain) is used to generate mode switching decision tree models and to show the benefits of proposed method.

Figure 3.3 shows the analysis algorithm proposed in this study. Data used in this study include the 2007/2008 TPB Household Travel Survey and Zone-to-zone Travel Time Skimming Matrices which, together, provide the trip and socio-demographic information necessary for mode choice analysis. After preprocessing, the two data sets are integrated to a single file and then randomly separate into two parts: training data set and test data set. Mode switching models are then built using training data set by the proposed DT induction method, which includes three steps: loss matrix selection, attribute selection and DT induction. Loss matrix is introduced to handle the class imbalance issues. Attribute selection is implemented to select a subset of relevant features for model construction. Ten-fold cross validation is used as a goodness-of-fit in loss matrix selection and attribute selection. Then, the mode switching models are validated with the test data set.

The key procedures of the DT induction method, loss matrix selection, attribute selection and DT induction will be discussed in the remainder of this section. Data and preprocess are described later in the report.



**Figure 3.3 Mode choice decision tree analysis procedure**

Ten explanatory variables are assumed to have significant effects on travelers' mode choice behavior, representing the characteristics of three categories: household properties, personal attributes and trip characteristics. Household properties include: housing tenure (*TENURE*) which includes owned, rented; household income (*INCOME*) which is divided into four cohorts; and number of vehicles (*VEH*) and bikes (*BIKES*) in the household, which takes value as 0, 1, 2 or 3+. Personal attributes contain age (*AGE*), employment (*EMPLY*) and driver license ownership (*LIC*). Trip characteristics are travel time (*TT*), trip distance (*DIST*) and trip purpose (*PURP*). Trip

purpose includes home, work, shop, school, passenger drop off/ pick up and other. Except for the model between driving and passenger, a new variable, unit travel time difference ( $\Delta TIME$ ,  $\Delta TIME = (TT_1 - TT_2) / DIST$ ), is defined to show the difference in the level of service between two modes, which also helps alleviate the fragmentation problem. Data cleaning is then applied to remove incomplete data. The explanatory variables are shown in Table 3. 2.

**Table 3.2 Explanatory Variables**

<b>Variable</b>	<b>Definitions</b>	<b>Values</b>
<i>TENURE</i>	Housing tenure	1 =Owned; 2 =Rented
<i>INCOM</i>	Household income	1=Less than \$50,000; 2=\$50,000 - \$99,999; 3=\$100,000 - \$149,999; 4 =\$150,000 or more
<i>VEH</i>	Number of vehicles in the household	0 = 0; 1 = 1; 2 = 2; 3 = 3+
<i>BIKES</i>	Number of bicycles in the household	0 = 0; 1 = 1; 2 = 2; 3 = 3+
<i>AGE</i>	Age group	0 = 0-4; 1 = 5-15; 2 = 16-18; 3 = 19-24; 4 = 25-34; 5 = 35-44; 6 = 45-54; 7 = 55-64; 8 = 65-74; 9 = 75+
<i>EMPLY</i>	Currently employed?	1=YES; 2=NO; -9 = Not Applicable
<i>LIC</i>	Have driver license?(Persons 16+)	1=YES; 2=NO; -9 = Not Applicable
<i>ΔTIME</i>	Difference in travel time divided by distance	Continuous (min/mile)
<i>TT</i>	Reported travel time	Continuous (min)
<i>DIST</i>	Estimated trip distance (Rounded to nearest 0.1 mile)	Continuous (mile)
<i>PURP</i>	Trip purpose	01 = Home; 02 = Work; 04 = Shop; 08 = School; 09 = Other; 11 = Passenger Drop off/Pick up

After the pre-process, a total of 72,536 observations are available for model construction. The data are then randomly separated in two parts: one contains 80 percent of the data as the training set, and the remaining 20 percent as the subsequent validation test set. Table 3.3 shows the composition information in different travel modes for the two groups of data. The percentages of the four modes, transit, driving, passenger in a car and walk/bike are 7.54 percent, 67.07 percent, 21.53 percent and 3.86 percent, respectively. It is necessary to apply the loss matrix method to handle the huge discrepancy in number of observations between driving and other modes.

**Table 3.3 Mode Splits for the Data Set**

<b>Mode</b>	<b>Total database</b>		<b>Training dataset</b>		<b>Test data set</b>	
	<b>Number</b>	<b>%</b>	<b>Number</b>	<b>%</b>	<b>Number</b>	<b>%</b>
<b>Transit</b>	5,467	7.54	4,368	7.53	1,099	7.58
<b>Driving</b>	48,653	67.07	38,995	67.20	9,658	66.57
<b>Passenger</b>	15,616	21.53	12,427	21.42	3,189	21.98
<b>Walk/Bike</b>	2,800	3.86	2,239	3.86	561	3.87
<b>Sum</b>	72,536	100	58,029	100	14,507	100

Four modes are considered in this study: transit, driving, vehicle passenger and walk/bike. To explore the underlying rules of travelers' switching decisions, DT models are built between the habitual mode and the newly identified alternative; as a result, a total of six DTs are established.

After tuning the loss weight ratio and feature set size, six DT models are extracted between any two modes using training data from the C4.5 algorithm. The test data set is used to validate the DT models. The chosen loss weight ratio  $R$  and the feature set size, as well as the performance of the DT classifiers, are shown in the left part of Table 3.4. Logit models are built as a comparison; the model performance is shown in the right part of the table. Three measures of success are presented in the table, including prediction accuracy ( $PA$ ), hit ratio ( $HT$ ) and the aggregate match rate ( $AMR$ ). Prediction accuracy and hit ratio, which reflect the modeling performance on an individual level, have been introduced and used in loss matrix selection and attribute selection. The aggregate match rate reflects the prediction accuracy on the mode aggregate level, defined as the ratio of the number of predicted observations (including correctly and incorrectly predicted observations) for one mode over the number of actual observations for the mode. The confusion matrices of the DT models are also included in the table. In the confusion matrix, each row shows the actual observations in the mode and each column shows the predicted number of observations in the mode. The diagonal elements give the number of correctly predicted observations of each mode and the non-diagonal elements give the number of misclassified observations of each mode. The team will first analyze the performance of the DT models and compare with logit models later.

The DT models perform very well in prediction accuracy; most of the models have prediction accuracy over 90 percent, with 91.31 percent for the model between transit and driving, 91.79 percent for transit and passenger, 93.92 percent for transit and walk/bike, 98.89 percent for driving and walk/bike and 97.63 percent for passenger and walk/bike. With an 84.01 percent prediction accuracy, the DT for driving and passenger performs the worst. This is acceptable because drivers and passengers share a lot of attributes, and behavior patterns are relatively harder to classify. The high value of the hit ratio and the aggregate match rate clearly shows that the loss matrix can effectively address the class imbalance issues. Although there are many more observations in driving, the DTs successfully predict the minority classes by adopting a suitable loss weight ratio. For the model between transit and driving, the hit ratios are 67.79 percent and 93.98 percent, respectively. For driving and passenger, the hit ratios are 88.19 percent and 71.37 percent, respectively. For driving and walk/bike, the hit ratios are 99.57 percent and 87.34 percent, respectively. The DTs perform very well in the aggregate match rate, with most values between 90 percent and 110 percent.

Logit models are developed and compared with the DT models to evaluate the performance of the proposed DT induction method. Linear utility functions are assumed for logit models:

$$U_m = \beta_{0,m} + \sum_{i=1}^n \beta_{i,m} x_{i,m} + \varepsilon_m \quad (13)$$

Where  $U_m$  is the utility of mode  $m$ .  $\beta_{0,m}$  denotes mode-specific constant.  $\beta_{i,m}$  is the coefficient of the  $i$ th explanatory variable in the utility function of mode  $m$ .  $x_{i,m}$  is the  $i$ th explanatory variable of mode  $m$ .  $\varepsilon_m$  represents the random component of the utility, which is assumed as an independently distributed random variable with a Gumbel distribution.

The same training and testing data are applied to estimate and validate the logit models. In order to eliminate the effects of using the loss matrix, the training data are also oversampled using the same loss matrix to estimate logit models. All 10 explanatory variables are used in estimations, with some dummy variables created to represent the category variables in the data set. For travel time, estimated travel time for each mode is used instead of the unit travel time difference. The measures of goodness of fit, i.e., prediction accuracy, hit ratio and the aggregate match rate, are calculated and compared with DT models, as shown in Table 3.4.

**Table 3.4 Performance of DT Models and Logit models**

Model	Loss Weight Ratio	Feature Set Size			Size of Tree	PA	PA (Logit)
T and D	3	10			4087	91.31%	90.70%
T and P	1	8			414	91.79%	87.31%
T and W	1	10			264	93.92%	87.95%
D and P	3	10			264	84.01%	90.52%
D and W	2	3			490	98.89%	96.52%
P and W	2	1			517	97.63%	88.64%
	<b>Confusion Matrix</b>	<b>T</b>	<b>D</b>	<b>HT</b>	<b>AMR</b>	<b>HT (Logit)</b>	<b>AMR (Logit)</b>
T and D	T	745	354	67.79%	120.66%	45.40%	81.80%
	Driving	581	9077	93.98%	97.65%	95.86%	102.07%
	<b>Confusion Matrix</b>	<b>T</b>	<b>P</b>	<b>HT</b>	<b>AMR</b>	<b>HT (Logit)</b>	<b>AMR (Logit)</b>
T and P	T	945	154	85.99%	104.00%	72.07%	93.63%
	P	198	2991	93.79%	98.62%	92.57%	102.20%
	<b>Confusion Matrix</b>	<b>T</b>	<b>W</b>	<b>HT</b>	<b>AMR</b>	<b>HT (Logit)</b>	<b>AMR (Logit)</b>
T and W	T	1068	31	97.18%	103.55%	89.72%	97.63%
	W	70	491	87.52%	93.05%	84.49%	104.63%
	<b>Confusion Matrix</b>	<b>D</b>	<b>P</b>	<b>HT</b>	<b>AMR</b>	<b>HT (Logit)</b>	<b>AMR (Logit)</b>
D and P	D	8517	1141	88.19%	97.64%	98.84%	110.29%
	P	913	2276	71.37%	107.15%	65.32%	68.83%
	<b>Confusion Matrix</b>	<b>D</b>	<b>W</b>	<b>HT</b>	<b>AMR</b>	<b>HT (Logit)</b>	<b>AMR (Logit)</b>
D and W	D	9616	42	99.57%	100.30%	99.87%	103.42%
	W	71	490	87.34%	94.83%	38.86%	41.18%
	<b>Confusion Matrix</b>	<b>P</b>	<b>W</b>	<b>HT</b>	<b>AMR</b>	<b>HT (Logit)</b>	<b>AMR (Logit)</b>
P and W	P	3153	36	98.87%	100.53%	93.10%	99.56%
	W	53	508	90.55%	96.97%	63.28%	102.50%

\* T, D, P, W is short for the four modes, transit, driving, passenger in a car and walk/bike respectively

For the overall prediction accuracy, the DT models outperform the logit models in five out of six cases. Only for the switching between driving and passenger does the logit model perform slightly better. In terms of the hit ratio, while the two methods both predict the majority modes at high accuracy rates, DT models perform much better in predicting the switching to minority modes, especially the walk/bike mode and transit. A good example is the switching model for driving and walk/bike; the hit ratio accuracy for walk/bike is improved from 38.86 percent (the logit model)

to 87.34 percent (DT model). The switching choices from driving to car passenger and transit are also predicted more accurately by the proposed DT models.

As for the aggregate match ratio, for driving and passenger, the aggregate match ratio for passenger is 68.83 percent and 107.15 percent for the logit and DT models, respectively. For driving and walk/bike, the aggregate match ratio for walk/bike is 41.18 percent and 94.83 percent for the logit and DT models, respectively. The DT models outperform the logit models in the two switching models, which also show a strong ability of handling the class imbalance issues with loss matrix. Otherwise, they perform similarly.

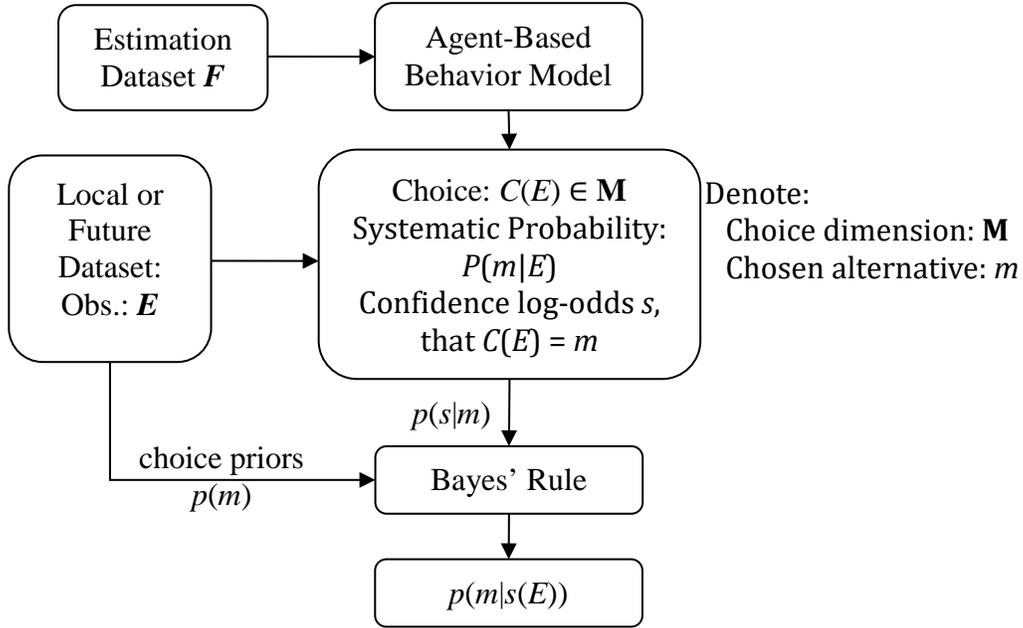
In comparison, DT models outperform logit models in most cases at both an individual prediction level and an aggregate prediction level. The DT models can classify the minority classes very well, demonstrating that they can better handle the class imbalance issues using the proposed decision tree induction method, compared to the logit models. Aside from a high accuracy and the ability to classify the minority classes using the loss matrix, there are many other benefits to employing DT models. They have high flexibility in the model specification. No IIA property needs to be assumed. Since DT comprises if-then rules, it can be easily interpreted and has high estimation efficiency (Xie et al. 2003). However, DT also has its shortcomings. In the Maryland team's DT models, some continuous variables are parsed too fine. For the model between passenger and walk/bike, only one variable—the unit travel time difference—is used to construct the DT, which contains 259 leaves. Discretization can be implemented before DT induction to alleviate this problem. In addition, the DT models can't capture the correlations among attributes or rule sets.

To summarize, this section documents the entire agent-based travel behavior modeling approach. Multidimensional decision-making process is modeled in a way that search gain and search cost interplay and thereby determine which dimension travelers will search and make behavior changes. This gain-cost mechanism also helps to determine the stopping criteria for the search. Various production rules are employed to explain searching and decision-making along each behavioral dimension. The aforementioned modeling components are estimated based on empirical observations collected via surveys. Validation/calibration of these models are discussed in the next section.

## **3.2 SILK AgBM: MODEL CALIBRATION AND VALIDATION**

### **3.2.1 Calibrating the Spatial Distribution of Travel**

A more flexible recalibration approach to gauge the spatial distribution of travel is proposed in this section. The proposed method relaxes the rigidity of using scale/constant parameters in conventional approaches. The proposed method directly recalibrates the systematic probabilities, which could potentially enhance the accuracy. The modeling flowchart is sketched in Figure 3.4.



**Figure 3.4 Bayesian calibration for travel demand model transfers**

The estimated model system predicts systematic probabilities based on a local or future-year dataset where the system is applied. Different probability values are converted into confidence scores for different alternatives (indicating the strength of the decision that the empirical observation chooses, alternative  $m$ ). Note that the score may not necessarily match the local data well. Typically, this score is represented by log-odds defined using the equation below. This measurement transfers the original scale of choice probability (i.e.,  $[0,1]$ ) to a space  $[-\infty, +\infty]$  where different continuous distributions are applicable.

$$s_m(E) = \log \frac{p(m)}{(1-p(m))} \quad (14)$$

The log-odd statistics and predicted probabilities may not match the observed probabilities. To recalibrate the model, it is necessary to perform a mapping of the model predictions to actual observations. This mapping is represented by a series of posterior calibration functions based on Bayes' Rule. Conditioned on each chosen alternative, a probability density estimator  $f$  is produced.

$$p(s|m) \sim f(s) \quad (15)$$

Here, various distributions can be employed to approximate the distribution  $f$ . For enhanced performance, asymmetric posterior functions can be adopted to improve accuracy. For instance, an asymmetric Laplace distribution can be produced by mixing two exponentials:

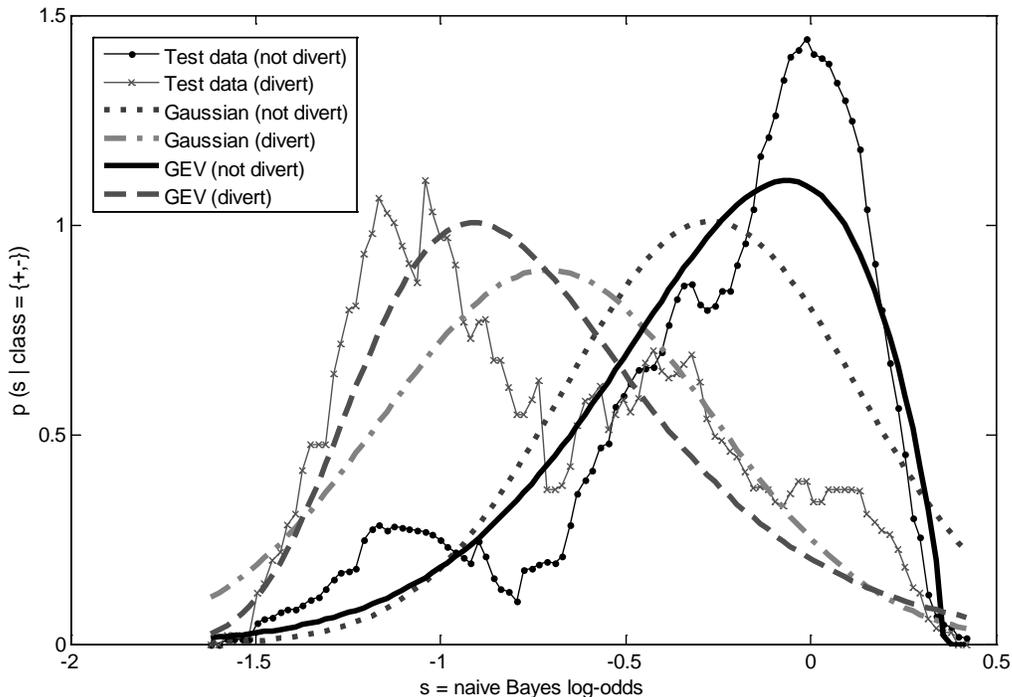
$$p(s|\theta, \alpha, \beta) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} \exp[-\alpha(\theta-s)] & s \leq \theta \\ \frac{\alpha\beta}{\alpha+\beta} \exp[\beta(\theta-s)] & s > \theta \end{cases} \quad (16)$$

These functions map score  $s$  to the actually observed probabilities. In other words, the family of posterior calibration functions adjusts the utilities estimated by the mode choice model in order to approximate the unobserved distribution of utility values for the local/future year dataset. Then, Bayes' Rule and the choice priors are used to obtain the estimate:

$$p(m|s) = \frac{p(m)p(s|m)}{\sum_{C \in \mathbf{M}} p(C) \cdot p(s|C)} \quad (17)$$

This spatial calibration method is then applied to recalibrate the AgBM behavioral modeling components: the routing and travel mode decision probabilities. The data sources used in the calibration and the calibration results are documented below.

In order to calibrate the routing decisions, real-world field observations on an often-congested commuting corridor are collected as the testing dataset. For the calibration function of the class-conditional densities, a Gaussian and a generalized extreme value (GEV) are fit to each of the class-conditional densities using the usual maximum likelihood estimates. The fits of these two functions represent a qualitative comparison between using symmetric distributions and using asymmetric distributions to approximate the class-conditional densities. Figure 3.5 shows the calibration function fits produced by these methods versus the testing behavior data. Performance measures are also offered for quantitative measurements.



**Figure 3.5 Estimated class conditional score densities versus the actual densities of the testing dataset**

In general, the calibration results agree with the empirical observation. The average value for the naïve Bayes log-odds is approximately -0.5, which is consistent with the low diversion rates perceived from the testing dataset. In other words, the optimistic prediction estimated by the en-route diversion model is well captured and recalibrated by this Bayesian calibration process. For the diversion class (+), the test data curve plotted in Figure 3.5 skews towards the left side, as the en-route diversion model gives these observations higher probability estimates to divert. The opposite occurs for the not-divert class (-).

The calibration function maps the estimated probabilities (i.e., log-odd scores) to the actual-observed diversion rates. Now the evaluation of the calibration results is of concern. There are at least two types of performance measures that have been typically used in data mining to assess the quality of probability estimates, specifically, log-loss (Good 1952) and squared error (Brier 1950; DeGroot and Fienberg 1983). While a better score according to these rules actually means an overall improved prediction quality, it has occasionally been termed loosely as improving “calibration” (Bennett 2003).

The actual classification for an empirical observation  $E$  (with class  $C(E) \in \{+, -\}$ ) in the testing dataset is observed. Let  $\delta$  denote the Kronecker delta function which equals 1 if the two arguments are equal to each other, and 0 otherwise. The log-loss and the squared error ( $Error^2$ ) are defined as Equation 18 and 19, respectively.

$$\log loss = \delta(C(E), +) \log P(+|E) + \delta(C(E), -) \log P(-|E) \quad (18)$$

$$Error^2 = \delta(C(E), +) (1 - P(+|E))^2 + \delta(C(E), -) (1 - P(-|E))^2 \quad (19)$$

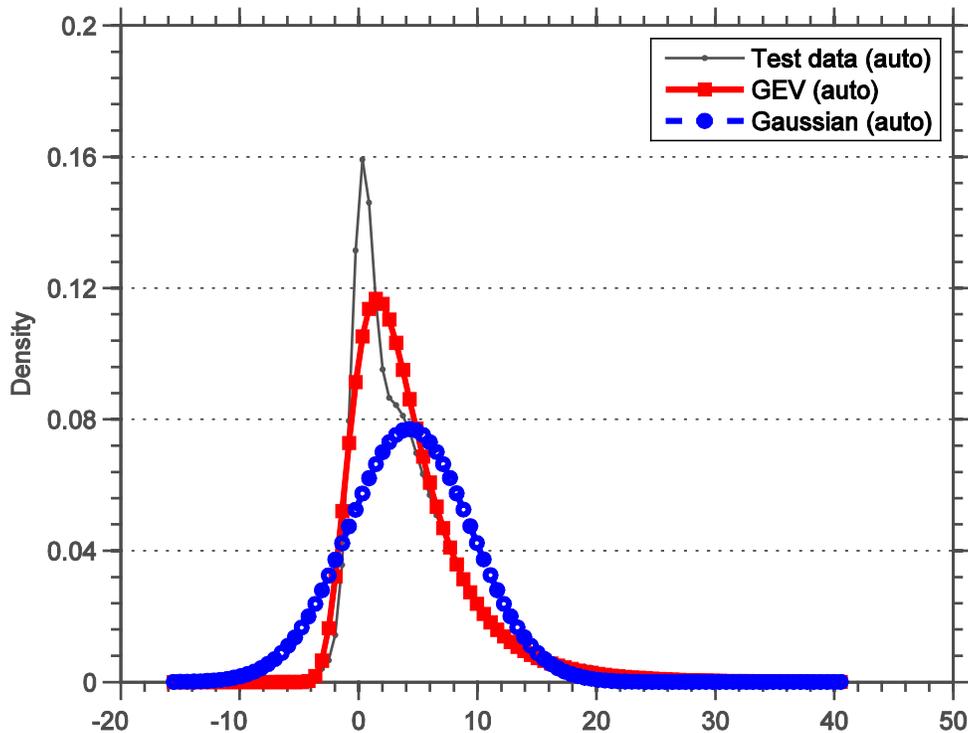
This research first reports the average log-loss and the mean squared error ( $MSE$ ) for the performance measure of the calibration. The results are given in Table 3.5. Both calibration functions result in significant improvement for the model’s prediction accuracy, as the average log loss statistic has been improved from  $-1.9909$  to  $-0.9750$  and  $-0.4792$ , respectively. The MSE has been reduced from  $0.2855$  to  $0.0921$  and  $0.0767$ , respectively. Overall, asymmetric distributions (GEV in this case) tend to be empirically preferable and outperform symmetric distributions in terms of prediction accuracy.

**Table 3.5 Results for Calibrating the AgBM Routing**

<b>Performance Measures</b>	<b>Naïve Bayes</b>	<b>Gaussian</b>	<b>GEV</b>
<b>Total Log-loss</b>	-2361.3	-1156.4	-568.34
<b>Avg. Log-loss</b>	-1.9909	-0.9750	-0.4792
<b>Total Squared Error</b>	338.66	109.34	90.947
<b>Mean Squared Error</b>	0.2855	0.0921	0.0767
<b>Predicting Accuracy</b>	0.5877	0.9039	0.9182

In order to calibrate the mode decision probabilities, TPB/BMC Household Travel Survey data is employed as the representative sample for the MITAMS study area. Two different continuous statistical distributions are applied in this research as Bayesian calibration functions: a symmetric Gaussian distribution and an asymmetric generalized extreme value (GEV) distribution. The calibration is conducted on each alternative specific utility function to perform a mapping from model predicted probabilities to actual observation. Log-odds are used to transfer the original scale of choice probability (i.e.,  $[0,1]$ ) to a space of  $[-\infty, +\infty]$  in order to suite different continuous distributions. Figure 3.6 illustrates the estimated posterior distributions for the calibration of probabilities for choosing drive-alone mode.

In the subfigures, the curves of the test data represent the actual densities of log-odds for choosing drive-alone mode. The curves are illustrated as nonparametric fixed-width kernels. In general, log-odds distribute over the positive spectrum and skew toward the left side, indicating a higher propensity of choosing drive-alone mode. The fits of the two statistical functions represent a qualitative comparison between using different distributions to approximate the conditional densities. The asymmetric GEV distributions seem to outperform the symmetric Gaussian distribution.



**Figure 3.6 Estimated mode choice conditional log-odd densities versus the actual densities of the test data (TPB/BMC Household Travel Survey)**

As mentioned earlier, the recalibration approach maps the estimated log-odd scores to the actually observed choice probabilities. Now that the concern surrounds how to evaluate this proposed

approach and compare it with existing approaches, such as transfer scaling. Measurements explained beforehand are employed here to assess the quality of the probability estimation. A better performance based on these measures can be termed improved “transferability” while it actually indicates overall higher prediction accuracy. Performance measures are reported in Table 3.6.

**Table 3.6 Performance Measures for Recalibrating the Mode Decision Rules**

	<b>AgBM</b>	<b>Transfer-Scaling</b>	<b>Gaussian</b>	<b>GEV</b>
<b>Avg. Log-loss</b>	-4.28	-2.85	-0.60	-0.48
<b>Total Log-loss</b>	-69399.20	-46220.76	-9751.87	-7831.75
<b>Mean Squared Error</b>	0.44	0.40	0.33	0.32
<b>Total Squared Error</b>	7161.50	6426.00	5321.36	5245.78
<b>Hit Ratio</b>	0.53	0.57	0.65	0.65

Based on the performance measures, both GEV and Gaussian calibration functions outperform the widely-used transfer-scaling method. The average log-loss and total log-likelihood statistics are reduced greatly in all four experiments. The mean squared error statistics are improved from near 0.40 to 0.29-0.34. Overall, asymmetric distributions (e.g., GEV) tend to be empirically preferable in terms of hit ratio and squared errors. It is worth noting that more favorable statistical distributions (e.g., compound distributions that mix two different types of distributions) can be further tested in order to enhance the goodness-of-fit.

### **3.2.2 Calibrating the Temporal Distribution of Travel**

Calibrating the temporal distribution of travel often involves a large number of parameters if discretizing the time frame into smaller time intervals. In this context, the Bayesian calibration approach is not the most suitable way for the calibration. In order to calibrate the temporal distribution of travel, a simulation-based optimization approach is proposed. The goal of the calibration is an optimization solving process to find the best decision variables so that the departure time profile of the behavioral user equilibrium (BUE) condition matches the real world temporal distribution data. Simulation-based optimization (SBO) is applied to achieve the calibration work.

Decision variables that are optimized here are specified as a set of mental cognitive functions. These functions are subjective weights that are used to weigh the past experiences for each agent. Representativeness measures the typical characteristics of a historical experience. For instance, if travel time at 7 AM is usually 30 min, but is 60 min on a particular day because of an incident, the experience of 60 min can be regarded as non-representative for the usual situation and ignored in perception updating. Recentness measures when an event takes place. More recent events receive higher weights. These two measurements can also reflect the subjective reliability of the transportation network.

$$w_t = \eta_t \cdot \zeta_t \tag{20-a}$$

$$\eta_t = \left(1 - \frac{u^* - u_t}{u^*}\right)^{\theta_1} \quad (20-b)$$

$$\zeta_t = \left(\frac{1}{N-t+1}\right)^{\theta_2} \quad (20-c)$$

$$u_{weight,N} = u_m, 1 \leq m \leq N, w_m = \max(w_1, w_2, \dots, w_N) \quad (20-d)$$

where  $w_t$  denotes the weight for day  $t$  among  $N$  days that has passed in the entire transportation experience. The weight is determined by a representativeness factor  $\eta_t$  and a recentness factor  $\zeta_t$ ;  $u^*$  is the theoretical maximum payoff level under free-flow travel condition, and assumes all individuals initially believe there is no congestion;  $u_t$  is the payoff on day  $t$ ;  $\theta_1$  and  $\theta_2$  are factors for representativeness and recentness. Thus,  $u_{weight,N}$  is the payoff of the day with highest weight (Eq. 20c).

SBO is an optimization method that mainly handles optimization problems that do not have an analytical solution and request simulation to obtain the value of objective function. In these optimization problems, the relationship between decision variables and objective function are not derivable directly through mathematical formulations. Once a vector of decision variables  $X = (x_1, x_2, \dots, x_k)$  is given, simulation approach is usually required to observe the performance of  $X$  on the objective variable  $Y$ . In SBO, a list of initial decision variables  $(X_1, X_2, \dots, X_N)$  will be tested to approximate a list of objective variables  $(Y_1, Y_2, \dots, Y_N)$ . These  $N$  initial sample points will be utilized through the SBO algorithm to regress the relation between  $X$  and  $Y$  to a number of particular parameterized basis functions. Once this relation is built, a new vector of decision variables  $X_{N+1}$  that may have a large potential to reach better objective function values will be generated. With the existing  $N$  initial points and 1 infill point, the relationship will be regressed once again to generate infill vector  $X_{N+2}$  and, iteratively, more infill vectors. The termination criterion of this infill process is often set as a maximum infill number ( $M$ ) of evaluations of objective function.

In this research, SBO is selected because the optimal solution cannot be obtained in analytical ways; the real-world travel demand modeling has to rely on traffic assignment, simulation software, or other traffic models. Once the real-world traffic network and an initial demand have been obtained, the model will simulate every traveler's travel experience and estimate their departure time switching process by the AgBM model system. The new departure times will be regarded as time-dependent demand. The process will be repeated until it reaches the BUE.

Since the search cost, search gain and decision rules have been validated in previous sections, this SBO calibration main considerations are representativeness factor ( $\theta_1$ ) and recentness factor ( $\theta_2$ ). There are 5 parameters to calibrate in this section:

- Representativeness factor (Commute Trips), range from 0 to 1

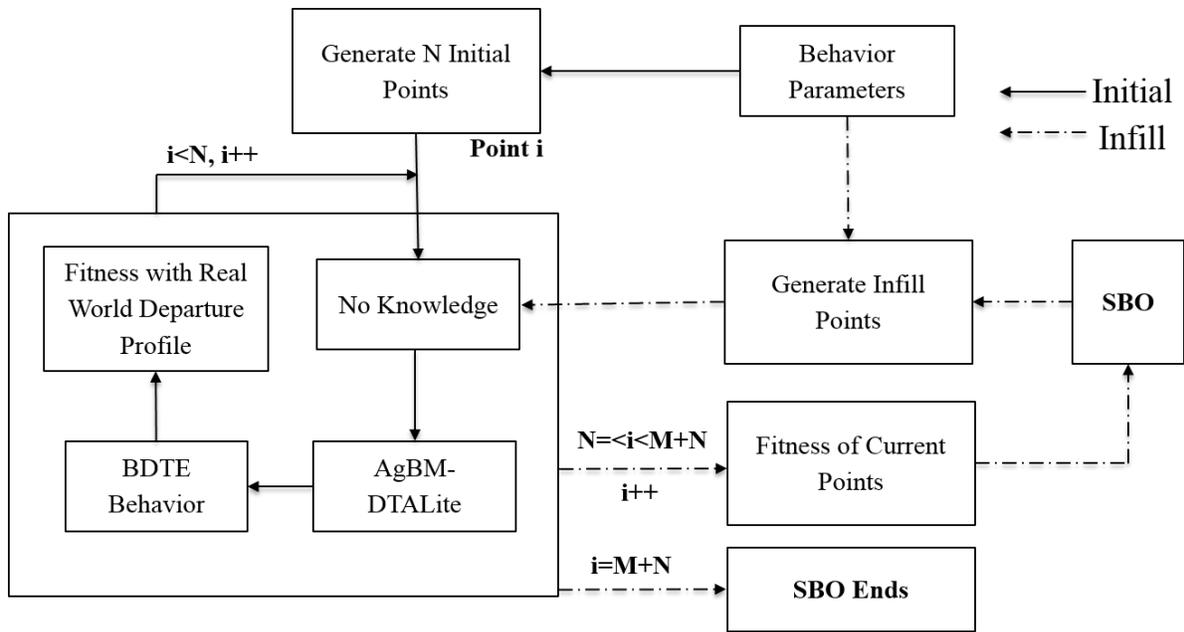
- Recentness factor (Commute Trips), range from 0 to 1
- Representativeness factor (Other Trips), range from 0 to 1
- Recentness factor (Other Trips), range from 0 to 1
- Search Scale factor, range from 0.6 to 1

The first four parameters are the factors in experience weight function in equation (20a) for both commuting trips and non-commuting trips; the last parameter is a scale factor for commuting trips, which means the commuting travelers will have smaller switches in search rules. The objective function is shown below:

$$Y = \sqrt{\frac{\sum_{i=1}^K (d_i - d_i^{BDTE})^2}{\sum_{i=1}^K d_i^2}} + \sqrt{\frac{\sum_{i=1}^K (cd_i - cd_i^{BDTE})^2}{\sum_{i=1}^K cd_i^2}} \quad (21)$$

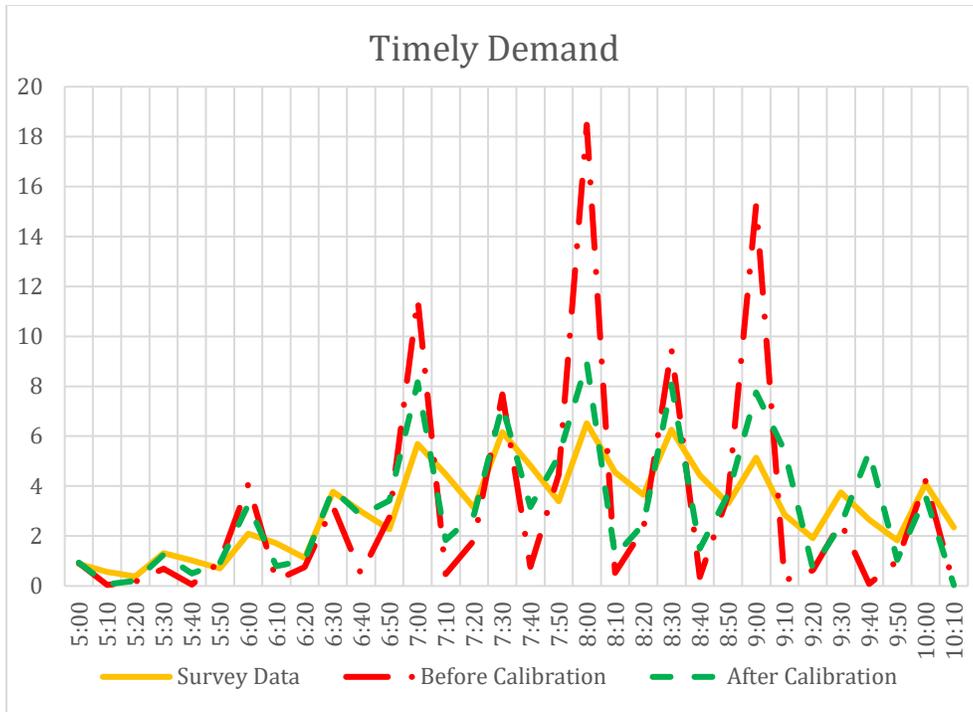
where the simulation horizon is divided to  $K$  time periods;  $d_i$  is the demand percentage of period  $i$  over the whole horizon in real world data;  $cd_i$  is the cumulative demand percentage from period 1 to period  $i$  in real world data;  $d_i^{BDTE}$  and  $cd_i^{BDTE}$  are the corresponding percentages of the condition from the AgBM model.

The flow chart of SBO calibration is shown in Figure 3.7. First,  $N$  initial decision vectors will be generated by using the Latin hypercube sampling method (Chen et al., 2015) to fill the five-dimensional range space of the decision variables. These initial vectors of behavior parameters will be loaded into the AgBM model to approximate the objective functions one-by-one. Next, the team applies the surrogate model modified by Chen et al. (2015) to regress the relation between  $X$  and  $Y$ . Based on the estimated parameter relationship, a new infill point with a large potential to reach better objective value will be generated for the AgBM to approximate its performance. After re-estimating the relation with the infill point, the system will generate one more infill point and repeat the infill process until the number of total infill points reaches  $M$ . After the termination of this SBO process, the system will pick the point with the lowest objective value to be the optimal solution. The capability of SBO to obtain the global optimal decision variables has been explained in this report.

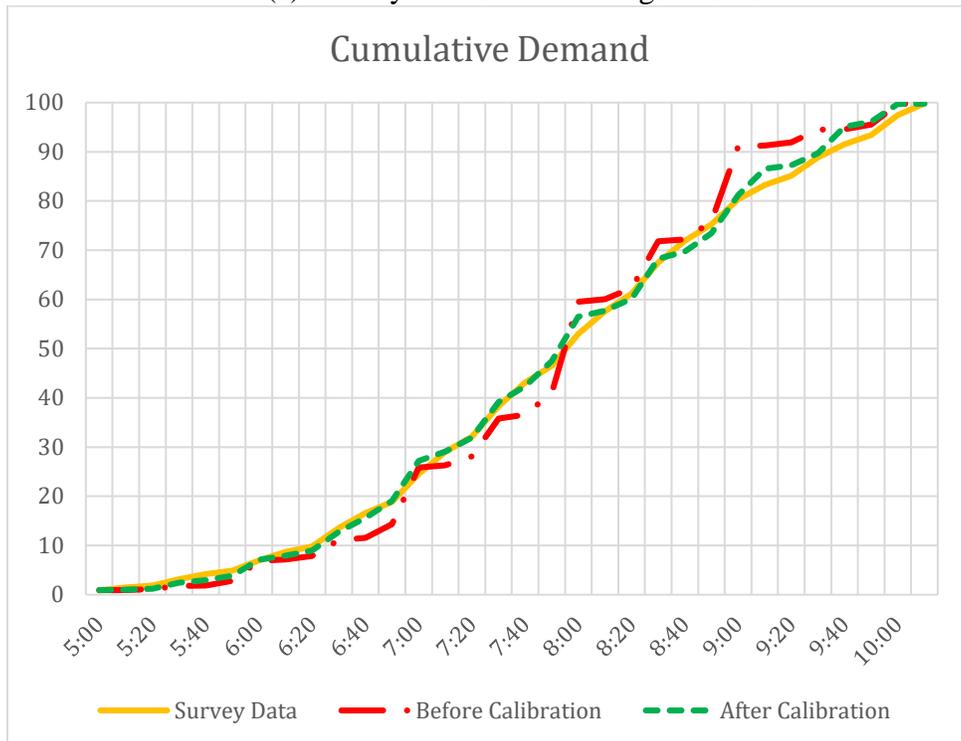


**Figure 3.7 Flow chart of SBO model calibration**

To enhance the applicability of this model, we applied the above calibration framework to a real-world traffic system located in Montgomery County, Maryland. The calibration horizon was from 5:00 AM to 10:00 AM, which contains 237,415 trips in the simulation model. The real-world demand profile is obtained from the 2007/2008 TPB/BMC survey. We generated 80 initial points and 40 infill points to complete the calibration. For each point, it takes around 20 simulation days to reach BUE. Before calibration, when the 5 parameters are all set to 1, the value of objective function is 7.559. After 120 iterations of SBO, the objective value drops down to 3.211, with the optimal decision vector [0.9925, 0.7724, 0.0247, 0.6123, 0.9953]. This calibration result indicates that commuting travelers weigh more on representativeness and non-commuting travelers weigh more on recentness. The search scale factor is close to 1, proving that the search rules are well validated in the previous approach. The comparison of the departure profiles is shown in Figure 3.8 (a) and Figure 3.8 (b).



(a) Timely Demand Percentage Profile



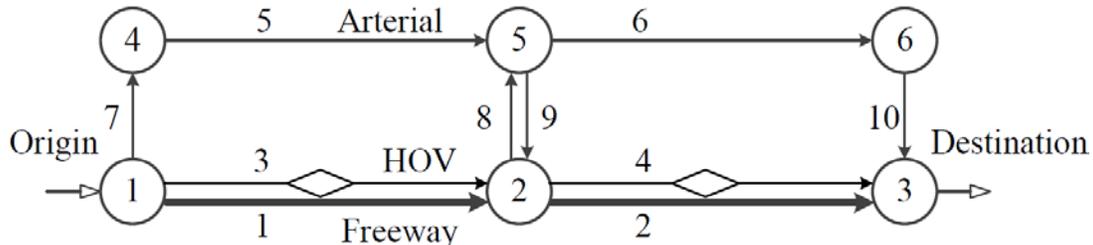
(b) Cumulative Demand Percentage Profile

**Figure 3.8 Temporal calibration results (a. timely demand percentage profile; b. cumulative demand percentage profile)**

## 3.3 SENSITIVITY TESTING

### 3.3.1 AgBM Simulation Setup

The proposed multidimensional behavioral theory and models have been estimated and implemented in an agent-based simulation to demonstrate the capability. A demonstration network with one origin-destination pair, four alternative routes and three travel modes (auto, carpool, and transit) is employed. The network is illustrated in Figure 3.9. It consists of 6 nodes and 10 links (including 2 HOV links).



**Figure 3.9 A demonstration network for sensitivity testing**

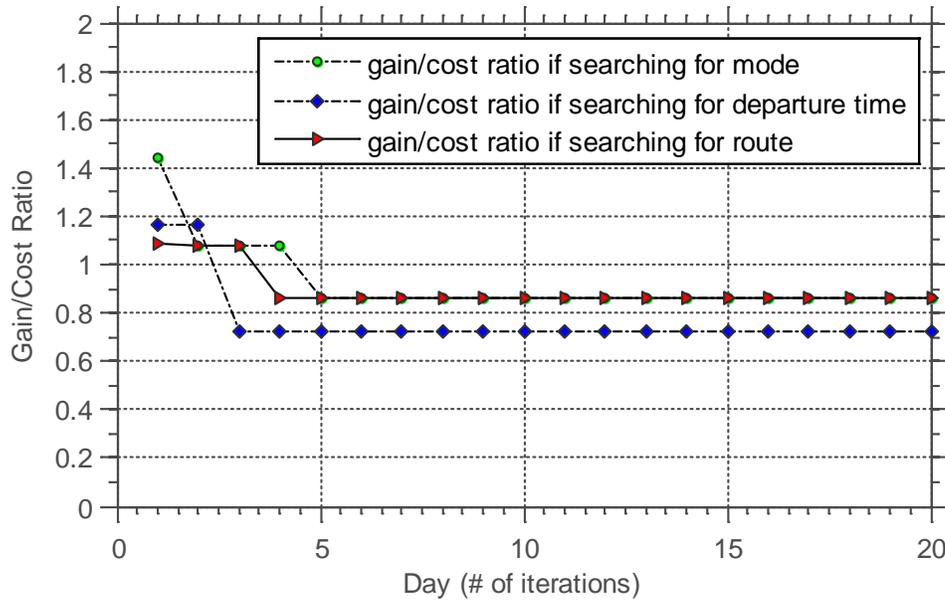
The scenario that is analyzed in this simulation is a future-year scenario with increased travel demand that creates excessive travel time and cost for the simulated agents, and prompts them to start the multidimensional behavior adjustments. This microsimulation of extended morning peak hours (5:00 am – 10:00 am) generates 90,000 agents. The agents' characteristics are synthesized based on Transportation Planning Board (TPB) – Baltimore Metropolitan Council (BMC) Household Travel Survey data collected in 2007 and 2008.

### 3.3.2 Individual Behavioral Dynamics

In the simulation, agents travel from the origin to the destination, accumulate experience, make behavioral adjustments on one or multiple dimensions, dynamically update beliefs, and eventually satisfy their decisions. The uniqueness of the model brings attention to each agent for whom the interplay of search gain and search cost is dynamically modeled in order to determine the behavioral dimension wherein the search and decision process occurs. Figure 3.10 illustrates the evolving gain/cost ratio for a particular agent.

On simulation day one, the agent initially believes that all dimensions are rewarding (with all gain/cost ratios above one) while the most profitable dimension is the mode dimension. She/he then employs search rules and decision rules to identify and examine one alternative mode. While the subsequent search reveals further information, this agent's knowledge and subjective beliefs on the mode dimension evolve significantly. On the second day, the departure time dimension emerges to be the one with the highest gain/cost ratio. Therefore, a search for alternative departure time is performed. Iterating this process, the agent forms a time-dependent search path for choosing behavioral adjustment dimensions: the mode-departure time-route-mode. On the fifth day, the gain/cost ratios of all dimensions drop down below 1, which indicates that this agent subjectively believes that no more searches are necessary. The agent is then satisfied and stops the search. Once a new turbulence emerges in the transport system, such as new policies or booming

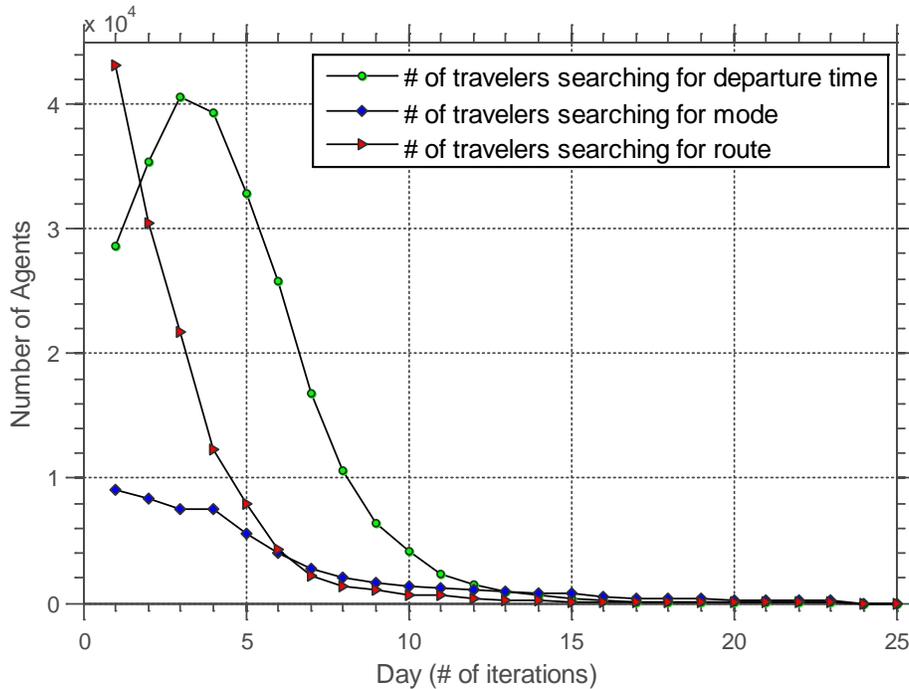
travel demand, the agent may be influenced by the growth of gain/cost ratios in certain dimensions, inciting them to seek further changes.



**Figure 3.10 The evolving gain/cost ratios of multidimensional travel behavior**

### 3.3.3 System Convergence and Sensitivity Testing

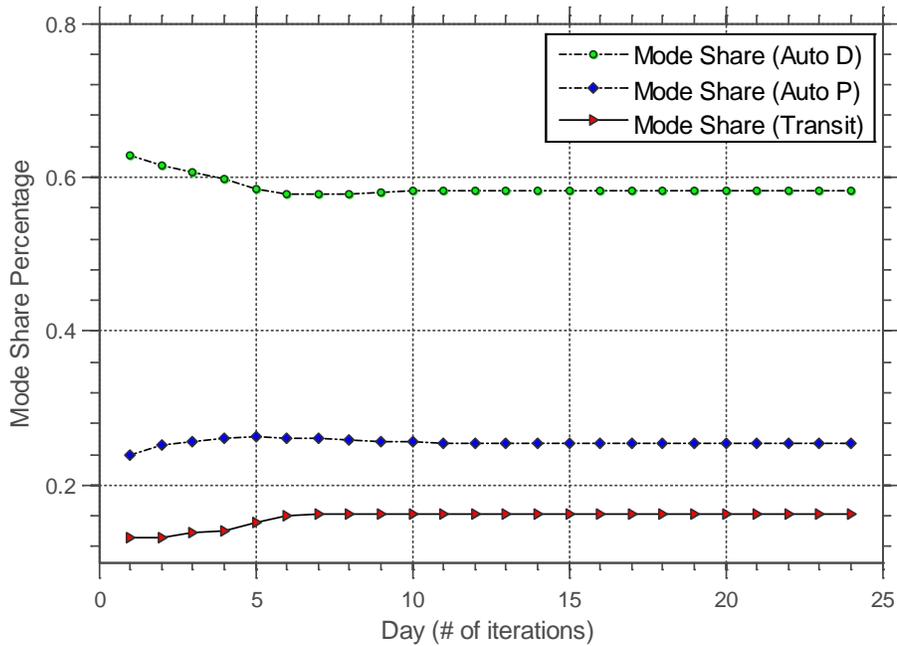
The convergence of the multidimensional behavior is illustrated in Figure 3.11. The scenario where all users stop making behavioral adjustments is defined as the Behavioral User Equilibrium (BUE). Overall, the model predicts active and reasonable agent behavior along the three behavioral dimensions. The convergence processes are smooth. With the innate bounded rationality and satisficing behavior, agents reach steady state and stop search within 25 search iterations. If each agent travels five days a week and all agents start searching at the same time, it would take five weeks for the traffic to stabilize and reach BUE on the network. This is an interesting finding in that, on the one hand, it allows us to model the gradual behavior adaptation to exogenous policies (e.g., pricing policy in Stockholm gradually nudges drivers to change behavior, Borjesson *et al.*, 2012). On the other hand, it suggests potential applicability of the proposed theory in large-scale planning models and simulation since it embeds multidimensional behavioral responses while maintaining a reasonable converging speed.



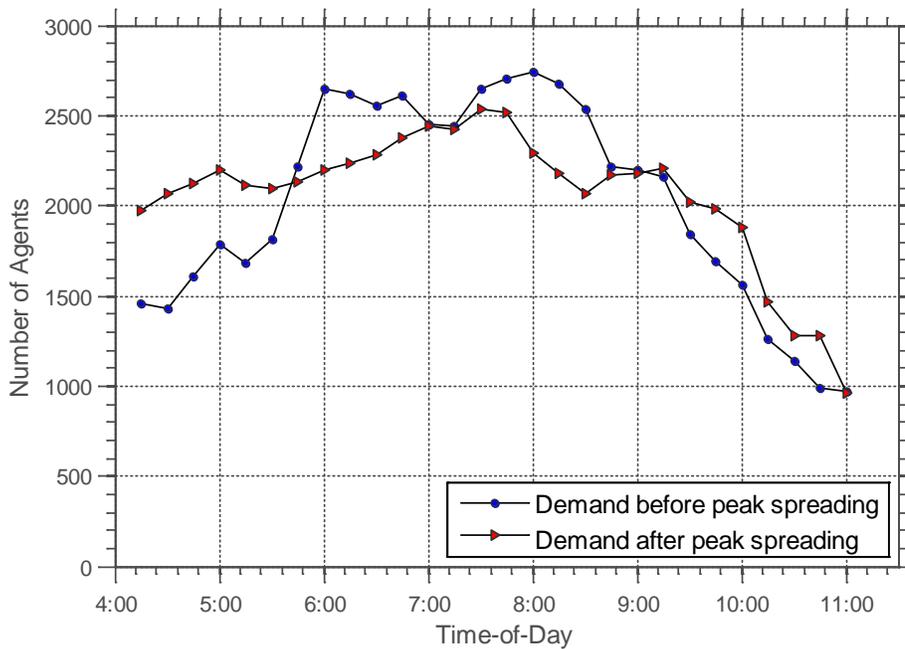
**Figure 3.11 The convergence of multidimensional travel behavior**

In response to the assumed demand increase, changing route and changing departure time are the most significant ways of adapting behavior. The initially high route-searching frequency cools down rapidly when agents have difficulty identifying better alternative routes under the assumed overall demand increase. Agents quickly learn this and update their subjective beliefs, which results in a decreasing search gain in the route dimension. Instead, agents search alternative modes and departure times, leading to an increase in the number of agents searching for alternative departure times in the second and third simulation days. A few agents search for alternative modes. Agents' mode searching and switching behavior is illustrated in Figure 3.12. Agents' departure time changes are illustrated in Figure 3.13.

By aggregating the individual behavior into travel patterns, we can observe that multidimensional learning and adaptation leads to a slight percentage decrease of auto drivers (Auto D in Figure 3.12). Those agents switch to auto passengers (Auto P) or transit users. The aggregate mode share of auto drivers drops from 63.4 percent to 58.3 percent. After six simulation days, the mode share tends to stabilize even though, from the microscopic level, there still exists some 3,000 travelers changing their travel modes. The active departure time changes lead to a significant peak spreading effect. The assumed demand increase results in more severe congestion and unreliable travel time, especially during peak hours. The excessive travel time, cost, and schedule delays make the departure time adjustments necessary in order for agents to gain an acceptable payoff through search. The model predicts that the dominating behavioral responses to the stimulus are route changes and departure time changes, which are consistent with the existing research (e.g., Arentze *et al.*, 2004). Meanwhile, the model predicts the behavioral dynamics and adaptive processes, which advance the current understanding of multidimensional travel behavior adjustments.



**Figure 3.12 Agents' mode share changes**



**Figure 3.13 Agents' departure time changes and peak spreading**

Travelers in the multidimensional agent-based model are not perfectly “rational,” in that they do not maximize their utility (or payoff). Instead, they are restrained by information acquisition cost,

decision cost, computational limitation, time budget and deadlines. Their irrationality also extends to their response to different intuitions and heuristic behavioral rules. Figure 3.14 demonstrates that through multidimensional learning and adaptation, agents search and improve their relative searching payoff. This term is defined as the ratio of the cumulative actual search gain and the cumulative subjective search gain (i.e., subjectively believed maximum payoff from the search) for all searchers. Judging by the curves, the departure time dimension is the most profitable dimension. Once agents begin searching in this dimension, they can retrieve the highest relative searching payoff. However, this learning and adaptation does not ensure that agents make decisions that result in maximum payoff. This example demonstrates the bounded rationality of the agents in search and changing their behavior.

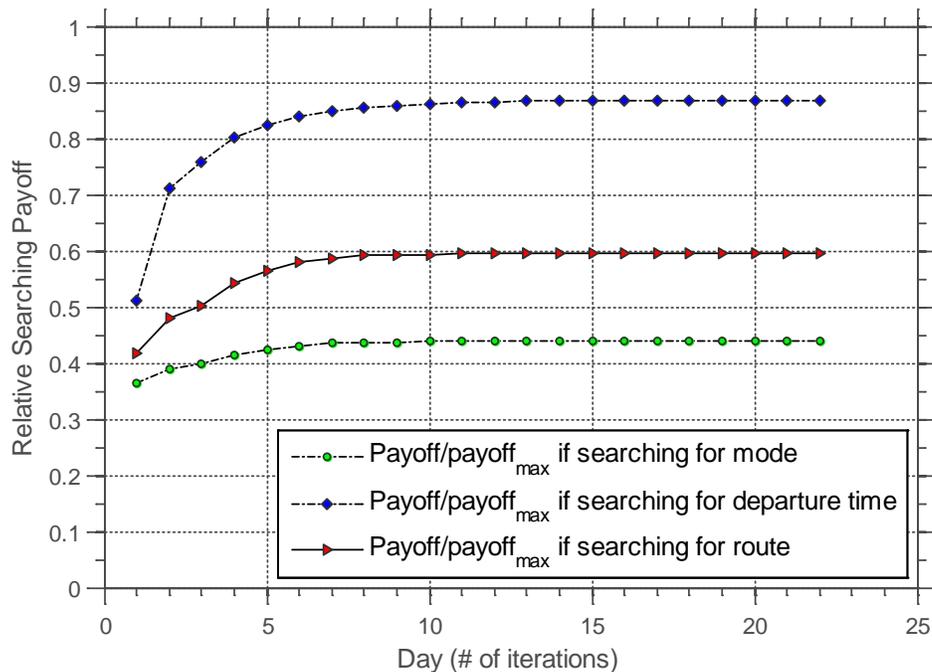


Figure 3.14 Agents' payoff dynamics

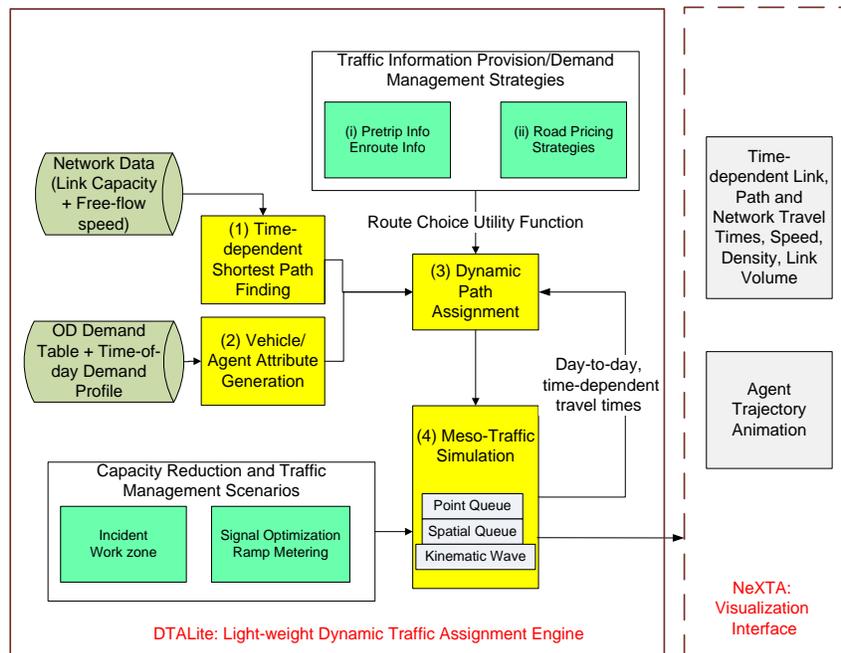
## 3.4 DTALite Traffic Simulation Engine

### 3.4.1 DTALite Simulation Engine

DTALite, an open-source light-weight, mesoscopic DTA simulation package, in conjunction with the Network eXplorer for Traffic Analysis (NeXTA) graphic user interface (GUI), has been developed to provide transportation planners, engineers and researchers with a theoretically rigorous and computationally efficient traffic network modeling tool. This fully functional, open-source dynamic traffic assignment model can be downloaded from <https://code.google.com/p/nexta/>. The software suite of DTALite + NeXTA seeks to:

- (1) Provide an open-source code base to enable transportation researchers and software developers to expand the range of capabilities to various traffic management applications;
- (2) Present results to other users by visualizing time-varying traffic flow dynamics and traveler route choice behavior in an integrated 2D/3D environment; and
- (3) Provide a free educational tool for students to understand the complex decision-making process in transportation planning and optimization processes.

Additionally, DTALite adopts a new software architecture and algorithm design to facilitate the most efficient use of emergent parallel (multi-core) processing techniques and exploit the unprecedented parallel computing power newly available on both laptops and desktops. The DTALite framework is illustrated in Figure 3.15. It integrates the four major modeling components, highlighted in yellow.



**Figure 3.15 DTALite Software System Architecture with Key Modeling Components**

DTALite requests at least two types of input data to conduct traffic assignment and/or traffic simulation: demand input data and network data. The demand input data includes a number of .csv files that describe the vehicle type (for emission and energy consumption estimation), demand type, OD tables and agent list. Both OD table and agent list are acceptable to provide the travel demand in the network. The OD tables store the time-dependent travel demand between different OD pairs. Most importantly for this integration effort, the agent list option provides the necessary linkage between SILK-AgBM and DTALite. The agent list predicted by SILK-AgBM can be used directly by DTALite. This agent list includes more detailed information, such as agents' departure time, origin, destination, modes and route trajectories. The network data used by DTALite include several .csv files to code the transportation network, including information about node location, link location, and link attributes (e.g., capacity, number of lanes, link type, speed limit, toll, etc.).

DTALite's four major modeling components include:

- (1) Time-dependent, shortest path finding, based on a node-link network structure.
- (2) Vehicle/agent attribute generation, which combines an origin-destination demand matrix with additional time-of-day departure time profile to generate trips.
- (3) Dynamic path assignment module, which considers major factors affecting agents' route choice or departure time choice behavior, such as (1) different types of traveler information supply strategies (e.g., historical, pre-trip, and/or en route information, and variable message signs), and (2) road pricing strategies where economic values are converted to generalized travel time.
- (4) A class of queue-based traffic flow models that can accept essential road capacity reduction or enhancement measures, such as work zones, incidents and ramp meters. The queue-based traffic simulation model in DTALite only requires basic link capacity and free-flow speed for operation, which are readily available from static traffic assignment models. By using simple input parameters, in addition to possible connections with common signal data interfaces, the proposed simulation package may enable state DOTs and regional MPOs to rapidly apply advanced DTA methodologies for large-scale regional networks, subareas, or corridors. Additionally, the modularized system design may help serve future needs by simplifying the process for transportation researchers and software developers to continue to build upon and expand its range of capabilities.

With DTALite, one could integrate travel demand/behavior models in different ways. One typical way is to use travel demand models to prepare demand input data for DTALite and then rely on its time-dependent, shortest path finding and dynamic path assignment module to assign trips. In this AgBM-DTALite integration, traffic assignment modules in DTALite are replaced by agent-based behavioral rules. In other word, SILK-AgBM will also do the routing for each agent. DTALite is employed as a traffic simulation engine (its meso-traffic simulation is mainly used in this integration).

## **3.5 INTEGRATED SILK AgBM-DTALITE**

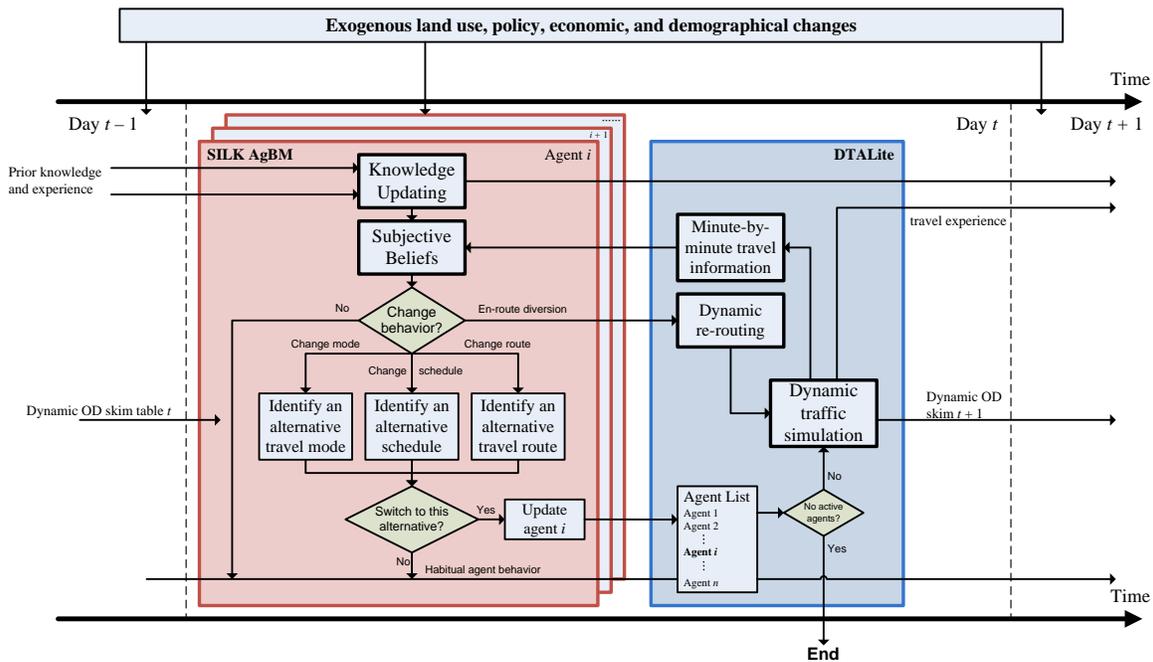
### **3.5.1 Overview of the Model Integration**

A transportation system typically has two major components: the transportation network and its users (potentially, decision-makers such as suppliers, logistics carriers, companies, and policy agencies can also be considered). The conditions of the transportation network are closely interrelated to the decisions made by its users, and vice versa. However, contemporary transportation models either fail to represent this or have to rely on unrealistic assumptions:

- Traditional, four-step transportation planning models assume a sequential process: trip generation, trip distribution, modal split and traffic assignment. When trips are assigned to paths, other behavioral dimensions such as departure times and modes are assumed fixed. In other words, there is no direct linkage from supply-side traffic conditions to estimation and prediction on the travel demand side.

- Activity-based models consider trips as induced demand from daily activity arrangement. There are certain “loops” built in activity-based models to reflect the linkage. However, activity-based models often assume perfect rationality in their discrete choice models for travel behavior. For instance, user equilibrium (or stochastic user equilibrium) is still assumed for route choices. This type of model focuses on modeling what travelers should do. It lacks behavioral realism and causes long computing time due to the calculation of equilibrium.

As briefly presented in section 3.5, agent-based travel behavioral models have the capability of mimicking and simulating the travel behavior changes of each user in the system. SILK-AgBM is an agent-based travel behavioral model system that emphasizes the role of searching, information, learning and knowledge (SILK). This modeling tool has been developed by UMD and already applied by MDSHA. It can also be efficiently interfaced with a dynamic traffic assignment model, such as DTALite. Once integrated with a traffic simulator, the system can be complete, given that all traffic conditions in the transportation network can be imitated by the simulator. This motivates the proposed integration of agent-based models and DTA simulator, as illustrated by the following flowchart (see Fig. 3). The traffic simulator used is the DTALite model (i.e., an open-source Lightweight Dynamic Traffic Assignment and Simulation Engine, <https://code.google.com/p/nexta/>). Therefore, the integrated model is named AgBM-DTALite for short.



**Figure 3.16 The systematic framework for the integrated AgBM-DTALite model**

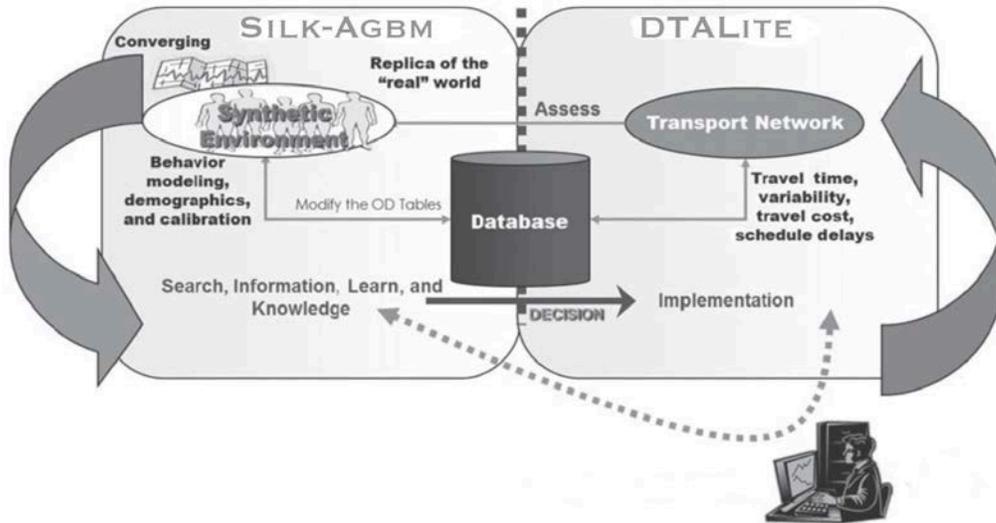
Travelers arrange their daily or recreational itineraries based on knowledge and various information sources: previous experience, social network, mass media, real-time traffic data sources (e.g., Google and INRIX), etc. Exogenous changes may result in different adjustments to the travel itinerary. AgBM models the travel behavior with the full consideration of information, learning, knowledge and searching. Here, the emphasis is given to the integration and, in particular,

the information exchange between AgBM and DTALite. DTA models are capable of simulating traffic in greater detail and producing various time-varying traffic information. A successful integration can provide a useful analysis tool to predict travel behavior in higher fidelity and accuracy and to evaluate various exogenous changes. The changes include relatively shorter-term, real-time information provision through advanced traffic information system (ATIS), as well as more long-term vehicular technology advances (e.g., ride-sharing, connected/autonomous vehicles). In the proposed AgBM-DTALite, two levels of integration are developed:

- (1) **Between-day integration:** On one particular simulation day, agents are able to acquire information from previous days and accumulate knowledge about the transportation system. For instance, when an autonomous vehicle is introduced to a household in a future year, members of the household will respond and rearrange their trips. Seniors and juveniles who previously rely on non-auto modes now may consider riding in the vehicle. Working adults may need to readjust departure time to accommodate the foreseeable increase in vehicle usage. These changes to each agent are modeled and outcomes are fed into DTALite to simulate dynamic traffic conditions, based on which agents will adapt their behaviors again.
- (2) **Within-day integration:** Within the same day, information is conveyed between AgBM and DTALite. Real-time information on congestion and different non-recurrent incidents has been made available to a certain percentage of agents, which reflects the fact that ATIS subscribers and Google/INRIX users have access to timely estimates of traffic congestion. This type of information exchange would trigger dynamic behavior adaptation. En-route diversion is a likely reaction and is incorporated in this integration. Future studies may also internalize dynamic modal shifts (park-and-ride options along major freeways, ride-sharing, etc.).

### **3.5.2 Day-To-Day Integration**

A unique aspect of SILK-AgBM and DTALite is that their data conversion and communication is fairly user friendly and thus the integration can potentially be developed flexibly in any programming environment. This aspect is a benefit to the integration. Nevertheless, certain data communication between AgBM and DTALite is necessary under the consistent data hub platform in order for comprehensive and streamlined analysis of the complex transportation system.



**Figure 3.17 Data flow chart of AgBM-DTALite day-to-day integration**

Travelers are synthesized in SILK-AgBM. Their travel behaviors (i.e., mode, departure time and route choices) are modeled and simulated based on concepts of search, information, learning and knowledge. These travelers' decisions are implemented in the transportation network and simulated using DTALite traffic simulation engine. Traffic conditions are predicted using DTALite and then again assessed by SILK-AgBM agents.

**Table 3.7 Comparison between SILK-AgBM and DTALite**

<b>Integration Components</b>	<b>SILK-AgBM</b>	<b>DTALite</b>
<b>Developed Functionalities</b>	CSV file reading and writing; integration convergence definition	CSV file reading and writing
<b>Input Data from Data Hub</b>	Day-to-day, time-dependent traffic skims for different modes	Traffic network; multimodal time-dependent agent list; agent path information
<b>Output Data to Data Hub</b>	Agent lists with OD information, mode choice, departure time choice and route choice	Speed/density/volume data for each link at different time-of-day

The method in which the DTALite model produces output has been adapted in order to meet the integration needs. A number of output files are generated. Therefore, for agent-based modeling, each agent is able to extract detailed day-to-day traffic information for her/his decision-making processes. The output files are in standard .csv format.

- **Output\_Agent.csv:** This file, which has been output from the agents' perspective, record all agents travel information, including their path sequence with exact simulation clock, their demand types, vehicle types, etc.

- **Output\_Skim.csv:** This skim file is the main feedback mechanism from DTA to SILK-AgBM, in which travel times, distance, toll cost and path node sequence between each OD pair are informed by the DTALite traffic simulator. These traffic skims are summarized for different time-of-day (i.e., traffic skims may differ in different time-of-day for the same OD pair) and different demand types (e.g., SOV, HOV, truck, light rail, etc.).
- **Output\_Routing\_Policy\_minXX.csv:** This routing policy output file can be generated depending on time periods. The major function is tracking agents on the network. To achieve that function, three references (distance, travel time, node sum) have been produced to distinguish and filter different agents. SILK-AgBM would receive all the output files from the data hub, analyze agents travel patterns and then modify input files for certain attributes for another day-by-day simulation.

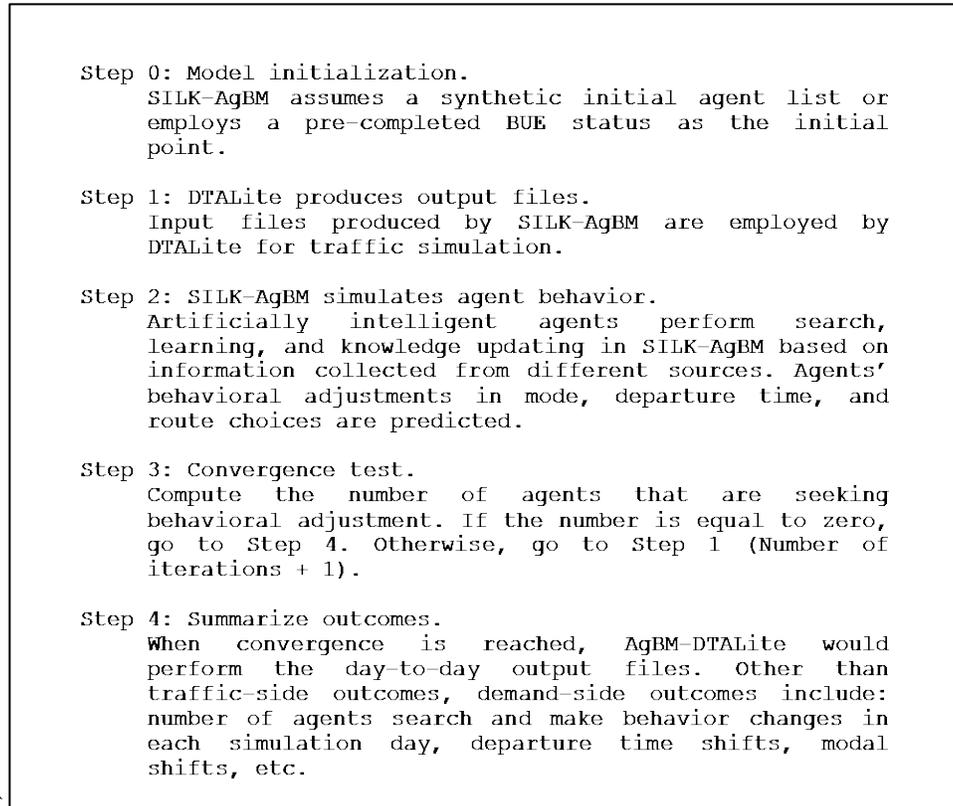
These output files will inform SILK-AgBM. Agents will search for information, learn and update their knowledge based on information collected from the simulation models. This agent-based knowledge and learning is maintained in the SILK-AgBM modeling platform. Based on the traffic output files and agent-based knowledge updating, the SILK-AgBM model will also predict multidimensional agent behavior responses in mode, departure time and route choices. The outcomes of SILK-AgBM have also been adapted and reorganized following DTALite-compatible format (e.g., Input\_Agent.csv). This list of agents will be used as input for the next round of DTALite traffic simulation.

A	B	C	D	E	F	G	H	I	J	
id	from_zone	to_zone	departure	demand_type	vehicle_type	information	value_of_time	toll	distance	Path_Node_Sequence
0	16	26	60	3	4	3	20	0	2.436	3311;3361;3362;3722;82;79;74;37;
1	57	39	60	2	1	0	10	0	11.179	3;298;3654;3649;12530;3793;112;:
2	37	12	60	1	1	3	20	0	7.783	3402;3403;3408;3401;3365;3364;3
3	60	57	60	1	1	0	1	0	4.947	3696;3687;139;3690;147;3686;365
4	41	39	60	3	5	3	20	0	1.651	286;285;3428;3427;3424;3423;
5	34	2	60	1	1	0	40	0	6.249	3357;3347;3511;3653;3262;3263;3
6	57	60	60	1	1	0	20	0	4.842	3;298;3654;3309;3317;3315;3320;:
7	61	33	60	3	4	0	20	0	7.138	3454;3448;3433;3434;3496;3424;3
8	10	37	60	1	1	0	15	0	6.829	3338;3339;3682;3833;3820;3340;3
9	17	29	60	1	2	0	20	0	4.423	105;106;3768;256;12542;3771;377
10	60	27	60	3	5	0	20	0	6.846	3696;3687;139;3690;3681;3674;14
11	57	15	60	3	3	0	30	0	6.133	3;298;3654;3649;12530;12531;125
12	44	49	60	3	4	0	30	0	3.326	3405;16500;3406;3797;282;3818;2
13	59	47	60	2	1	0	15	0	13.426	3083;3055;3325;142;146;140;3681

**Figure 3.18 Screenshot of an example Input\_Agent.csv produced by SILK-AgBM and fed into DTALite (SILK-AgBM provides input information regarding agent ID, OD, departure time, demand type, vehicle type, information provision type, value of time, toll cost, path distance and path node sequence)**

A key component of the day-to-day integration is the redefinition of convergence in the integrated AgBM-DTALite model. As mentioned earlier, user equilibrium is no longer defined. Considering travelers' actual behavior and adaptation to traffic conditions, behavioral user equilibrium (BUE) is defined in this integration. BUE is a convergence state where no agents are seeking behavioral adjustments in any behavioral dimensions. Therefore, the convergence testing mechanism is reconstructed using this BUE concept.

A pseudo code for the day-to-day integration is provided below:



**Figure 3.19 Pseudo code for the day-to-day integration of AgBM-DTALite**

### **3.5.3 Within-Day Real-time Integration**

Day-to-day integration allows us to analyze various planning and policy applications. Nevertheless, short-term policy making such as active traffic management would require tighter integration between SILK-AgBM and DTALite at a fine-grained level, such as minute-by-minute. Based on different application needs, this level of detail can be customized. Therefore, we hereby refer to this tighter integration as “within-day” real-time integration.

In this integration, the focus is given to dynamic behavior adjustment, such as en-route diversion. In response to real-time changes in traffic conditions, traffic management strategies and/or non-recurrent incidents, certain travelers are more willing to make dynamic route changes in order to optimize their traveling. There is also passive diversion stimulated by dynamic message signs, advanced traffic information, etc. Drivers may convey to the detour instructions. At the same time, travelers could also make other behavior changes, such as pre-trip departure time adjustment, pre-trip or en-route modal switching, etc. These behavior considerations should also be incorporated in AgBM-DTALite.

In order to achieve this level of integration, a number of DTALite output files are generated: Real time link performance measures are reported in RT\_Output\_LinkMOE.csv. Attributes in this file sufficiently provide AgBM with link traffic situation in predefined time intervals, including speed, density and link travel time.

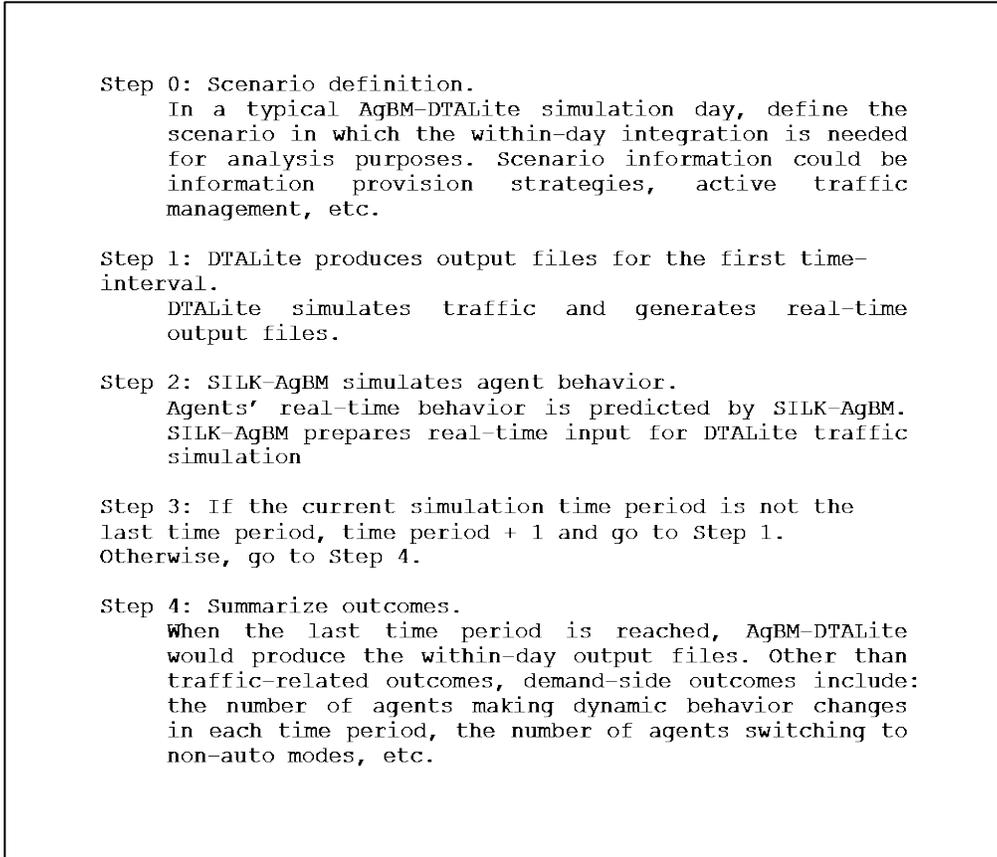
- Real time information on certain predefined agent-groups is reported in RT\_Output\_Tagged\_Agent.csv. Using agent tagging, DTALite can track particular agents, including those who use certain link(s) or who are traveling between a critical OD pair. This enables AgBM-DTALite to track agents who are actively changing behavior in real-time.
- Real time information on agents who complete their trips is reported in RT\_Output\_Complete\_Agent.csv.

These output files will be processed in real-time by SILK-AgBM, demonstrating that on-time traffic information is considered by travelers. Similar to the day-to-day integration, the following list of output files will be used as input by DTALite.

- A real-time, updated list of agents is provided by RT\_Input\_Update\_Agent.csv. This real-time input file created through AgBM would be deployed as a DTALite input file. Updated agents have new values of attributes, including departure time, path sequence, information type, etc. All of these new attributes of certain agents could create new link features, depending on how the agents behave or how AgBM works.
- Real time updated link attributes are provided by RT\_Input\_LinkAttribute.csv. Adjustments in agent behavior could also affect link attributes. Therefore, the integration uses this file to inform AgBM-DTALite about the link-level changes.

The time interval of one round of data exchange performed by SILK-AgBM and DTALite can be predefined in input\_simulation\_scheduling.csv. For instance, 300 sec for RT\_Output\_LinkMOE indicates that real-time link performance measures will be produced every 300 seconds. Note that different input/output can have different time intervals. For instance, an integrated model can produce RT\_Output\_LinkMOE in 300 seconds and read RT\_Input\_Updated\_Agent as input every 600 seconds. This way, the simulation clock can be synchronized.

A pseudo code for the within-day, real-time integration is provided below:



**Figure 3.20 Pseudo code for the within-day integration of AgBM-DTALite**

**3.5.4 The Convergence Property of the Integrated AgBM-DTALite**

Convergence of this agent-based model can be directly measured by the remaining number of travelers still searching for alternative routes. Performance of the network can be measured both after the equilibrium is reached and during the equilibrium process. This is an advantage of the agent-based model over network equilibrium models. This capability would be valuable in situations when the duration of the equilibrium process is quite significant in comparison with the planning horizon, or when it is clear that an equilibrium does not exist in the modeled system. Understanding how the AgBM-DTALite tracks the system evolution over time should be intuitive; therefore, the following memo focuses on the behavioral user equilibrium (BUE) and convergence property of the model.

In this subsection, a memo that demonstrates the finiteness of behavioral adjustment is provided. This is the theoretical foundation of the existence of behavioral user equilibrium and the definition of the convergence criteria of the AgBM-DTALite. The condition for the BUE is that for all users, perceived search costs exceed expected search gains. In order to demonstrate that this equilibrium indeed exists, it is necessary to revisit the definition of expected search gain  $g_{di}$  for traveler  $i$  in dimension  $d$ . The traveler's perceived search cost ( $sc_{di}$ ) in the same dimension  $d$  is always a positive constant. Therefore, there must exist a positive integer  $N_{di}$  for a specific dimension  $d$ , such that after  $N_{di}$  searches,  $sc_{di} > g_{di}$ , because:

$$\lim_{N_{di} \rightarrow \infty} g_{di} = \lim_{N_{di} \rightarrow \infty} \frac{C_{\min, d-1}^d - C_0}{N_{di} + 1} = 0 \quad (1)$$

For the same reason, there must exist a positive integer  $N_i$ , such that after  $N_i$  searches,  $sc_{di} > g_{di}$  for all dimensions  $d$ . Let  $N^*$  be the maxim of all  $N_i$ . The BUE will surely be reached after  $N^*$  search iterations, which demonstrate its existence. The BUE exists because users adjust their expected search gains to accommodate unsatisfactory performance of the transportation system. Eventually, one individual stops searching, either because a good alternative is identified, or because repeated experience with unsatisfactory alternatives leads to decreased expectations.

While the existence of BUE is guaranteed in AgBM-DTALite, it is not practical to perform the integration until every single agent stops behavioral adjustments, especially under the circumstances of slow convergence or large-scale simulation. To facilitate the integrated modeling and analysis, we also define the gap of behavioral user equilibrium as the percentage of travelers who are still searching for behavior changes.

$$\text{gap}(\%) = \frac{\sum_d n_{search}^d}{n_{total}} \times 100 \quad (2)$$

Where  $n_{search}^d$  denotes the number of agents who are searching in dimension  $d$ .  $n_{total}$  denotes the total number of simulation agents. Based on this measure, researchers are also able to quantify the speed of convergence and how fast a transportation system will re-equilibrate after disturbance. Applications of the AgBM-DTALite modeling system can choose the convergence criteria differently according to various analysis needs.

When the multidimensional search and switching rule sets are implemented as deterministic rules (e.g., the same antecedent condition of an if-then rule always yields the same behavior, etc.), the BUE behavioral pattern is always unique, given an initial system state, because travelers' multidimensional behavior adjustments are completely determined by their searching, learning, decision rules and the initial state. In this sense, the BUE with the deterministic implementation is also unique. However, if probabilistic rules are deployed in conjunction with a randomization procedure (e.g., a pseudo random number generator), the AgBM-DTALite system will exhibit Markovian variability and different BUE patterns will be obtained. Nevertheless, a specific random seed and initial condition will still lead to the same equilibrium state.

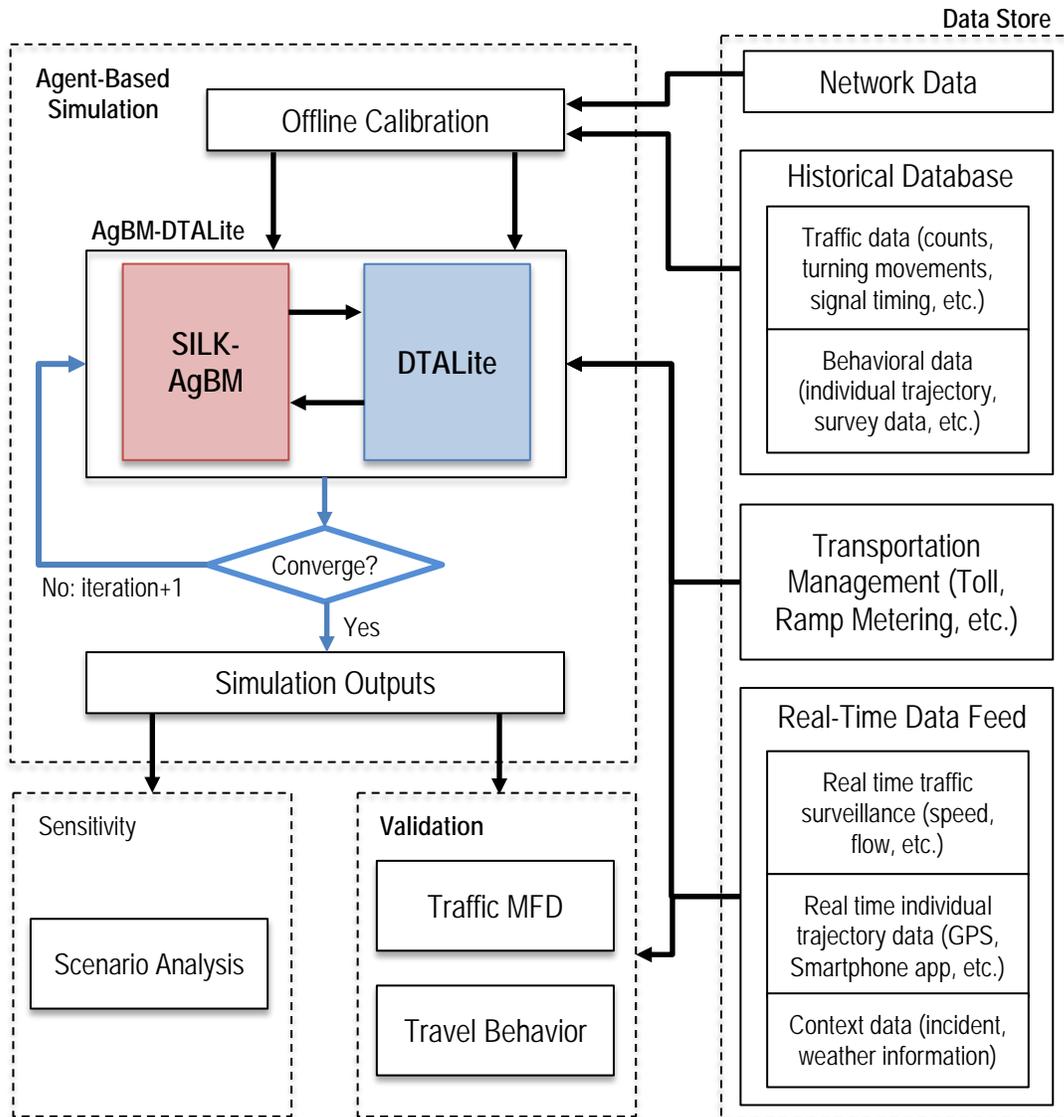
An equilibrium is considered stable if the system can return to the same equilibrium state after a perturbation. A small change in the network may cause individuals in the AgBM-DTALite system to start searching for new alternatives because their perception about search gains increases just slightly to exceed the perceived search cost in one or several behavioral dimensions. For the same reason, some travelers may change routes, departure times or modes when a small perturbation is introduced to the system. Even if just a few travelers actually change behavior, they may not switch back to their previously chosen routes, departure times or modes after the perturbation disappears, because the search and behavioral adjustments will enable them to learn new information and

update their spatial knowledge. Therefore, the BUE is not stable, which is expected for a system with historical dependency. The question of whether or not a small noise causes consequential impacts in the multidimensional travel behavior cannot be answered analytically. Simulation experiments are necessary.

The nature of the AgBM-DTALite dictates that the final BUE depends on the initial state specified. In this section, it is assumed that the network is initially empty and users are loaded onto the network stochastically based on their departure time frame (specified in the DTALite engine). In the real world, the network is constantly changing and population fluctuates. In order to apply the AgBM-DTALite and BUE for transportation planning or policy analysis, it is necessary to properly identify the initial conditions. A common method to deal with this issue is to first run the AgBM-DTALite with an arbitrary initial state and save a later system state as the new initial state for analysis. Another way of defining the initial state is to train the agents with knowledge/information that is based on available ground truth data sources (e.g., 24-hour traffic counting, link travel times based on probe vehicles, etc.).

## **3.6 CALIBRATION AND VALIDATION OF SILK**

The core of the system is the AgBM-DTALite integrated model, which simulates agent behavior and traffic trajectories. AgBM-DTALite adopts an integrated offline calibration model to adjust all of its internal model parameters based on historical observations. After the BUE is reached, the outputs of the agent-based simulation are used for model validation against real-time data feeds, including fixed traffic flow detectors and probe vehicle sensors. Travel behavioral patterns and dynamic traffic conditions are validated. Finally, model sensitivity is analyzed through scenario analysis.



**Figure 3.21 A systematic framework of the calibration, validation and sensitivity analysis of the integrated AgBM-DTALite model**

A two-stage offline calibration approach is adopted in the integrated AgBM-DTALite. In the first stage, OD matrices are calibrated using an iterative path-based OD adjustment algorithm. In the second stage, the multidimensional agent-based behavioral model parameters are adjusted using a Bayesian calibration method, taking full advantage of the historical database.

At this stage, online parameter calibration is not implemented. Real-time and online parameter calibration will certainly add value to the time-dependent performance of AgBM-DTALite. On the other hand, notable influence on the computational efficiency due to real-time adjustments will undermine the overall applicability of the integrated AgBM-DTALite in multi-corridor and regional case studies. In the current phase of research, the team focuses on the implementation of integration. Developing an online calibration module that uses the developed data store—

especially the real-time data feed—can be a useful extension in the next phase of this research. In this research, dynamic transportation control/management (toll rates, ramp metering rates etc.) and real-time data feeds (including real-time traffic surveillance, individual trajectories and context data) are used as inputs for the integrated model and for the model validation. A comprehensive validation of dynamic traffic conditions and multidimensional travel behavioral patterns will ensure the model performance is satisfactory.

Finally, the sensitivity analysis is included in the system to test how sensitive the model is to changes in the inputs and weights adopted in the integrated model. The amount of sensitivity analysis that can be performed is often application-specific and limited by budget and time constraints. In practice, it is recommended that decision-makers elicit preferences on what shall be considered the most important factors to examine in the sensitivity analysis. We demonstrate in this memo a number of sensitivity tests based on presumed scenarios. When implementing the model for real-world applications, a variety of sensitivity tests will be included, according to the needs of different application contexts.

### **3.6.1 Offline Calibration of AgBM-DTALite**

Before calibration begins, the complete 24-hour signal timing information for all signalized intersections in the study area must be accurate. Signal timing plans for these intersections are obtained from the state and county departments of transportation and then implemented in the simulation model. Signal timing plans are in standard Synchro data format, including information on phase timing (minimum initial/split, total split, yellow time, all-red time, etc.), phase sequence, detector locations, and so forth, for different time-of-day. 24-hour field count data are used for model calibration. The data comes from freeway and local arterial sensors, and are collected for multiple days. Using the hourly OD, the calibration algorithm (details are listed in Zhang *et al.* 2013) evaluates demand adjustment factor  $\alpha_{ij,r,t}$  associated with each path  $r$  between an OD pair  $i, j$  and, for a given time period,  $t$ , by using the following equation:

$$\alpha_{ij,r,t} = \frac{\sum_{a \in S(ij,r,t)} \zeta_{ij,r,a,t} \left[ \left( F_{a,t} + \Delta t_{ij,r,a,t} \right) / \left( f_{a,t} + \Delta t_{ij,r,a,t} \right) \right]}{\sum_{a \in S(ij,r,t)} \zeta_{ij,r,a,t}} \quad (3)$$

Where  $ij$  denotes OD pair from origin  $i$  to destination  $j$ ;  $r \in R(ij, t)$  where  $R$  denotes the complete set of all used paths of OD pair  $ij$  at time  $t$ ;  $S(ij, r, t)$  denotes the link set of path  $r$  at time  $t$ .  $F_{a,t}$  denotes the actual link flow on link  $a$  at time  $t$ ;  $f_{a,t}$  denotes simulated link flow on link  $a$  at time  $t$ ;  $\Delta t_{ij,r,a,t}$  denotes travel time from origin  $i$  to link  $a$  starting at time  $t$ .  $\zeta_{ij,r,a,t}$  is an indicator which equals 1, if  $a \in S(ij, t)$  and 0 otherwise.

Various performance measures have been applied to evaluate the accuracy of the match between field data and simulated counts:

#### 1. Root Mean Square Deviation (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (f(\mathbf{x}^{(i)}) - \hat{f}(\mathbf{x}^{(i)}))^2} \quad (4)$$

#### 2. Normalized Root Mean Squared Error (MSE)

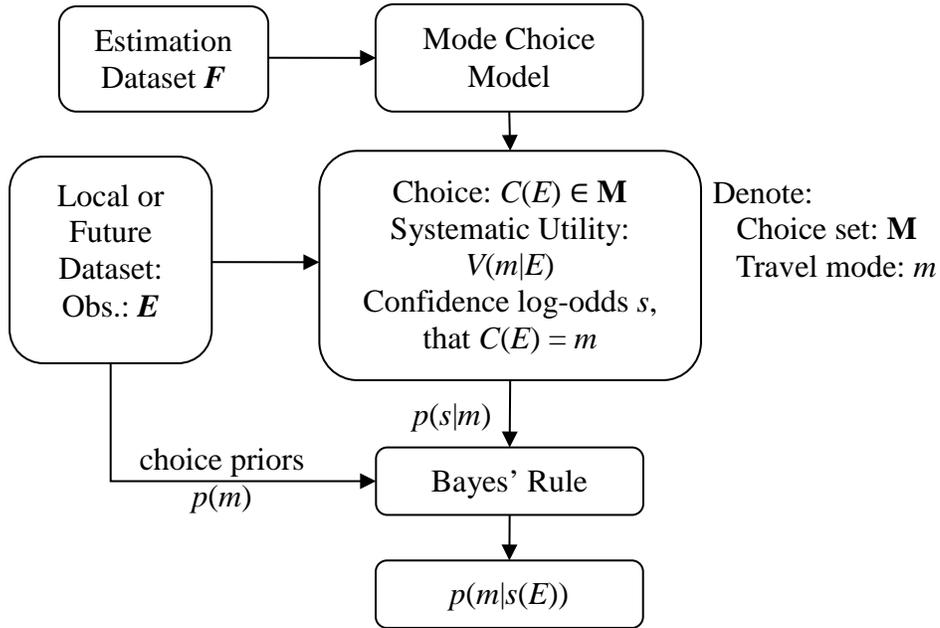
$$\text{NRMSE} = \sqrt{\frac{\sum_{i=1}^N [f(\mathbf{x}^{(i)}) - \hat{f}(\mathbf{x}^{(i)})]^2}{\sum_{i=1}^N (f(\mathbf{x}^{(i)}))^2}} \quad (5)$$

### 3. Pearson Correlation Coefficient (PCC)

$$\text{PCC} = \left( \frac{N \sum_{i=1}^N \hat{f}(\mathbf{x}^{(i)})^2 - \sum_{i=1}^N f(\mathbf{x}^{(i)}) \sum_{i=1}^N \hat{f}(\mathbf{x}^{(i)})}{\sqrt{[N \sum_{i=1}^N f(\mathbf{x}^{(i)})^2 - (\sum_{i=1}^N f(\mathbf{x}^{(i)}))^2][N \sum_{i=1}^N \hat{f}(\mathbf{x}^{(i)})^2 - (\sum_{i=1}^N \hat{f}(\mathbf{x}^{(i)}))^2]}} \right)^2 \quad (6)$$

Where  $N$  is the number of independent set data to be compared,  $f(\mathbf{x}^{(i)})$  and  $\hat{f}(\mathbf{x}^{(i)})$  denotes the observed and simulated count at sensor  $i$ . RMSE represents the sample standard deviation of the difference between simulated and observed counts; MSE indicates the relative deviation, where observed counts are weighted by volume; PCC is a measurement indicating the correlation between field counts and simulated counts. If  $r^2 = 1$ , the model is exactly predicting the test data, while  $r^2 = 0$  indicates there is no correlation between the model results and the field measurements.

After the OD calibration, the second stage of the offline calibration is to recalibrate model parameters of the agent-based travel behaviour model (AgBM). The research team develops a more flexible recalibration approach which relaxes the rigidity of using traditional recalibration parameters. The proposed method directly recalibrates the systematic utilities, which could also potentially enhance the accuracy.



**Figure 3.22 Bayesian calibration for SILK-AgBM models**

The estimated choice model predicts systematic utility functions based on the local or future-year dataset where the model is transferred. Different utility values are converted into confidence scores for different travel modes (indicating the strength of the decision that the empirical observation chooses mode  $m$ ). Note that the score may not necessarily match the local data well. Typically, this score is represented by log-odds defined using the following equation. This measurement transfers the original scale of choice probability (i.e.,  $[0,1]$ ) to a space  $[-\infty, +\infty]$  where different, continuous distributions are applicable.

$$s_m(E) = \log \frac{p(m)}{1-p(m)} \quad (7)$$

The log-odd statistics and predicted probabilities may not match the observed probabilities. To transfer the choice model, it is necessary to perform a mapping of the model predictions to actual observations. This mapping is represented by a series of posterior calibration functions based on Bayes' Rule. Conditioned on each mode alternative, a probability density estimator  $f$  is produced.

$$p(s|m) \sim f(s) \quad (8)$$

Here, various distributions can be employed to approximate the distribution  $f$ . For enhancing the performance, asymmetric posterior functions can be adopted here to improve accuracy. For instance, an asymmetric Laplace distribution can be produced by mixing two exponentials:

$$p(s|\theta, \alpha, \beta) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} \exp[-\alpha(\theta-s)] & s \leq \theta \\ \frac{\alpha\beta}{\alpha+\beta} \exp[\beta(\theta-s)] & s > \theta \end{cases} \quad (9)$$

These functions map score  $s$  to the actual choice probabilities. In other words, the family of posterior calibration functions adjusts the utilities estimated by the mode choice model in order to approximate the unobserved distribution of utility values for the local/future year dataset. Then, Bayes' Rule and the choice priors are used to obtain the estimate:

$$p(m|s) = \frac{p(m)p(s|m)}{\sum_{C \in M} p(C) \cdot p(s|C)} \quad (10)$$

Various measures of effectiveness (MOEs) are established in this study to evaluate the transferability as well as the recalibration results. Two typical measures for goodness-of-fit are employed to assess the transferability effectiveness: log-loss (Good, 1952) and mean squared error (Brier, 1950). Two other frequently used MOEs in the travel forecasting literature are also included in this study: transferability test statistics (Atherton and Ben-Akiva, 1976) and transferability index (Koppelman and Wilmott, 1982).

- Log-loss is employed when the model output is a numeric probability, which mainly performs as a gauge of prediction confidence. Log-loss measures the accuracy of a prediction and it can be considered as the cross entropy between the distribution of the true class and the prediction. Log-loss is calculated by Eq. (11):

$$\log loss = \sum_{C=\{m\}} \delta(C(E), C) \log P(C|E) \quad (11)$$

where  $\delta$  denotes the Kronecker delta function, which is equal to 1 if the two arguments are identical and 0 otherwise, and  $C(E)$  denotes the actual class of an empirical observation  $E$ . When the class of an observation is correctly predicted with a probability of 1, log-loss is 0. By minimizing the log-loss, one maximizes the accuracy of the model.

- Mean squared error (*MSE*) describes the difference between the prediction and the observation. When the class of an observation is correctly predicted with a probability of 1, MSE is 0. MSE is computed using Eq. (12):

$$MSE = \sum_{C=\{m\}} \delta(C(E), C) (1 - P(C|E))^2 \quad (12)$$

- Transferability test statistics (*TTS*) measures the difference between the uncalibrated model and the model recalibrated. The null hypothesis is that the parameters of the transferred model do not vary substantially from the parameters estimated from the application context data. *TTS* is calculated as twice the difference in log-likelihood, as defined by Eq. (13):

$$TTS = -2(l_j^*(\theta_i) - l_j^*(\theta_j)) \quad (13)$$

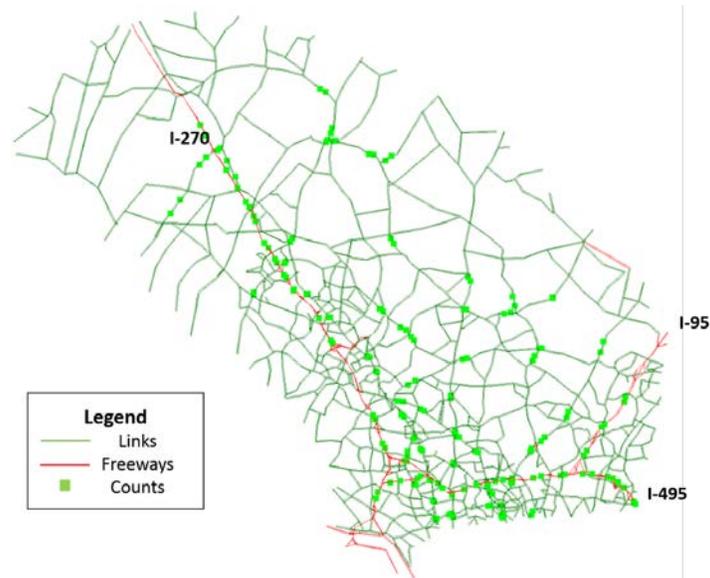
where  $l_j^*(\theta_i)$  is the log-likelihood of the transferred model and  $l_j^*(\theta_j)$  is the log-likelihood of the application context model. The null hypothesis is rejected when the *TTS* is greater than critical chi-square with degrees of freedom equal to the number of model parameters.

- Transfer index (*TI*) describes how the log-likelihood of the model exceeds a null model. *TI* is calculated as:

$$TI = \frac{l_j^*(\theta_i) - l_j^*(C)}{l_j^*(\theta_j) - l_j^*(C)} \quad (14)$$

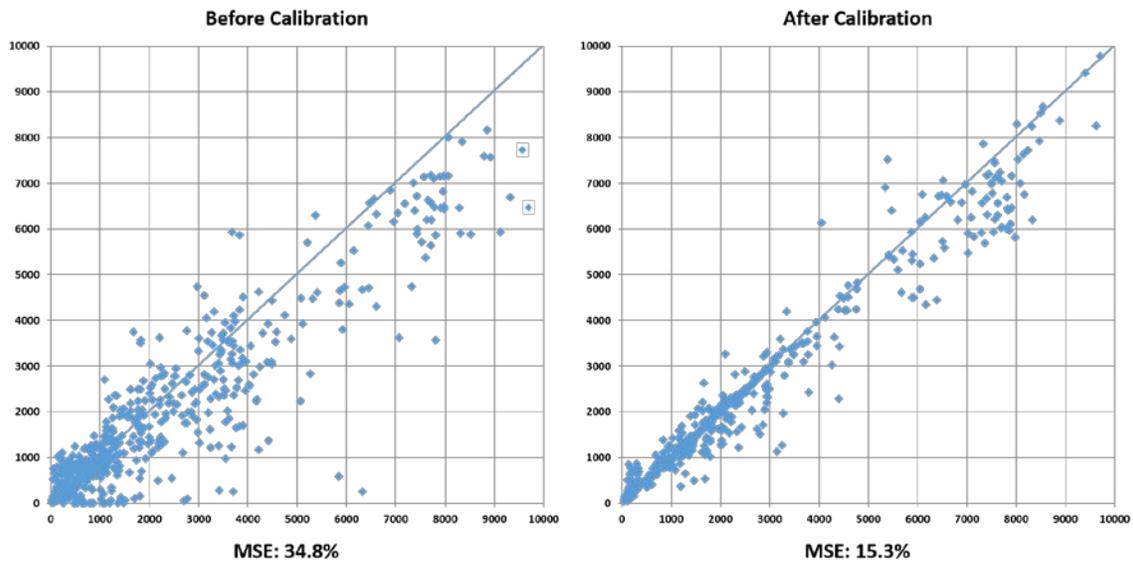
where  $l_j^*(C)$  is the log-likelihood of the market share model. In general, *TI* is a relative measure and *TTS* is a statistical test measure (Ortúzar and Willumsen, 2011).

In one of the proposed real-world applications, the simulation network covers the critical I-270 corridor and extends to I-95 and I-495 in Maryland. The network contains 474 internal zones, 102 external zones, 6,450 links and 2,373 nodes. The base year model (2010) is well calibrated with hourly volume data provided by 179 sensors from the SHA's traffic monitoring system, which provides two continuous day's hourly volume data (green points in Figure 3.23).



**Figure 3.23 Mesoscopic transportation network employed in the MITAMS C10**

The result of calibration is shown in Figure 3.24. It demonstrates that the mean squared error decreases from 34.8 percent to 15.3 percent in the calibration process. This calibrated outcome can be used for a wide range of applications, either through applying the mesoscopic model directly for regional analysis or for smaller case studies through sub-area analysis.



**Figure 3.24 Calibration results of the integrated AgBM-DTALite model**

### 3.6.2 Validation of AgBM-DTALite Simulation Outputs

The comprehensive validation involves the validation of dynamic traffic conditions as well as the validation of travel behavioral patterns. The validation of traffic conditions seeks to provide a macroscopic understanding of urban traffic dynamics. Using both fixed sensing data of freeways

and probe vehicle data of major arterials, the research team validates the simulation outputs on the network level, investigates the relationship of the accumulation of vehicles in a network with the exit outflows, and examines the equivalent relationship of the freeway network-wide weighted average density and flow rate. We have

$$N_t = \sum_{a \in A} k_{a,t} l_a \lambda_a \quad (15)$$

where  $N_t$  is the time varying number of vehicles in a network denoted by  $A$ , each individual link is  $a \in A$ ,  $k_{a,t}$  is the traffic density of link  $a$  at time  $t$ , and  $l_a$  and  $\lambda_a$  are the length and the number of lanes of link  $a$ .

$$K_t = \frac{N_t}{L} = \frac{\sum_{a \in A} k_{a,t} l_a \lambda_a}{\sum_{a \in A} l_a \lambda_a} \quad (16)$$

where  $K_t$  is the space mean density (vehicle per mile per lane) at time  $t$ , and  $L$  is the total length (lane-miles) of the network.

Analogously, we have

$$Q_t = \frac{\sum_{a \in A} q_{a,t} l_a \lambda_a}{\sum_{a \in A} l_a \lambda_a} \quad (17)$$

where  $Q_t$  is the space mean flow rate (vehicle per hour per lane) of the network, and  $q_{a,t}$  is the traffic flow rate of link  $a$  at time  $t$ . Both empirical observations (Geroliminis & Daganzo, 2008) and dynamic traffic assignment experiments on a real, large-scale urban network (Mahmassani & Saberi, 2013) concluded that  $Q_t$  is robust linear with the trip completion rate that is the sum of finished and exiting trips for the whole network.

The network-wide weighted average speed is given by

$$V_t = \frac{\sum_{a \in A} v_{a,t} k_{a,t} l_a \lambda_a}{\sum_{a \in A} k_{a,t} l_a \lambda_a} \quad (18)$$

where the weighted quantity is the number of vehicles on an arbitrary link  $a$  at time  $t$ . According to the traffic variables relationship in the MFD, as well as equations (2-4), the network average speed is estimated by

$$V_t = \frac{Q_t}{K_t} \quad (19)$$

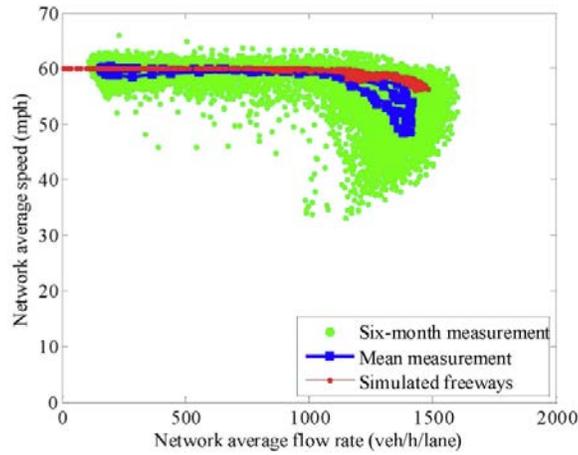
The spatial standard deviation of densities in the network is formulated by

$$\sigma_t = \sqrt{\frac{\sum_{a \in A} l_a \lambda_a (k_{a,t} - K_t)^2}{\sum_{a \in A} l_a \lambda_a}} \quad (20)$$

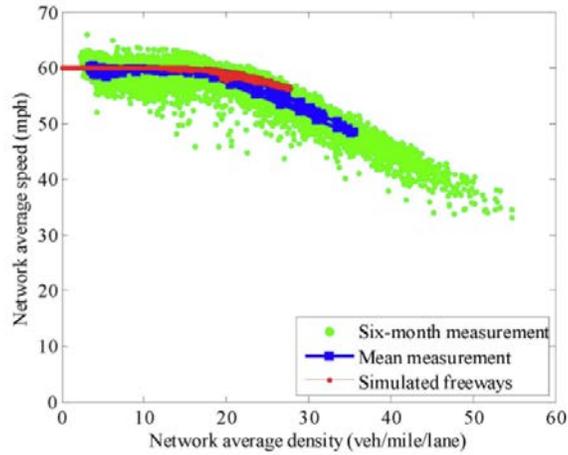
Without loss of generality, the coefficient of variation (CV) of network-wide densities is

$$\delta_t = \frac{\sigma_t}{K_t} = \sqrt{\frac{\sum_{a \in A} l_a \lambda_a \left( \frac{k_{a,t}}{K_t} - 1 \right)^2}{\sum_{a \in A} l_a \lambda_a}} \quad (21)$$

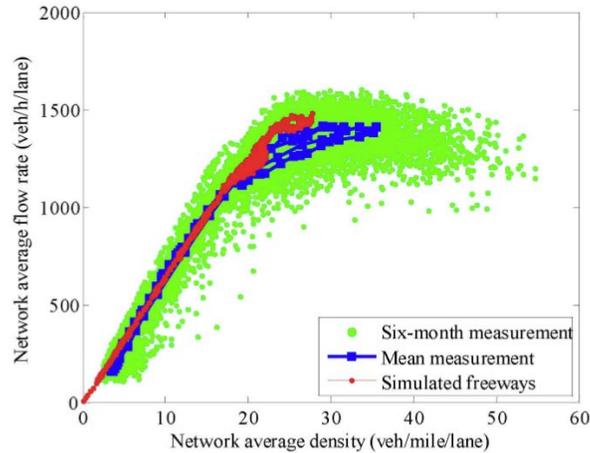
where the network CV  $\delta_t$  at time  $t$  is a normalized measure of spatial dispersion of vehicles in the network. The absolute value of the CV is expressed as a dimensionless percentage of the standard deviation to the mean value. CV can be used to identify the time-dependent reliability on a transportation network level.



(a)



(b)



(c)

**Figure 3.25 Comparisons of the simulated and measured MFD curves**

We run the AgBM-DTALite model for two replications. It validates simulation results by comparing the network average speed, density and flow with six-month traffic flow data at fixed sensors. The square blue curves represent historical average network-wide statistics across 130 weekdays (information provided by INRIX). Based on the freeway link output files (minute-by-minute link speed, density and flow) of the simulation, we can see the simulation matches well with historical measurements, considering that the simulation model was calibrated by an independent dataset (i.e., annual average daily traffic counts and turning movement).

In INRIX datasets, five-minute speed data of weekdays over the course of 2012 are available for every road segment of most arterials in the study area, but road densities and vehicle counts are not known, because probe vehicles consist of only a small amount of the road users.

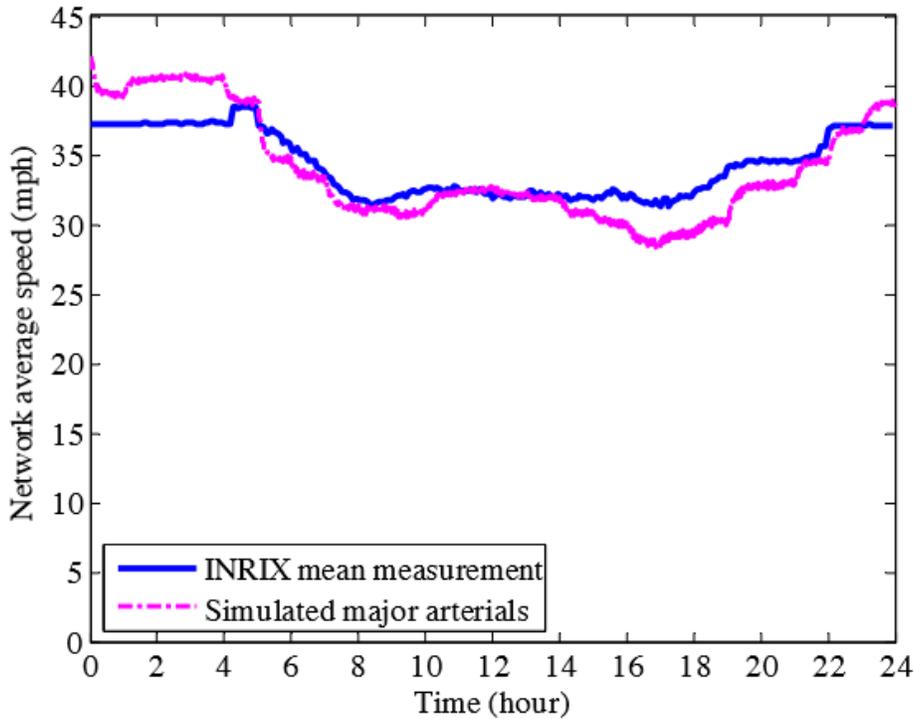
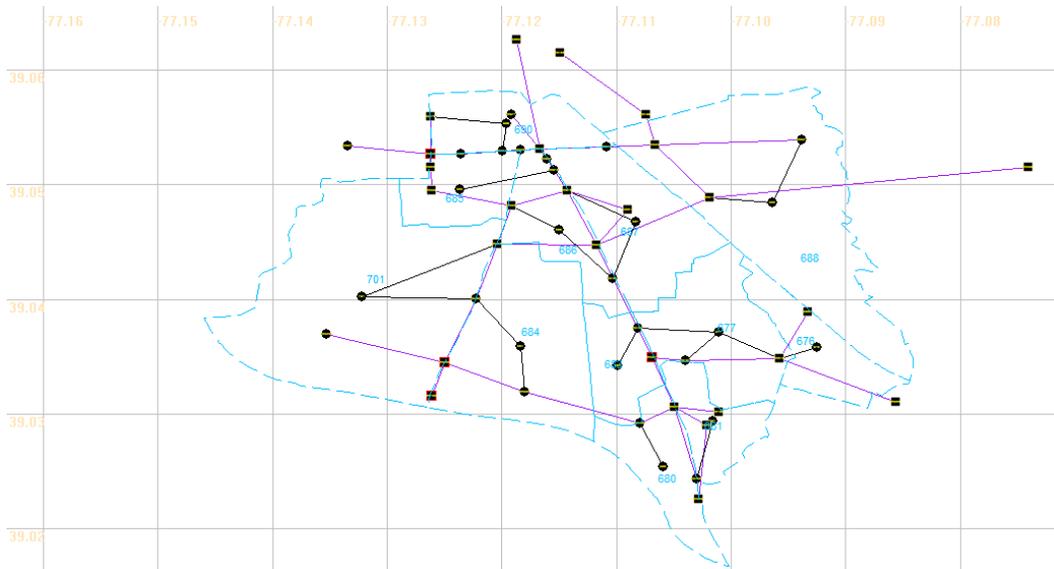


Figure 3.26 Comparison of the simulated and measured major arterial average speed

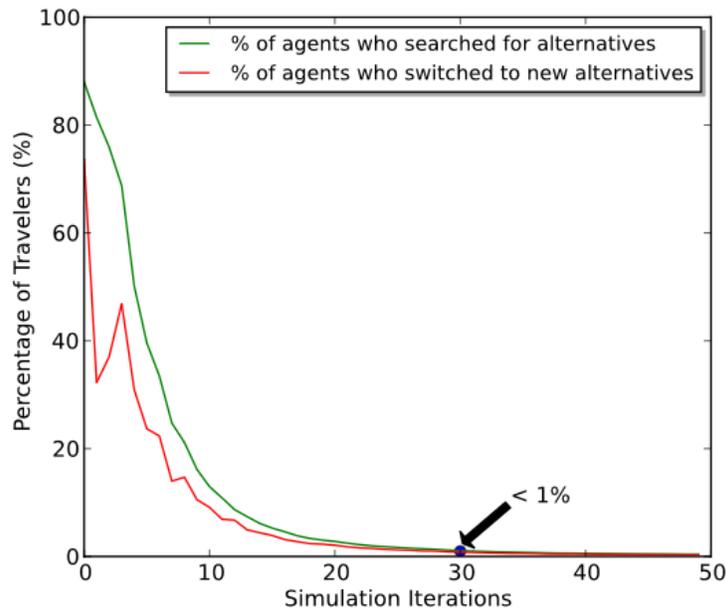
### 3.7 AgBM-DTALITE DEMONSTRATION

The developed integrated AgBM-DTALite model is tested in a real-world case study using a mid-size transportation network. The study area is the White Flint region of Montgomery County, Maryland. White Flint lies between Rockville and Bethesda, two highly urbanized areas located in the Northwest quadrant of the Washington D.C. Metropolitan Area. I-270 and MD-355 serve the majority of northbound and southbound traffic in the area. There are a few east-west corridors, including Randolph Road. Multiple bus lines and [SEP] Washington Metrorail’s Red Line also serve this vivid area. Mixed land development and transit-oriented development are currently ongoing in White Flint. Business, residential and shopping units are being built in close proximity to the Metro Station. While reshaping a dense and multi-functional urban region, the traffic impacts caused by these development projects draw attention and need to be analyzed.

24 traffic analysis zones, 55 roadway links, and 136 nodes are included in this network. A total number of 40,140 traveling agents are generated to represent the travel demand pattern in the morning peak hours of a typical workday.



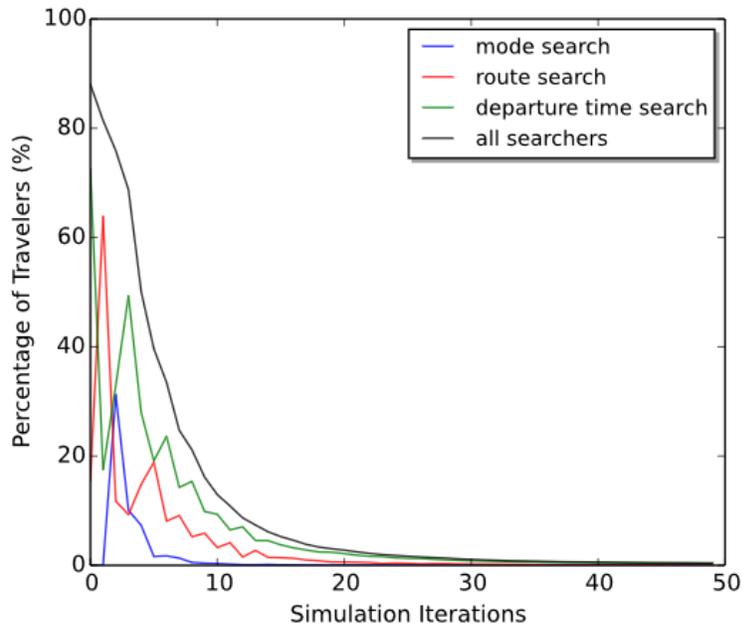
**Figure 3.27 Study area for demonstrating the integration of SILK-AgBM and DTALite**



**Figure 3.28 The convergence of the AgBM-DTALite on the demonstration network**

A dynamic user equilibrium cannot be achieved within 50 iterations when running a typical DTA with this size network and demand files. In fact, the larger the network, the greater number of iterations are required to reach DTA convergence. Based on the proposed SILK multidimensional travel behavioral model, another equilibrium, behavioral user equilibrium (BUE), has been defined in this research as the situation where all agents stop making behavioral adjustments. Initially, travelers will follow the travel option that yields the lowest generalized cost. Congestion during

the a.m. peak hours results in the discrepancy between the expected and realized travel conditions. As a result, over 70 percent of travelers decide to search, learn and adapt to the network by adjusting modes, departure times and/or routes. Among these travelers, more than half are searching routes. As time goes by, agents reach satisfaction either because a more promising travel alternative has been identified or because of the decreasing expectation on travel condition after excessive searching. Therefore, the number of searchers decreases. After ten iterations, only a very small number of users are still actively searching for alternatives. If we define the convergence criteria of BUE as 1 percent of the travel behavioral gap, the integrated model reaches convergence after thirty simulation iterations. Defined by the SILK theory and bounded rationality, BUE convergence is guaranteed regardless of the size of the network and the scale of the study.



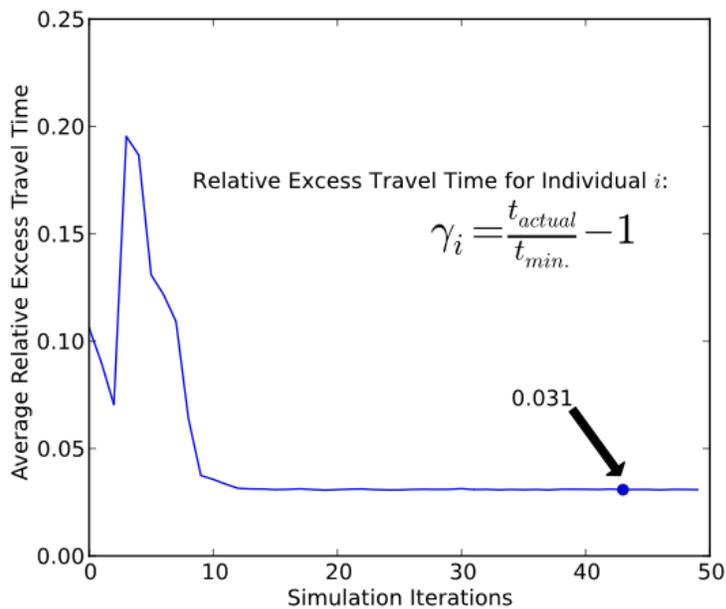
**Figure 3.29 The convergence of the day-to-day integration of AgBM-DTALite**

In response to the actual experienced congestion level in the suburban Washington D.C. metropolitan area, changing route and changing departure time are the most significant ways of behavioral adaptation. The initially high route searching frequency tapers off rapidly, since agents have trouble identifying better alternative routes under the overall peak-hour congestion condition. Agents quickly learn the fact and update the subjective beliefs, which results in a decreasing search gain in the route dimension. Instead, more agents turn to search alternative modes and departure times. Thus we can observe in the simulation an increasing number of agents searching for alternative departure times in the second and third simulation days. A few agents search for alternative modes.

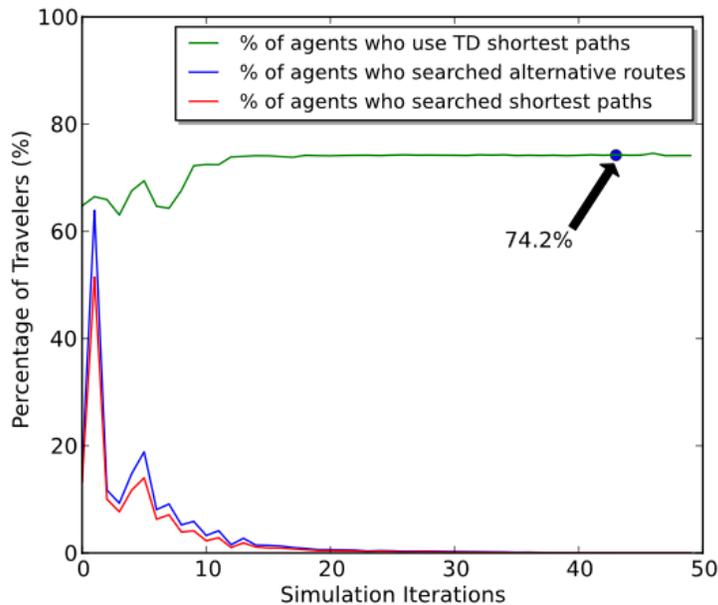
Travelers in the AgBM-DTALite model are not perfectly “rational,” in that they do not minimize travel time. They are restrained by information acquisition cost, decision cost, computational limitation, time budget and deadlines. They also follow intuitive and heuristic behavioral rules, which together represents a descriptive approach. If planners and engineers take a normative view

on system performance, the non-rational behavior may have a cost because it is not at optimum defined by normative standards. If minimizing system-wide travel cost is the normative goal from a system manager’s perspective, the cost of not being “rational” is the excess travel time, defined as the difference between the actual travel time and the minimum possible travel time on the network. The relative excess travel time is the excess time normalized by the minimum possible time (i.e., actual minimum divided by minimum).

When many travelers search and change behavior in an ever-changing environment, the traffic pattern shifts continuously and it is difficult for individuals to find the best alternatives. The average relative excess travel time of all modeled travelers is very high during the first several iterations. As more and more travelers stop searching and switching, a steady decrease in the relative excess time is observed. At the BUE, the average relative excess travel time is 3.1 percent, which implies that an average traveler could reduce travel time by 3.1 percent in the demonstration network, provided that other travelers do not subsequently change behavior.



**Figure 3.30 Convergence of the average relative excess travel time**



**Figure 3.31 Convergence of the route search and switching behavior**

Similarly, not being perfectly rational also indicates travelers' limited capability of identifying and switching to the theoretically shortest paths. In this small demonstration model, only a small percentage of travelers are able to identify time-dependent (TD) shortest paths as their alternative. Overall, there are only 74.2 percent of travelers that have ever tried or used TD shortest paths before they stopped behavioral adjustments. It is believed that this number is even higher than normal situations; in this small demonstration network, there are a limited number of alternative paths. In a much larger network representing real-world situations, there often exist a large number of alternative paths between each OD pair. It is expected that travelers cannot always identify a TD shortest path when they plan their daily travels.

Other than the behavioral foundation and the convergence property, another merit of the proposed AgBM-DTALite lies in its superior computational efficiency when compared to typical disaggregated travel demand models. Two unique characteristics of the integrated model ensure a promising computational performance:

- Without the time-consuming log-sum calculation, learning, searching and decision rules can be executed within relatively shorter CPU time.
- BUE changes the way of defining relative gaps and thereby reduces the number of simulation iterations.

It is important to note that the second characteristic does not differ with respect to the size or the scale of the system. Unlike DTA models that have an exponentially increasing number of alternative paths w.r.t. network size, AgBM-DTALite assumes agents neither have the capability nor are willing to consider every alternative. BUE only relies on the individual's travel experience and the information gathered to determine the start and stop of each search. Therefore, AgBM-

DTALite can maintain its computing performance even if applied to a very large-scale transportation analysis.

The integration of SILK-AgBM and DTALite has been demonstrated in this section. Multidimensional behavioral modeling and prediction are conducted by SILK-AgBM. Time-dependent traffic dynamics are simulated using DTALite traffic simulator. This integrated model is able to capture day-to-day behavioral adjustment in response to the evolving traveling conditions. Based on the SILK travel behavioral theory, another more behaviorally realistic equilibrium, BUE, is defined to replace UE and DUE. This newly defined equilibrium and its convergence property are also demonstrated in this illustration example.

## 3.8 Applications of SILK AgBM-DTALite

This chapter describes the applications of SILK AgBM-DTALite integrated modeling for MITAMS. As elaborated in the previous chapters, the SILK AgBM-DTALite integrated model emphasizes individual-level Search, Information acquisition, Learning, and Knowledge updating (SILK), which are collectively modeled by an agent-based model (AgBM). It is then integrated with a lite-weighted traffic simulation engine, DTALite, wherein second-by-second traffic dynamics are simulated. This integrated modeling system has several advantages. First, it is a truly agent-based simulation with well-represented traffic dynamics and multi-dimensional travel behavioral sensitivities, including departure time adjustments, route changes, modal shifts, and en-route diversion. Second, several unique features of the integrated model ensure a relatively more efficient computational performance. The behavioral user equilibrium (BUE) embedded in the SILK-AgBM presumes a behaviorally sound converging process, which largely reduces the number of iterations for the system to reach equilibrium. The DTALite traffic simulator employs parallel computing technique and an efficient traffic simulation engine, which enables large-scale computation for transportation networks with much greater size.

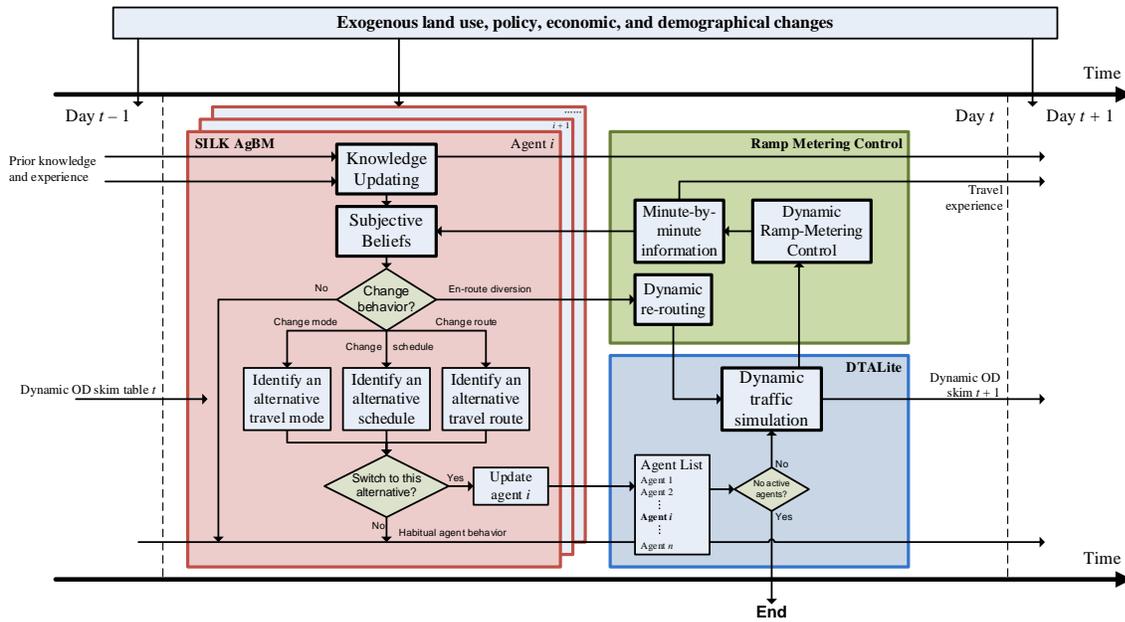
In order to demonstrate the capabilities of the SILK AgBM-DTALite, and showcase how it can be applied to different transportation planning and operations analysis, the research team presents two application studies in this chapter: 1) Prince George Plaza road-diet project; 2) dynamic ramp metering on MD-100 Corridor.

### 3.8.1 Methodology

Agent-based travel behavioral models are capable of mimicking and simulating travel behavior changes of each user in the system. SILK-AgBM is an agent-based travel behavioral model system that emphasizes the role of Searching, Information, learning and knowledge. It can also be efficiently interfaced with a dynamic traffic assignment model, such as DTALite. Once integrated with a traffic simulator, the system can be complete, given that the simulator can replicate all traffic conditions in the transportation network. The AgBM-DTALite framework is well-suited for the modeling and evaluation of various types of scenarios. While detailed modeling specifications can be found elsewhere in this final report, key features are highlighted as follows:

- The DTALite module (i.e., an open-source Light-weight Dynamic Traffic Assignment and Simulation Engine, <https://code.google.com/p/nexta/>) can simulate the overall traffic evolution on a day-to-day basis, which provides the foundation of evaluating mid/long-

term transportation planning scenarios, such as future-year land development plans, road-construction and transit-oriented development (TOD) strategies. It is also capable of simulating dynamic ramp metering by controlling the link attributes in a time-dependent manner (Zhou and Taylor, 2014). It simulates traffic dynamics in greater detail and estimates various time-dependent traffic conditions such as volume, density, etc. In a simulation environment, these pieces of information in a minute-by-minute setting are critical to modeling/controlling operational strategies dynamically, such as the ramp metering rates.



**Figure 3.32 The integrated model of SILK-AgBM and DTALite**

- Behavioral responses in departure time, travel mode, and route adjustments can be modeled and assessed with the integrated model of SILK-AgBM and DTALite. Between simulation days, agents arrange their daily or recreational itinerary based on knowledge and various information sources: previous experience, social network, mass media, real-time traffic data sources (e.g., Google, WAZE and INRIX), etc. Exogenous changes, such as the road diet construction and the implementation of ramp metering, may result in a different adjustment to the travel itinerary. AgBM models the travel behavior with the full consideration of search, information, learning, and knowledge. Within the same simulation day, traffic information is conveyed between AgBM and DTALite. Agents who possess real-time information could react to the dynamic ramp metering when driving (Xiong and Zhang, 2013). This en-route diversion is incorporated into this integrated model (Amini, et al., 2015). An agent's travel experience under different scenarios will both inform and update their knowledge base, as well as alter their subjective beliefs/expectations about travel conditions. A higher expected gain compared to the search cost (modeled as the mental/physical effort spent on searching) will lead to behavioral adjustments.

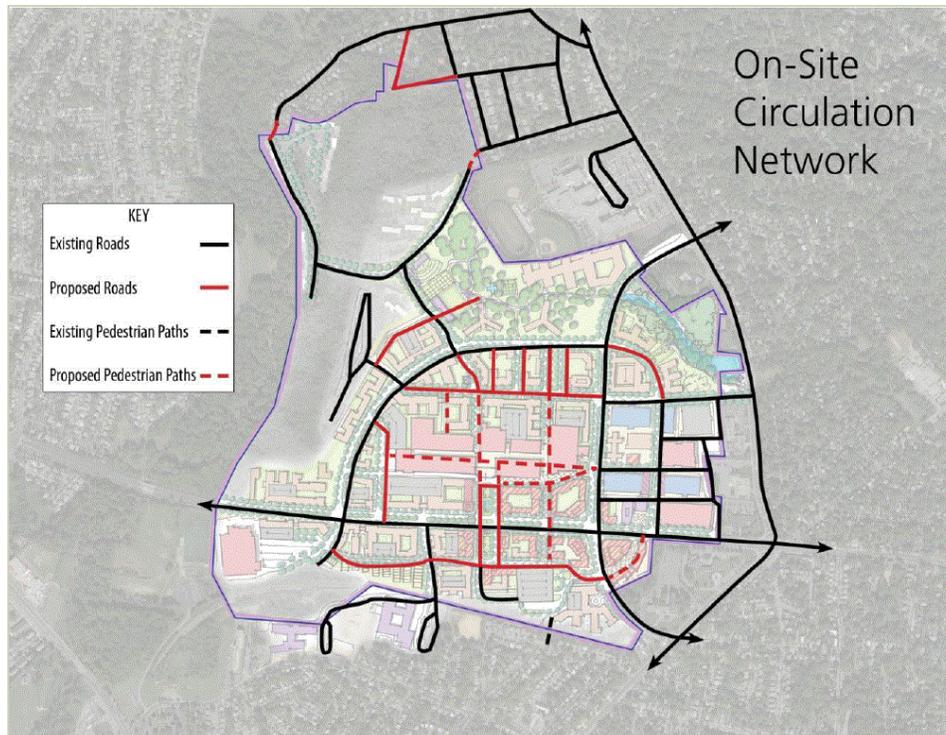
- A Behavioral User Equilibrium (BUE, where all users stop seeking behavioral changes) is developed to guide the model convergence process. The existence of BUE indicates that travelers will only search and adjust their behavior in limited times when responding to different long-term and short-term scenarios. Compared to traditional user equilibrium or dynamic traffic assignment, this BUE process incorporates more behavioral dimensions (route choice, departure time, and travel mode choice) and is guaranteed to converge.

### **3.8.2 Application 1: Transit Oriented Development (TOD) and Road-Diet Transportation Planning**

The integrated SILK AgBM-DTALite model is first applied to analyze a typical transportation planning scenario: a land use development plan. New land use development, such as new residential areas, commercial areas, and mixed land use areas will change the travel demand patterns by influencing the incoming and out-going traffic. In some cases, the scenario may also involve the removal, addition, and/or reconstruction of roadways. In addition to changing the origin-destination travel demand in the traditional travel demand analysis, these land use changes could lead to more dynamic and individual-level travel behavior changes, including modal shifts, departure time adjustments and route changes. This application applies the SILK AgBM-DTALite model in an analysis of a land use transportation planning scenario where transit oriented development (TOD) and road-diet are involved. The following subsections describe the scenario defined in this application, the applied models, and the numerical findings from AgBM-DTALite.

#### **3.8.2.1 Scenario definition**

Prince George's (PG) Plaza in Prince George's County, MD has long been a major destination for leisure, recreational, and commuting purposes. The plaza is a mixed-use development consisting of a shopping center, multimedia entertainment amenities, retailers, dining places, and a multimodal transportation hub with a Washington Metrorail station and its associated park-and-ride facilities. Recently, a transit oriented development (TOD) plan was proposed to the PG Plaza neighborhood. Several proposed roads and pedestrian paths would connect the community and enhance the accessibility and walkability. The major arterial, East-West Highway (MD-410), would undergo a road-diet re-construction, reducing the total number of traffic lanes from six lanes to four lanes. These developments have been proposed to calm traffic and encourage multimodal and eco-friendly travel.



**Figure 3.33 On-site circulation network for the transit oriented development (TOD) and road-diet planning at the Prince George's Plaza**

To apply SILK AgBM-DTALite and analyze the transportation impact of this proposed TOD and road diet plan, we must first define the three scenarios:

- (1) No-Build Scenario: this benchmark scenario is defined to set the baseline. The network and travel demand will be extracted from a regional transportation planning model (i.e., the Metropolitan Washington Council of Governments model), through a path-based subarea analysis method developed by the research team (Zhang et al. 2013). The models are calibrated and validated based on data collected from the year of 2015 using the calibration/validation framework (details can be found in the calibration/validation chapter of this report).
- (2) TOD/Road Diet without Behavioral Modeling: based on the No-Build Scenario, we redefine the transportation network based on the land development plan. Traveler's multidimensional behavioral adjustments are not considered.
- (3) TOD/Road Diet with AgBM-DTALite: based on the No-Build Scenario, we redefine the transportation network based on the land development plan. The multidimensional travel behavioral adjustments of the travelers in the region will be modeled using the AgBM-DTALite system, in order to assess the modal shifts, departure time changes and route adjustments of the impacted travelers.

### **3.8.2.2 Model Components of the Application**

As mentioned earlier, this network, along with the travel demand in the No-Build Scenario, is extracted from the MWCOG regional transportation planning model. The study contains over

55,000 travelers during the study time period (3 – 7 p.m.). This small to mid-size transportation network involves 640 links (major arterials, minor arterials, collectors, local roads and neighborhood roads), 228 nodes (i.e., intersections with or without signals/signs), and 81 traffic analysis zones (TAZ, 56 internal TAZs and 25 external TAZs).



**Figure 3.34 The transportation network for the SILK AgBM-DTALite Application 1**

The transportation system in this application also includes multiple bus lines and the D.C. Metrorail, which can be accessed from the Prince George’s Plaza Metro Station. These public transit options are included using the General Transit Feed Specification (GTFS) data, including the information of transit line geometry, schedule, station nodes, etc. This is used to produce transit skims (travel time and costs) for the SILK AgBM-DTALite (i.e., agents in the model could learn the transit skims information through their own experience and make a travel behavior decision according to the learning). Walking and biking skim information is presumed for trips with distance shorter than three miles and five miles, respectively. These modeling components are then plugged into the integrated AgBM-DTALite.

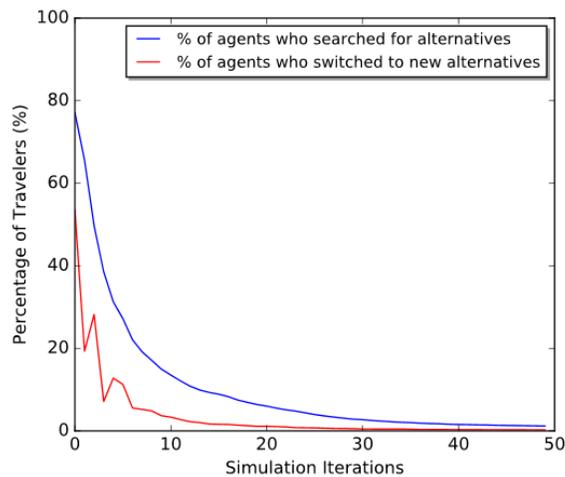
### **3.8.2.3 Results**

Overall, PG Plaza could face additional traffic and higher congestion from land use development and the road diet. However, travelers may adapt by adjusting their own behavior, such as switching to shoulder hours instead of peak hours, switching to public transit or active modes etc. If there is an organic mechanism that functions well in incentivizing peak spreading, smart-mobility travel options and eco-friendly travel, the negative impact on the roadway system can be greatly mitigated.

**Table 3.8 Simulated Traffic Impact Summary of Application 1**

Scenarios:	No-Build Scenario	TOD/Road Diet Without AgBM-DTALite	TOD/Road Diet With AgBM-DTALite
<b>Network-Wide Avg. Travel Time (min.)</b>	6.36	6.58	6.41
<b>Difference</b>	-	+3.5%	+0.8%

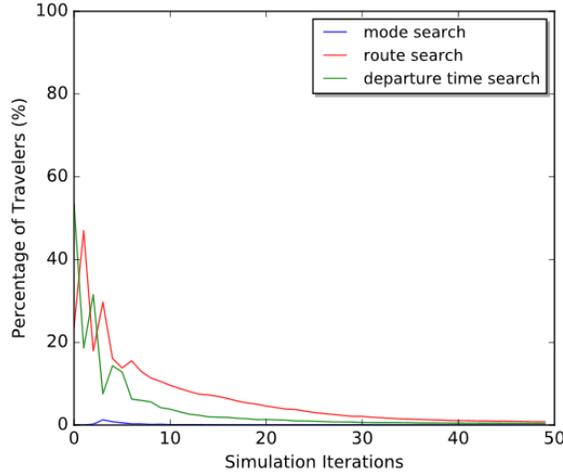
The behavioral user equilibrium (BUE) has been demonstrated in this application. The higher level of congestion in the defined TOD/Road Diet scenario results in higher initial search gain for many agents. Therefore, over 70 percent of agents are searching for new solutions. The PG Plaza region is usually very congested during p.m. Peak Hours. Travelers are having difficulty in identifying promising alternatives. As the model predicts, significantly fewer agents have made decisions to switch to new alternatives. As time goes by (day-by-day traffic evolution and agent-based learning), agents reach satisfaction either because a more promising travel alternative has been identified or because of the decreasing expectation on travel conditions after excessive searching. After about 40 iterations (40 simulation days, representing roughly 8 weeks), only a small number of users are still actively searching for alternatives. The integrated SILK AgBM-DTALite reaches convergence after 46 simulation iterations. This prediction suggests that it could take about two months for the travelers in PG Plaza region to fully adapt to the TOD and road diet, if the proposed plan is to be realized.



**Figure 3.35 The convergence of the AgBM-DTALite for application 1**

This integrated modeling system can simulate travelers’ multidimensional behavioral responses. Changing route and changing departure time are the most significant ways of behavioral adaptation. A small percentage of agents has been considering modal shifts. At BUE, 0.7 percent of the agents

have changed from car/carpool to other modes (0.34 percent switch to public transit, 0.21 percent switch to walk, and 0.15 percent switch to bike).



**Figure 3.36 The multidimensional behavioral responses predicted by SILK AgBM-DTALite in application 1**

### **3.8.3 Application 2: Real-Time Ramp Metering Control for MD-100 Corridor**

#### **3.8.3.1 Algorithms for Real-Time Ramp Metering Rates**

Based on the integrated AgBM-DTALite modeling framework, a dynamic ramp metering control module is added to the system. The control algorithm employs the time-dependent inputs produced by the dynamic traffic simulation to control the metering rates of those affected ramps. Then, simulated traffic dynamics are fed to the agent-based travel behavioral model to update the travel behavior decisions. While the process of developing the linkage between dynamic traffic assignment models and real-time ramp metering control is documented in (Zhu et al., 2017) in detail, the research team chose to outline only the key components in this paper. The dynamic ramp metering rate of a particular freeway on-ramp is controlled by a series of equations derived using the ALINEA algorithm. Based on the dynamic traffic assignment performed in DTALite, minute-by-minute density, queue length and volume of each link are simulated. These pieces of information are then employed in ALINEA. The employed ALINEA algorithm estimates the time-dependent ramp metering rate  $r(t)$  at time  $t$ :

$$r(t) = r(t-1) + K_R [\hat{O} - O_{out}(t)] \quad (1-a)$$

$$\text{s.t. } r_{\min} \leq r(t) \leq r_{\max} \quad (1-b)$$

In this formulation,  $K_R$  denotes the regulatory parameter;  $\hat{O}$  denotes the critical occupancy;  $O_{out}$  denotes the downstream occupancy; and  $r_{\min}$  and  $r_{\max}$  are the predefined minimum metering rate and maximum metering rate, respectively. ALINEA/Q algorithm extends the ALINEA method to further consider the queue length on the freeway mainline.

$$r'(t) = -\frac{1}{T} [\hat{w} - w(t)] + d(t-1) \quad (2-a)$$

$$r''(t) = r'(t-1) + K_R [O' - O_{out}(t)] \quad (2-b)$$

$$\text{s.t. } r_{\min} \leq r(t) = \max \{r'(t), r''(t)\} \leq r_{\max} \quad (2-c)$$

In this formulation,  $T$  denotes the time interval;  $\hat{w}$  denotes the maximum permissible queue length on ramp;  $w$  denotes the current queue length on ramp; and  $d$  denotes the demand flow entering a ramp. This formulation allows ALINEA/Q to further consider the required entering rate ( $r'$ ) from ramps to freeway mainline in order to prevent excessive queues.

### 3.8.3.2 Application Network, Subarea, and Ramp Metering Locations

The integrated AgBM-DTALite model has been applied to the Washington, D.C. and Baltimore Metropolitan Regions shown in Figure 3.37, including the majority of Montgomery County, Prince George’s County and Howard County. and part of Anne Arundel County, Frederick County and Baltimore County. Critical corridors in the region, such as I-95, I-495 (the Capital Beltway), I-695 (the Baltimore Beltway), as well as the toll facility (MD-200, the Intercountry Connector) are all included in the model. The simulation system takes care of nearly 10,000 links, 4,300 nodes, 712 traffic analysis zones and 2 million agents. The initial origin-destination travel demand information is extracted from the regional transportation planning model developed by the Metropolitan Washington Council of Government (MWCOCG). Calibration methodology and results are documented in details in the authors’ other working papers (Lee et al, 2017; Zhu et al, 2017). The mean squared error (MSE) rate after calibration is controlled at 15.3 percent. The computational time for running this simulation system is around 600~720 sec per iteration on a regular server (8-core CPU and 16GB RAM). It may take 150 to 200 minutes for the entire system to reach convergence (i.e., BUE).

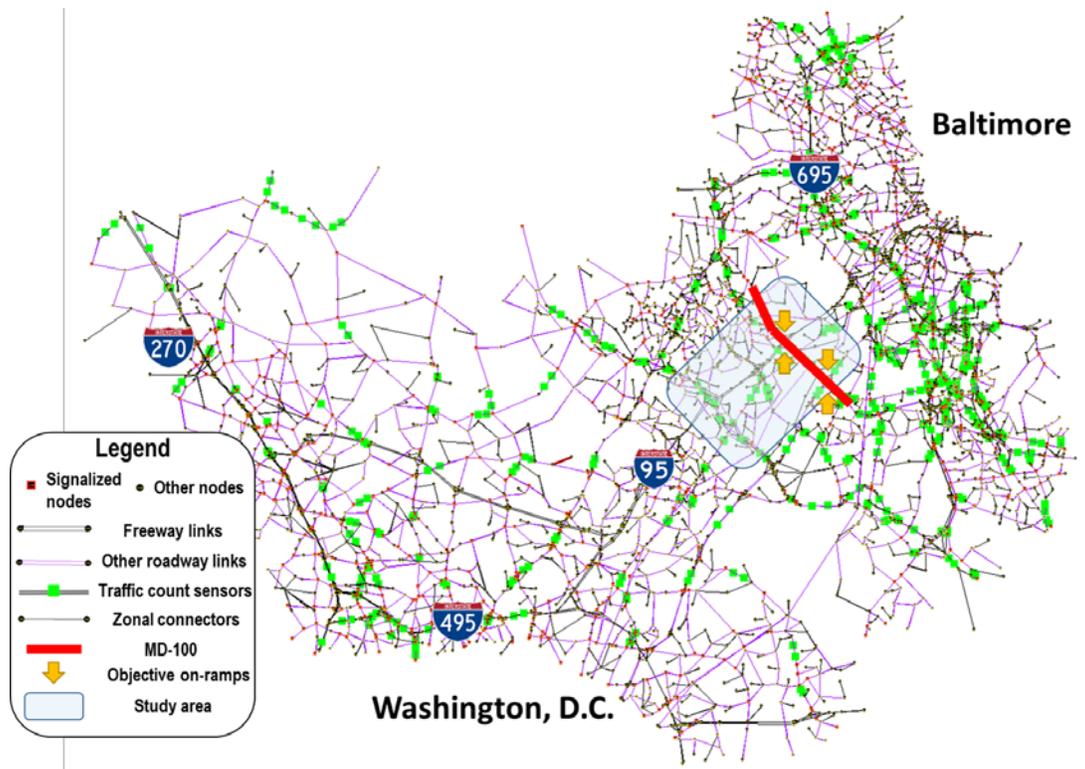
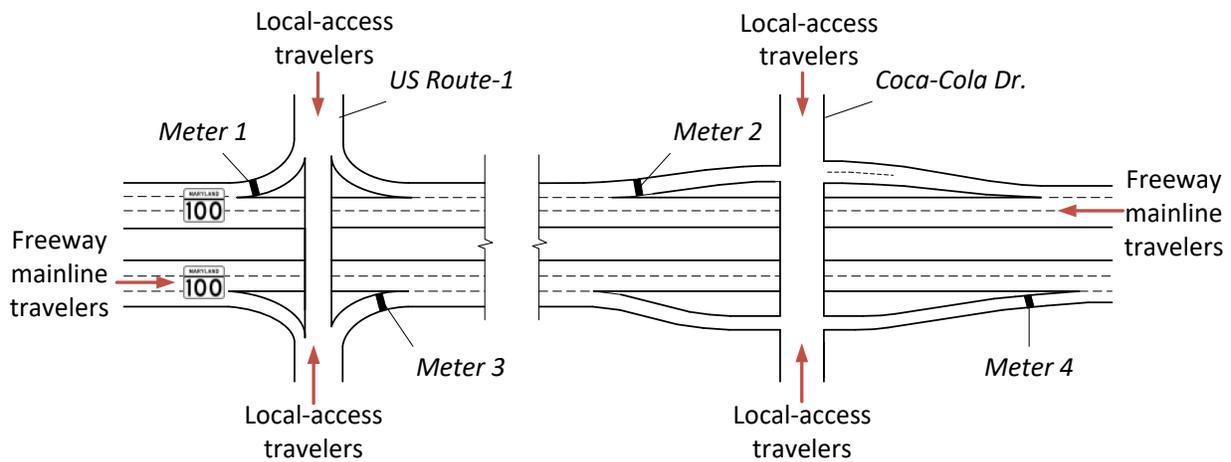


Figure 3.37 Regional network and the subarea cut for the ramp metering analysis

This large-scale regional model is used as the basis for the ramp metering analysis. MD-100 is one of the critical arterials serving east-west travelers in the Baltimore-Washington commuter corridor. The team focuses on the subarea near MD-100 (the area is identified by the square in Figure 3.38) in order to use the agent-based microsimulation to quantify the direct impact of ramp metering. The subarea model involves a number of highly congested roadways including I-95, MD-100, US-1, MD-32, and MD-295 (Zhu et al., 2017) and over two hundred thousand travelers in the p.m. peak hours (4:00-7:00 p.m.) under the direct influence of the ramp metering control scenarios.

Four on-ramps to MD-100 are metered. On MD-100 Westbound and Eastbound, the ramp metering system controls the on-ramps linked to Washington Blvd (US Route-1) and Coca-Cola Dr. Those are heavily used on-ramps, especially during peak hours.



**Figure 3.38 A detailed layout of the ramp metering locations and local-access and freeway-mainline traffic under the impact**

Under the influence of ramp metering, local-access travelers from US Route-1 and Coca-Cola Dr. are more likely to face a higher level of congestion than before. The effective ramp meters can cause excessive delays on the ramps and the on-ramp queues may even spillback to local roads. In contrast, freeway mainline travelers who travel through this area via MD-100 have a stronger possibility of light traffic. Since the meters restrain the incoming vehicles, the vehicle density is reduced and disturbance from weaving, merging etc. is also mitigated. How will these different agents respond to the ramp metering control? What are the impacts on their short- or long-term behaviors? These findings can be obtained from the simulation and will be summarized in the following sub-sections.

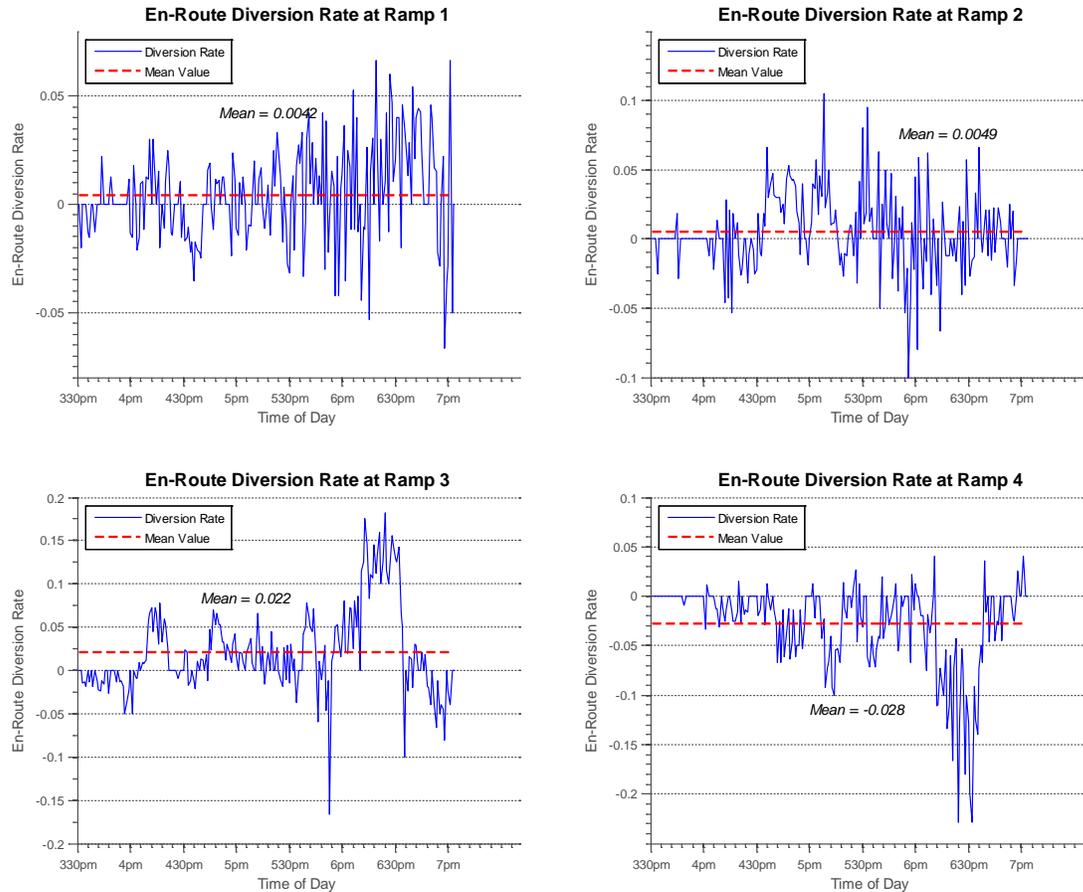
### **3.8.3.3 Dynamic Ramp Metering and Within-Day En-Route Diversion Responses**

Route changes and dynamic en-route diversion are modeled on a minute-by-minute basis. The agent-based microsimulation assumes that 19.8 percent of the agents possess real-time traffic information. This number is based on a parallel travel behavior study conducted by the research

team (Xiong et al., 2016). Summarized from 2,050 representative samples, it finds that 19.8 percent of the daily travelers residing in D.C.-Baltimore metropolitan regions are accessing real-time traffic information via smartphone apps, radio and other information sources while traveling. Therefore, these travelers can dynamically switch routes if they observe notable delays on their normal routes. This percentage is presumed in the simulation. It can be gauged when ground truth data on information-provision rate was observed or obtained. Based on the real-time traffic information, agents in the simulation model will respond and make an en-route diversion decision (simulated en-route diversion rates at the four metered ramps are illustrated in Figure 3.39). This decision is governed by a set of decision rules estimated and calibrated based on actual behavioral data (Xiong and Zhang, 2013). A positive en-route diversion rate (i.e., the percentage of agents who change routes to divert from this interchange) indicates that more agents decide not to use the ramp and dynamically switch to other routes. A negative rate indicates more agents are switching their route to the particular ramp.

It is preferable that the AgBM-DTALite model does not lose too much resolution in en-route diversion rates. We can observe this kind of fluctuation minute-to-minute and over longer periods as well. This can be best explained by the behavior of well-informed travelers, who regularly update their information based on the current traffic conditions and make diversion decisions dynamically. When the on-ramps are controlled under a higher metering rate, resulting in longer on-ramp queues, more travelers will divert to other alternatives. This leads to a positive en-route diversion rate and in return, helps the congestion dissipate. The highly-congested freeway mainlines of MD-100 and the ramp metering control, in general, encourage en-route diversion (the mean diversion rates are positive in three out of the four on-ramps). Nevertheless, after a certain amount of time, the higher metering rate will have an enduring effect in managing the incoming flows to the freeway segments. Reduced density on the freeway mainline will trigger a decreased metering rate. The mitigated congestion on the freeway mainline, as well as the dissipated on-ramp queues, will again attract informed travelers to utilize these on-ramps. This will dynamically lead to a negative diversion rate.

Another interesting finding from the agent-based simulation is the spatial correlation of the diversion behavior. The most obvious case is ramp 3 and 4 between 6 p.m. and 6:30 p.m. ramp 4 is adjacent to ramp 3 and at the downstream. Between 6 p.m., and 6:30 p.m., a positive and significant diversion rate is observed on Ramp 3. Meanwhile, agents show a negative diversion rate on ramp 4. One explanation could be that the high diversion rate at ramp 3 (10-15 percent) results in improved traffic conditions on MD-100 downstream segments, and therefore, agents who have access to real-time information are dynamically re-routed to MD-100 via ramp 4, causing a rebounding higher ramp usage.



**Figure 3.39 Agents’ en-route diversion rates at the freeway on-ramps where ramp metering control is effective**

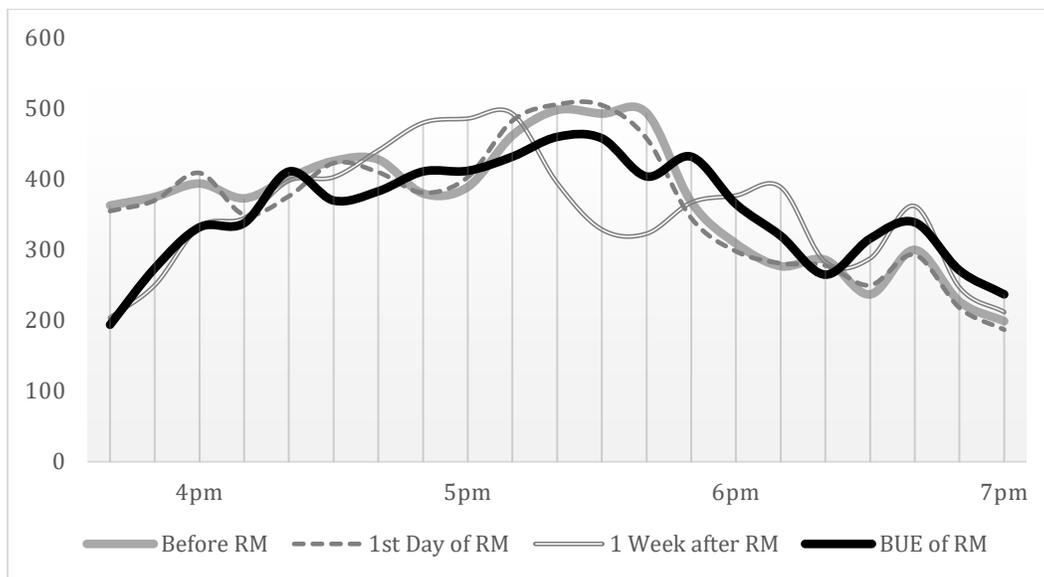
### **3.8.3.4 Day-to-Day Departure Time Shifts**

One unique aspect of applying the AgBM-DTALite model to evaluate ramp metering control, is that the day-to-day travel behavioral adjustments can be captured. This agent-based microsimulation sheds light on how travelers adapt to newly built infrastructures and implementations. Besides en-route diversion, another significant behavioral change in response to dynamic ramp metering control is day-to-day departure time adjustment. The two subgraphs show the influences on local-access travelers and freeway mainline travelers, respectively. Departure time adjustments in the short-term, mid-term and long-term are all displayed. The light blue curve, labeled “Before RM,” represents the status-quo departure time pattern (i.e., number of travelers departing at each simulation time period) before ramp metering (RM) is implemented. The departure time pattern of the first day of ramp metering and the pattern after one week’s practice is also displayed. Generally, it is observed that travelers’ departure time patterns on the first day of ramp metering are similar to those in the before-RM scenario, but become distinctly different after one week’s practice. Finally, the orange curve, “BUE of RM”, denotes the pattern after the model reaches convergence and behavioral user equilibrium (BUE) is reached. It implies that travelers may keep adjusting their departure time for a longer period than one week.

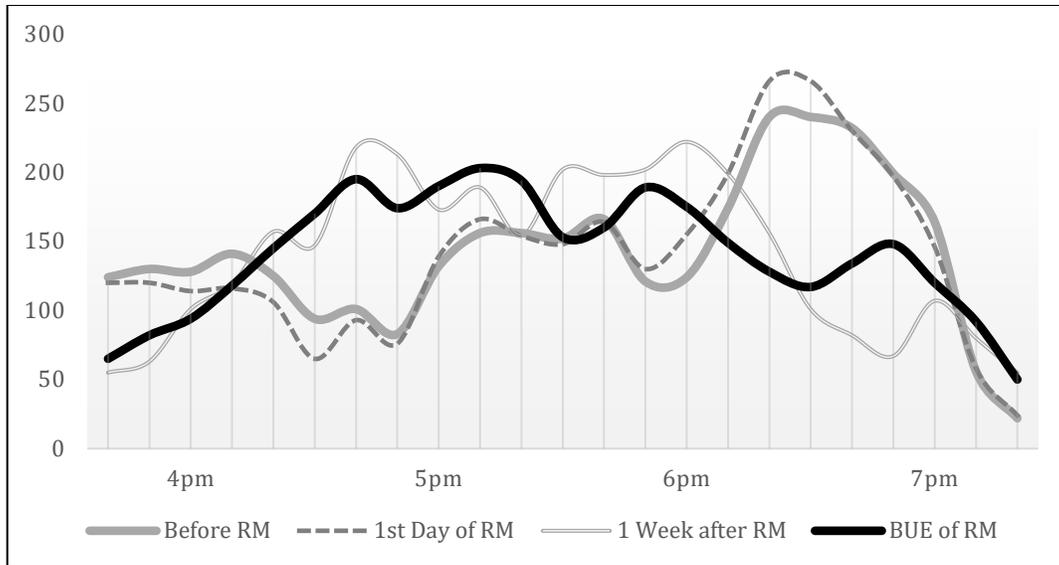
Agent characteristics such as preferred arrival schedule, flexibility and departure time search scope are incorporated in the integrated AgBM-DTALite model. The simulation result suggests quite a few travelers will change departure time to avoid the longest ramp delays appearing at around 5:30 p.m. On the first day of the exercise, there is a slight increase in travelers departing in the first half of the peak hour. One week after the RM implementation, a significant number of travelers decide to either depart earlier, at 5 p.m., or later, after 6 p.m., to avoid the emerging delays. Hence the number of travelers depart at 5:30 p.m.—the previous peak time—has dropped significantly. Last but not least, the convergence line leads to a milder pattern shift than the one-week change, and to a flatter and smoother departure time distribution, compared to the status-quo.

The agent-based model tracks different heterogeneous agent groups, including the freeway mainline users, in comparison to local-access travelers. Figure 3.40 (b) reveals the changes of departure patterns of those users on the MD-100 corridor. The freeway users are expected to benefit from the RM scenarios resulting from reduced incoming flows at the affected on-ramps. Mitigated congestion level during the p.m. peak (4-7 p.m.) has encouraged travelers who previously departed late in the status-quo model to leave earlier. After one week, a large number of travelers who initially departed between 6 p.m. and 7 p.m. have shifted to an earlier time (similar to an inverse process of peak spreading). When the new equilibrium under the RM is reached, the p.m. peak is pushed up to 4:30 p.m. The overall travel demand pattern is distributed more evenly.

Although travelers from local accesses and on the mainline have distinguished departure patterns, it can be derived from both figures that the ramp metering application on MD-100 encourages travelers to depart ahead of or after the peak hour, thus contributing to the mitigation of congestion on the mainline.



(a) Departure Time Adjustment of the Local-Access Travelers



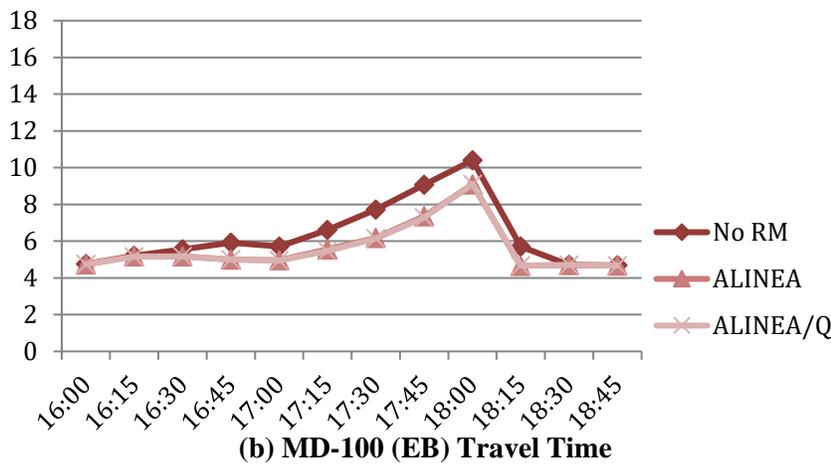
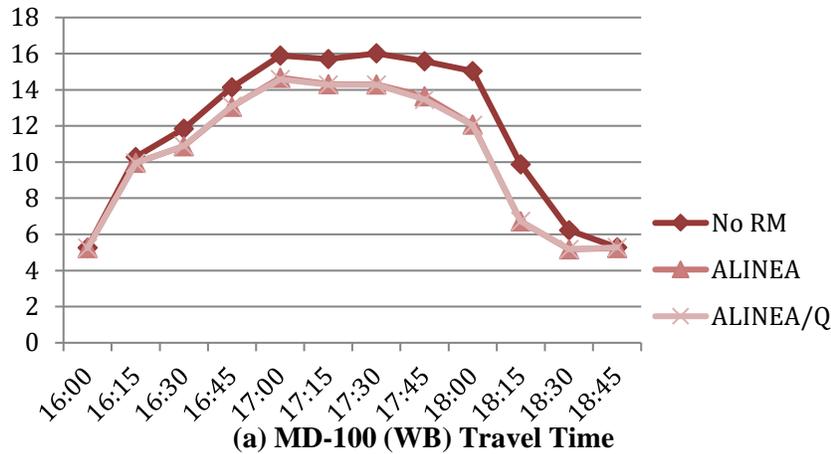
(b) Departure Time Adjustment of the MD-100 Freeway Mainline Travelers

**Figure 3.40 AgBM-DTALite simulated day-to-day departure time adjustments and peak spreading behavioral patterns**

### **3.8.3.5 Regional Traffic Impact of the Dynamic Ramp Metering**

Ramp metering analyses are typically conducted at the corridor level. With the integrated AgBM-DTALite model, we are able to extend the evaluation to the regional level, which is rarely seen in ramp metering analyses. In this section, we present the traffic impact of the dynamic ramp metering. Travel time is employed as the measure of performance in the case study.

The corridor-level traffic impact of the dynamic ramp metering control is illustrated in Figure 3.41. The effectiveness of ramp metering using ALINEA or ALINEA/Q is evident. During the p.m. peak, the average travel times on MD-100 corridor have been shortened by 1.32 minutes on the Westbound lanes and by 0.74 minute on the Eastbound lanes. A maximum decrease of 20 percent is realized on westbound MD-100 at 6:00 p.m. ALINEA and ALINEA/Q tend to present similar corridor-level travel times, since the two algorithms become differentiated only if there are queues exceeding the predetermined threshold on the controlled ramps. Although it is minor, ALINEA/Q leads to slightly longer corridor travel times on average than ALINEA, while it induces less waiting times for drivers on the controlled on-ramps.



**Figure 3.41 MD-100 travel time change**

Besides corridor-level direct impact, the integrated model also allows the evaluation of indirect and network-wide traffic impact. Table 3.9 summarizes the overall network performance. In addition to the substantial travel time decrease on MD-100, the average travel times among the network also experience moderate improvements. Although applying ALINEA or ALINEA/Q will only lead to about 0.6 percent travel time saving at the network level, aggregating that savings for all simulated agents in the region could result in 28,000 minutes of savings daily (or 170,333 hours of savings annually). Considering the average value of commuting travel times in the D.C.-Baltimore region, the ramp metering control scenarios considered in the study can bring significant indirect benefit for travelers, which agrees with the promising potential of applying dynamic ramp metering control in a relatively congested metropolitan area.

**Table 3.9 Simulated Traffic Impact Summary of the Dynamic Ramp Metering Control**

Scenarios:	Status-Quo (no ramp metering)	Ramp Metering with ALINEA		Ramp Metering with ALINEA/Q	
		Time	Difference	Time	Difference
Network-Wide Average Travel Time (min.)	23.03	22.89	-0.61%	22.90	-0.56%
MD-100 (WB) Average	11.76	10.44	-11.2%	10.42	-11.4%

---

<b>Travel Time (min.)</b>					
<b>MD-100 (EB) Average</b>					
<b>Travel Time (min.)</b>	6.34	5.60	-11.7%	5.59	-11.8%

---

## 4. Integrated InSITE-DTALite

This integrated model is being developed as part of the MITAMS C10 project, and has close ties with two agencies: Maryland State Highway Administration (MD SHA) and the Baltimore Metropolitan Council (BMC). This integrated model is composed of the InSITE activity-based model (ABM) system, and the mesoscopic traffic simulator, DTALite. These two components are described in this section. In addition, the key innovations of the integrated model are highlighted.

The primary objective of the model integration is to produce a working model of the Baltimore region that is disaggregate for auto travel on both the demand and the supply side. This model is intended to be “application ready,” meaning that it has been validated and includes all functionality required for typical planning analyses. Therefore, the integration pursued between InSITE and DTALite falls in the sequential integration paradigm. In this paradigm, the travel demand model (InSITE) and the DTA model (DTALite) are run in iterations, with a single “big loop” consisting of an iteration of the travel demand model followed by a run of the DTA model. The travel demand information (e.g., agents with their activities and travel decisions as well as characteristics) is passed to the DTA model within each big loop, and travel time information resulting from the dynamic assignment is passed back to the travel demand model for use as input (e.g., update activity pattern choices) in the next big loop. Note that DTA models themselves are run iteratively, and so there are several “small loops” of the DTA run within each “big loop.” A number of “big loops” are run until a measure of convergence is achieved, in which the change in travel times (or some other measure) from one big loop to the next is within a specific tolerance. The data exchange between InSITE and DTALite is such that an agent is preserved during the exchange of data from InSITE to DTALite and vice versa. Thus, the integrated InSITE-DTALite model is fully disaggregated. There is also the tight integration paradigm in which both the travel demand model and DTA model share information continuously (i.e., the analyst specifies the temporal resolution for the exchanges) within each big loop. This paradigm is not considered for the integrated model InSITE-DTALite.

### 4.1 BMC InSITE ABM

InSITE is an activity-based model system composed of interconnected, discrete choice models representing choices at distinct dimensions (e.g., travel mode, destination) that focus on decisions related to daily activity and mobility for a typical weekday. InSITE adopts the day activity-schedule approach, where a day activity schedule is defined through the concepts of activity pattern and activity schedule. The activity pattern defines the participation in activities as primary and secondary. Primary activities are the anchors (e.g., home-to-work trip and work-to-home trip represent a tour, with work as primary activity) of a tour, and secondary activities are intermediate stops within a particular tour (i.e., stopping for shopping during the work-to-home half of the tour). The activity schedule adds detailed information to the activity pattern about tours, such as timing, travel mode, destination of primary activity and the stops for secondary activities within the tours.

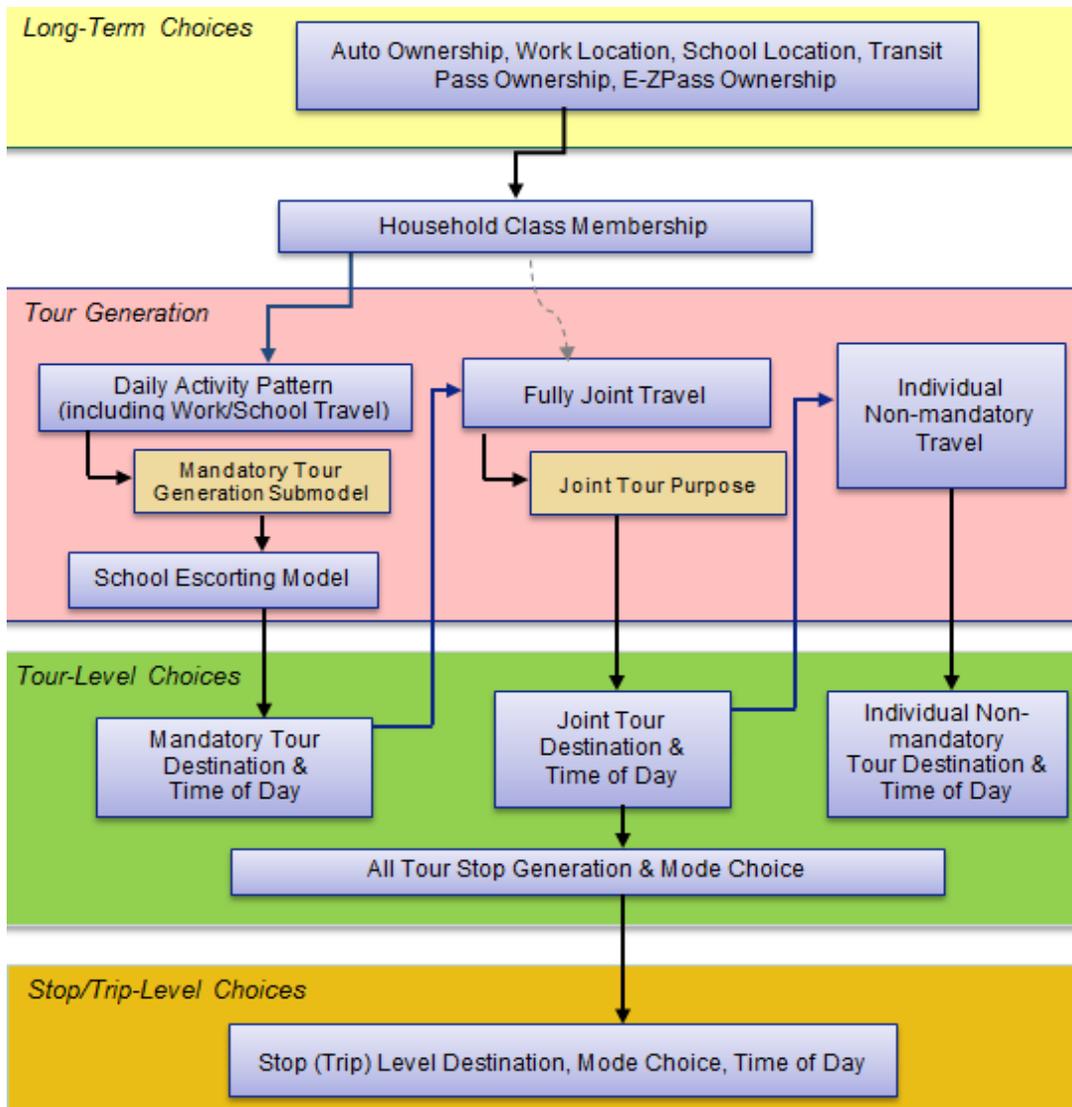
The model covers an area including the entire BMC region, plus the District of Columbia and the Maryland portion of the region covered by the Metropolitan Washington Council of Governments (MWCOCG). The portion of Maryland in the model region consists of Baltimore City and Anne

Arundel, Baltimore, Carroll, Harford, Howard, Frederick, Montgomery, and Prince George's Counties.

InSITE models travel for a typical weekday. The choices made by households and individual travelers are simulated using probabilities from a series of logit models. The model components and the sections of this report where they are discussed are listed in Table 4.1; Figure 4.1 shows the structure of the InSITE. (In the interest of keeping the figure readable, the logsum relationships from subsequent to previous model steps are not shown in the figure, nor are the nonactivity-based components, such as trip assignment.)

**Table 4.1 Components of InSITE**

Model Name	Level	What is Predicted
Synthetic Population Generator	Households	Household size and composition, household income, person age, gender, employment status, student status
Regular Workplace Location	Workers	Workplace location zone
Regular School Location	Students	School location zone
Auto Ownership	Households	Number of autos owned
Transit Pass Ownership	Households	Whether the household owns a transit pass
E-ZPass Ownership	Households	Whether the household owns an E-ZPass transponder
Household Class Membership	Households	Class in which household falls, for Daily Activity Pattern Model
Daily Activity Pattern	Person Day	0, 1, or 2 tours for each activity purpose 0, 1, or 2 stops for each activity purpose
Joint Travel	Households	Number of fully joint tours with 2 or more household members Which household members participate in each joint tour
School Escorting	Person (Household) Day	On which half tours a student is escorted to/from school Which household member escorts the student Whether escorting is done on a mandatory tour
Work Tour Destination Choice	Work Tours	For work tours – regular workplace or other work location (and its zone)
Work-Based Sub-tour Generation	Work Tours	Number and purpose of any sub-tours made during a work tour
Work Mode Choice	Work Tours	Main tour mode
School Mode and Time-of-Day Choice	School Tours	Main tour mode, the time period arriving at school and the time period leaving school (all school tours are to regular school location)
Work Time-of-Day Choice	Work Tours	The time period arriving at work, and the time period leaving work
Other Tour Time-of-Day Choice	Other Tours	Time period arriving at the primary destination and the time period leaving the primary destination
Other Tour Mode and Destination Choice	Other Tours	Primary destination zone and main tour mode
Intermediate Stop Generation	Half-tour	Number and activity purpose of any intermediate stops made on the half-tour, conditional on day pattern
Intermediate Stop Location	Trip	Destination zone of each intermediate stop, conditional on tour origin and destination, and location of any previous stops
Trip Mode Choice	Trip	Trip mode, conditional on main tour mode
Trip Departure Time	Trip	Departure time, conditional on time windows remaining from previous choices
Special Generators	Zone	Number of trips, trip end location, mode choice
Commercial Vehicle	Zone	Number of trips, trip end location
External Travel	Zone	Number of trips, trip end location
Highway Assignment	Vehicle Trip Table	Link volumes and travel times/speeds
Transit Assignment	Person Trip Table	Transit trips/boardings by route/stop



Source: Cambridge Systematics, Inc.

**Figure 4.1 InSITE model structure**

#### **4.1.1 Model Inputs**

The four main inputs into InSITE are:

- (1) Synthetic population: the synthetic population generator, PopGen, developed at Arizona State University (by a team led by MITAMS team member Ram Pendyala) produces a synthetic population representing every person and household in the model region.
- (2) Land use parcel database: the parcel database includes information on the land use type and size (square footage and/or residential units) for all land parcels in the model region.

- (3) Zone level socioeconomic and other data: the model region is divided into traffic analysis zones (TAZs), which are used as the analysis units for some model components.
- (4) Highway and transit networks: these are standard networks representing the major roadways in the region and all public transit routes. They are maintained using the Cube software, a product of Citilabs for which BMC, SHA and members of the consultant team have licenses. In the original InSITE model, which uses static equilibrium highway assignment, the highway network is used for the highway assignment process; for the integrated model, this will be replaced by the DTA. The static highway network is still expected to be used to create the initial travel time, distance and cost “skims” that are used as inputs to the demand model components.

### **4.1.2 Model Structure**

As shown in Figure 4.1, the model begins by simulating long term choices that are made before the travel day, including auto ownership, workplace location for workers, school location for students, transit pass ownership, and E-ZPass toll transponder ownership.

Another feature of InSITE is a class membership model that is applied prior to the daily activity pattern model. The class membership model determines which segment an individual household belongs to via a multinomial logit model. Each class membership model alternative represents a distinct segment, and the model uses attributes of the household (such as household size and income) to generate probabilities that the household is a member of a specific class. The parameters of the daily activity pattern model, the fully joint tour model and the class membership model were estimated simultaneously, since there are components of each model that affect one another. A more in-depth look at the class membership model is documented by Lemp (2014).

The next set of models estimates a daily activity pattern for each person. Whether the person has work tours (with or without stops), has school or university tours, has non-mandatory activity tours only, or does not travel within the region is simulated. If the simulated pattern has mandatory tours, they are generated. For students making school tours, InSITE simulates whether they are escorted by a household member and if so, by whom. At this point, the destinations and times of day for the mandatory tours are simulated.

Next, fully joint tours among household members are simulated, including who participates, the activity purpose, the destination and the time of day. After the details of the mandatory and joint tours are known, individual non-mandatory tours are simulated, with their destinations and times of day. The final tour level models are the generation of stops for each tour and tour level mode choice.

The stop and trip level choices are simulated next. These include destination choice for each stop, mode choice for the trips between stops (conditional on tour mode), and the times for the stops. From the results of these models, auto vehicle and transit person trip tables are assembled for input to highway and transit assignment, respectively.

## 4.2 BMC InSITE Validation and Sensitivity Testing

This section summarizes the validation of the activity-based model developed for the Baltimore region. This model was developed for the Baltimore Metropolitan Council (BMC) by a team led by Cambridge Systematics, Inc. (CS), and including Gallop Corporation, AECOM and Sabra-Wang Associates. The model estimation results are documented by Cambridge Systematics, Inc. (2016). User documentation is provided in a separate document.

The model is applied disaggregately using a synthetic population, generated by the PopGen synthetic population generator (Konduri and Pendyala, 2015). It represents the population of the model region, which includes the entire BMC region, the District of Columbia and the Maryland portion of the region covered by the Metropolitan Washington Council of Governments (MWCOG). The portion of Maryland in the model region consists of Baltimore City and Anne Arundel, Baltimore, Carroll, Harford, Howard, Frederick, Montgomery and Prince George's Counties.

The activity and travel choices made by each household and person in the synthetic population are realized through Monte Carlo simulation, with the choice probabilities determined by the individual model components.

A model validation plan (Cambridge Systematics, Inc., 2014) was developed prior to model development. This plan outlined the process that was followed for the model validation and the specified the tests performed by the research team. A few tests changed slightly or were more specifically defined for the final model validation, but in general, the plan was faithfully followed. The tests in the plan included verification of the input highway, the transit skim data and the synthetic population data; checks of the results of all model components compared to the 2007-2008 regional household survey data set; checks of the highway and transit assignment; and tests of the sensitivity of the model to changes in input data.

### 4.2.1 Model Component Validation

The tests consisted of comparisons of model results for various market segments to the expanded household survey data. These comparisons were done in Excel spreadsheet files. The model application software, TourCast, outputs .dbf files that were imported into a relational database and processed with stored procedures using MySQL. The processed summaries were exported to comma delimited files that can be read directly into the Excel spreadsheets, which were populated in advance with the survey data results. The model results presented in this chapter are based on a model application with three iterations of speed feedback.

The comparisons described in this section reflect model calibration adjustments. In some cases, model parameters were adjusted to produce more reasonable results, although there was not a universal attempt to match all results from the expanded household survey for all market segments by adjusting model constants or other parameters. This type of adjustment was only made when the uncalibrated model results did not appear reasonable and the survey data results were based on

a substantial number of observations. The specific calibration adjustments are documented in the Excel files.

Because of the extensive number of comparisons, the spreadsheet files themselves are incorporated as appendices to this report. The remainder of this chapter summarizes the validation results as presented in these spreadsheet files.

#### **4.2.1.1 Long-term Choice Models**

##### **Vehicle Availability Model**

The vehicle availability model simulates the number of vehicles owned by each household in the synthetic population. The Excel file with the results of the vehicle availability model is *VehicleAvailability.xlsm*. On a regional basis, the number of households by number of vehicles owned matches well.

**Table 4.2 Vehicle Availability Model – Regional Validation**

<b>Vehicles</b>	<b>Expanded Household Survey Data</b>		<b>Model Results</b>		<b>Percentage Point Difference</b>	
	<b>Households</b>	<b>Percentage</b>	<b>Households</b>	<b>Percentage</b>	<b>Point Difference</b>	<b>Percentage Difference</b>
0	231,695	11.2%	208,160	10.0%	-1.2%	-10.3%
1	690,202	33.3%	652,543	31.4%	-1.9%	-5.6%
2	753,072	36.3%	806,458	38.8%	2.5%	6.9%
3+	398,131	19.2%	409,036	19.7%	0.5%	2.6%
<b>Total</b>	<b>2,073,100</b>		<b>2,076,197</b>			

##### **Regular Workplace Location**

The regular workplace location model simulates whether each worker in the synthetic population has a regular workplace and the location of that workplace. The Excel file with the results of the regular workplace location model is *UsualWork.xlsm*.

The more detailed comparisons in the Excel files show the following results:

- The modeled percentage of household vehicle ownership closely matches the survey data for each county in the model region. The model slightly underestimates vehicle ownership in Carroll County, the smallest county in the model region.
- Vehicle availability levels were compared for cross-classifications of household size (1, 2, 3, 4+) by income level (<\$15,000, \$15,000-\$29,999, \$30,000-\$49,999, \$50,000-\$99,999, >\$100,000). The model results closely match the expanded survey data.
- Vehicle availability levels were compared for cross-classifications of number of workers (0, 1, 2, 3, 4+) by income level (<\$15,000, \$15,000-\$29,999, \$30,000-\$49,999, \$50,000-\$99,999, >\$100,000). The model results accurately

match the expanded survey data, considering the relatively low number of households surveyed for many of the cells.

- Vehicle availability levels were compared for cross-classifications of number of workers (0, 1, 2, 3+) by number of children (0, 1, 2+). The model results accurately match the expanded survey data, again considering the relatively low number of households surveyed for many of the cells, especially those representing households with zero vehicles.

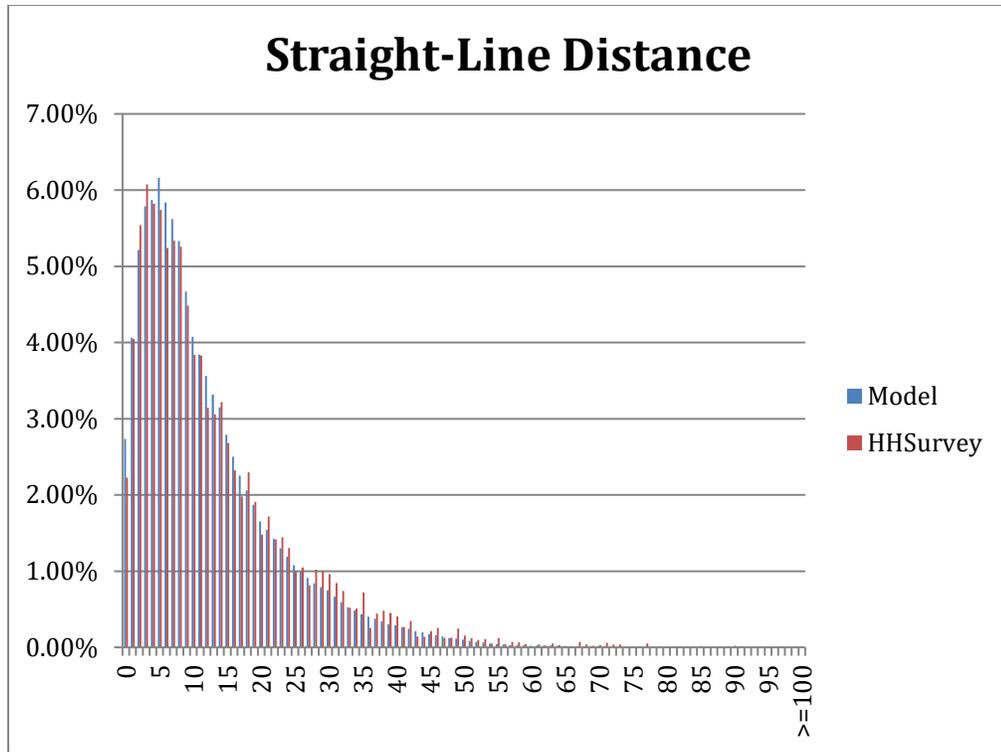
**Table 4.3 Percentage of Workers by Type with Regular Workplaces**

<b>Worker Status</b>	<u>Expanded household survey data</u>			<u>Model</u>			
	No Usual Workplace	Total	Percentage	No Usual Workplace	Total	Percentage	Diff.
Full-time	317,530	2,143,942	14.8%	175,681	2,317,432	7.6%	-7.2%
Part-time	84,267	266,015	31.7%	72,338	339,193	21.3%	-10.4%
Total	401,796	2,409,957	16.7%	248,019	2,656,625	9.3%	-7.3%

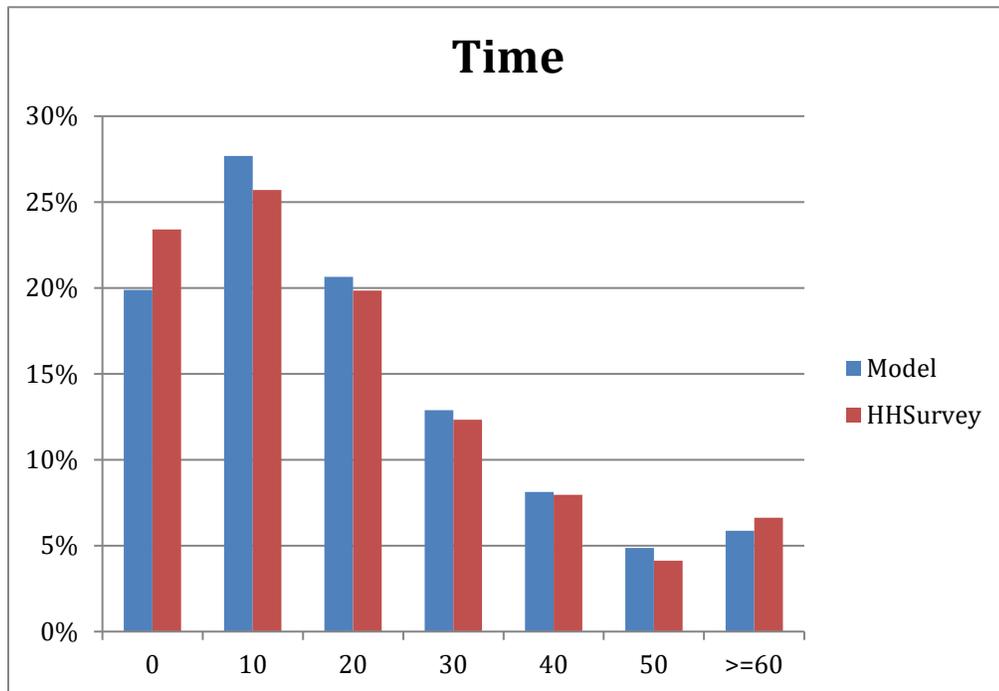
While there are some differences in the distributions, the fits are good; the coincidence ratios are 92 percent for both distance and time. The average tour times are 25.3 minutes (observed) and 25.5 minutes (modeled); the average tour distances are 13.1 miles (observed) and 12.7 miles (modeled).

The more detailed comparisons in the Excel files show the following results:

- Full time workers have longer tour lengths than part time workers. The model results were a good match to the survey results in this case.
- The distance between home and the regular workplace increases with income. In the survey data, this increase is a little steeper than in the model results.
- The distance between home and the regular workplace increases as the home location becomes less urban; the survey data trend is reflected in the model results.
- The modeled percentage of workers whose regular workplaces are in the same zone as their homes (the “intra-zonal percentage”) is 1.8 percent, compared to 2.0 percent in the survey data.
- The modeled and observed intra-zonal percentages are slightly lower for full-time workers than for part-time workers.



**Figure 4.2 Home to regular workplace tour length frequency distribution (distance).**



**Figure 4.3 Home to regular workplace tour length frequency distribution (time).**

### School Location

The school location model simulates the school location for each child in the synthetic population. The Excel file with the results of the school location model is *SchLocation.xlsx*.

While there are some differences in the distributions, the fits are good; the coincidence ratios are 82 percent for distance and 70 percent for time. The average tour times are 11.4 minutes (observed) and 11.5 minutes (modeled); the average tour distances are 6.0 miles (observed) and 6.0 miles (modeled).

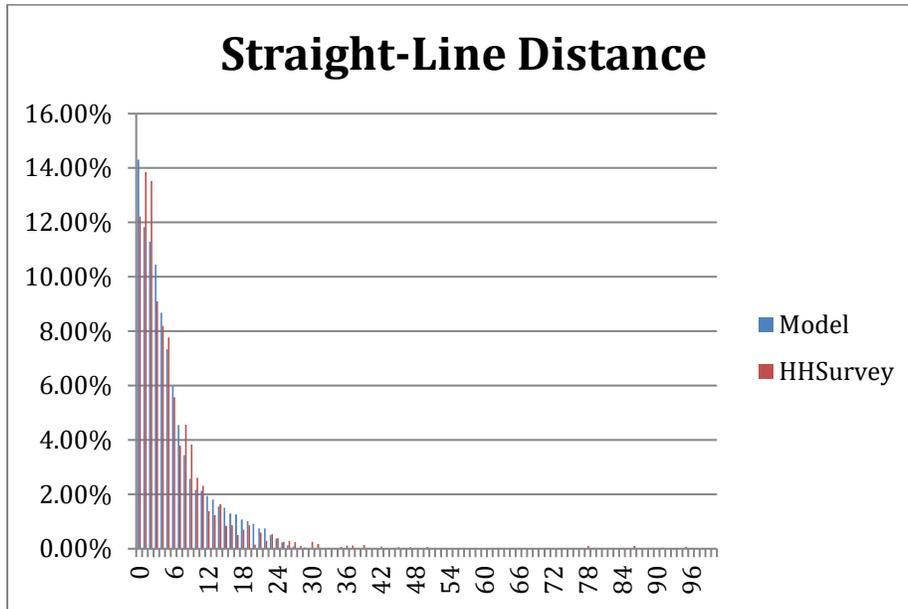
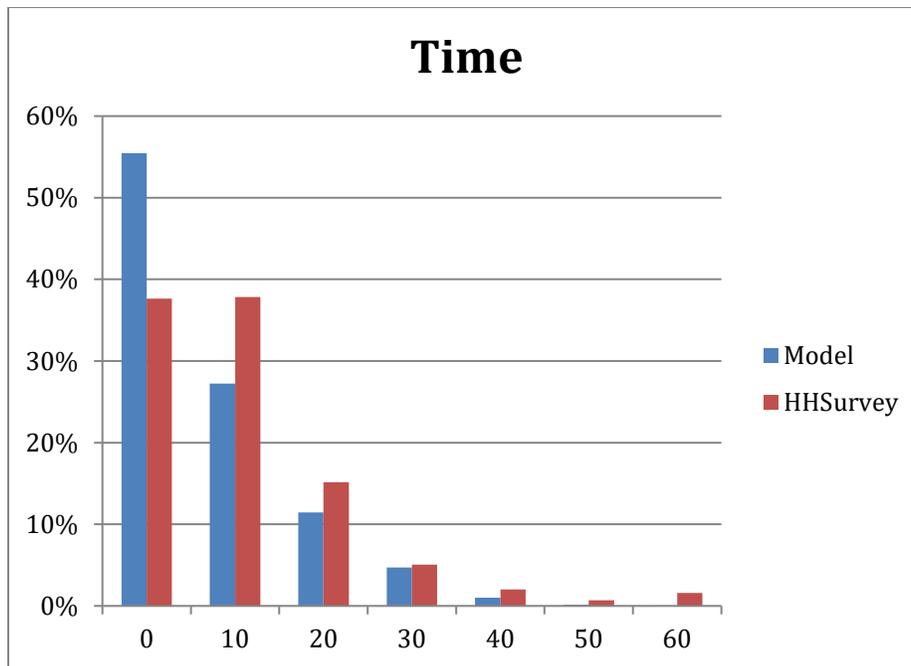


Figure 4.4 Home-to-school tour length frequency distribution (distance)



**Figure 4.5 Home-to-school tour length frequency distribution (time).**

The more detailed comparisons in the Excel files show the following results:

- Both the survey data set and the model results show that children between the ages 5 and 15 have shorter school trips than younger children, who have shorter trip lengths than students age 16 and older. This reflects that high school students often travel longer distances to school than younger children, and that pre-school children may travel farther to day care than the distance to elementary school.
- The model results show a slight increase in trip length to school as income increases. This is generally true in the survey data, although the trend is not consistent (and seems illogical).
- The highest percentage of students who go to school in their zone of residence is for students age 6 to 15, in both the survey data and model results. The overall percentage of students who attend school in their residence zone is 12 percent (survey) and 11 percent (model).

**Transit Pass Ownership**

The transit pass ownership model simulates whether each household in the synthetic population has a transit pass. The Excel file with the results of the transit pass ownership model is *Transit Pass Ownership.xlsx*.

The modeled percentage of households with transit passes matches the observed percentages well within the BMC region. The modeled percentages are low in the MWCOG region, especially in Washington, D.C.

**Table 4.4 Comparison of Modeled and Observed Transit Pass Ownership by County**

County	Survey		Survey	Model Results		Model	Difference
	Yes	No		Yes	No		
Baltimore City	24,120	228,088	10%	20,705	216,077	9%	-0.8%
Baltimore County	10,528	236,218	4%	14,516	295,706	5%	0.4%
Anne Arundel	8,717	164,813	5%	7,727	190,810	4%	-1.1%
Howard	6,456	119,521	5%	4,729	100,455	4%	-0.6%
Carroll	1,415	96,741	1%	2,021	60,317	3%	1.8%
Harford	1,515	122,941	1%	4,387	85,272	5%	3.7%
Montgomery/Prince George's/Frederick	111,908	665,605	14%	45,925	696,530	6%	-8.2%
D.C.	71,095	203,421	26%	26,310	234,828	10%	-15.8%
<b>Total</b>	<b>235,754</b>	<b>1,837,348</b>	<b>11%</b>	<b>126,320</b>	<b>1,879,995</b>	<b>6%</b>	<b>-5.1%</b>
<b>BMC region</b>	<b>52,751</b>	<b>968,322</b>	<b>5%</b>	<b>54,085</b>	<b>948,637</b>	<b>5%</b>	<b>0.2%</b>

### E-ZPass Transponder Ownership

The E-ZPass transponder ownership model simulates whether each household in the synthetic population has a transponder. The Excel file with the results of this model is *E-ZPass Ownership.xlsm*.

It should be noted that the household survey data set did not include E-ZPass transponder ownership. Therefore, observed data for comparison was obtained by BMC from the Maryland Transportation Authority (MDTA). It provided the number of transponders owned by zip code. The number of “commercial” transponders (for example, the “Standard Business Plan”) was removed from the totals prior to comparison.

The use of this alternate observed data source means that the observed data does not exactly correspond to the number of households with transponders. Most notably, the MDTA data set counts the *number of transponders*, not the *number of households with transponders*. If a household had more than one transponder, the MDTA data set would count multiple transponders. In addition, some households may own E-ZPass transponders obtained from agencies in other states. Furthermore, some commercial vehicles may have transponders that do not fall into the excluded categories, while some personal vehicles may have transponders that are counted in the commercial categories. There are also some differences between the survey period (2007-2008) and the relatively recent MDTA data set.

The model results show some noticeable differences from the observed data by county. While some model calibration was performed, the model results were not adjusted too much, given the different nature of the MDTA data set.

**Table 4.5 Comparison of Modeled and Observed E-ZPass Transponder Ownership by County**

County	Households	Observed		Households	Modeled		% Diff.
		Number of Passes	% Households with Passes		Number of Passes	% Households with Passes	
Baltimore City	252,208	66,654	26%	252,718	46,247	18%	-8%
Baltimore County	246,748	165,833	67%	318,820	151,812	48%	-20%
Anne Arundel	173,529	90,189	52%	202,188	148,250	73%	21%
Howard	125,976	45,981	36%	107,719	75,412	70%	34%
Carroll	98,156	9,816	10%	63,098	36,723	58%	48%
Harford	124,455	80,360	65%	91,762	68,410	75%	10%
Montgomery/Prince Georges/Frederick	777,512	177,348	23%	764,828	390,920	51%	28%
DC	274,517	23,852	9%	275,064	46,614	17%	8%
Total	2,073,100	660,033	32%	2,076,197	964,388	46%	15%
BMC Region	1,295,588	482,685	37%	1,311,369	573,468	44%	6%

### **4.2.1.2 Daily Activity Pattern and Related Models**

#### **Daily Activity Pattern Model**

The daily activity pattern model simulates whether each person in the synthetic population has mandatory (work, university or school) activities, has non-mandatory activities only, or makes no travel within the region (i.e., stays at home, is temporarily out of the model region, or has only external travel only between home and locations outside the model region). If a mandatory activity pattern is chosen, the number of mandatory tours (zero, one, or two) is simulated, including whether any simulated work tours have stops.

Excel files summarizes the results of the daily activity pattern model for each person type:

- Senior - *DAP\_Senior.xlsm*
- Full-time worker - *DAP\_FTW.xlsm*
- Part-time worker - *DAP\_PTW.xlsm*
- Adult (university) student - *DAP\_Adult Student.xlsm*
- Non-working adult - *DAP\_NWA.xlsm*
- Child age less than 5 - *DAP\_Child1.xlsm*
- Child age 5-15 - *DAP\_Child2.xlsm*
- Child age 16 or older - *DAP\_Child3.xlsm*

The Excel files show the results segmented by various variables of interest, including county of residence, household size, income level, vehicle availability and gender. These comparisons show only minor differences between the survey data and the model results (though in many cases, the large number of alternatives in the daily activity pattern model means that the survey data has few observations for several of the alternatives of many of the market segments).

**Table 4.6 Regional Comparison of Daily Activity Patterns**  
*Fulltime Worker*

<b>Daily Activity Pattern Type</b>	<b>Expanded household <u>survey data</u></b>		<b><u>Model Results</u></b>	
	<b>Count</b>	<b>Percentage</b>	<b>Count</b>	<b>Percentage</b>
1 Work Tour, No Stops	986,563	42.6%	955,922	41.2%
1 Work Tour, with Stops	729,477	31.5%	783,134	33.8%
2 Work Tours, No Stops	47,491	2.0%	36,614	1.6%
2 Work Tours, Stops on One	36,882	1.6%	30,623	1.3%
2 Work Tours, Stops on Both	10,098	0.4%	9,168	0.4%
1 Univ. Tour/1 Work Tour, No Stops	5,966	0.3%	4,599	0.2%
1 Univ. Tour/1 Work Tour, Stops on Work Tour	3,892	0.2%	3,355	0.1%
1 School Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%

1 School Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 Univ. Tour	10,445	0.5%	10,477	0.5%
2 Univ. Tours	0	0.0%	0	0.0%
1 School Tour	0	0.0%	0	0.0%
2 School Tours	0	0.0%	0	0.0%
Non-mandatory Travel Only	244,707	10.6%	267,483	11.5%
Stay at Home/Out of Area/ External Travel Only	241,912	10.4%	216,057	9.3%
<b>Total</b>	<b>2,317,432</b>	<b>100%</b>	<b>2,317,432</b>	<b>100%</b>

**Table 4.7 Regional Comparison of Daily Activity Patterns**  
*Part-time Worker*

Daily Activity Pattern Type	Expanded household <u>survey data</u>		<u>Model Results</u>	
	Count	Percentage	Count	Percentage
1 Work Tour, No Stops	72,243	21.3%	70,730	20.9%
1 Work Tour, with Stops	62,267	18.4%	69,263	20.4%
2 Work Tours, No Stops	3,353	1.0%	2,187	0.6%
2 Work Tours, Stops on One	4,477	1.3%	3,047	0.9%
2 Work Tours, Stops on Both	2,081	0.6%	1,532	0.5%
1 Univ. Tour/1 Work Tour, No Stops	0	0.0%	1,724	0.5%
1 Univ. Tour/1 Work Tour, Stops on Work Tour	108	0.0%	99	0.0%
1 School Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 Univ. Tour	277	0.1%	405	0.1%
2 Univ. Tours	0	0.0%	0	0.0%
1 School Tour	0	0.0%	0	0.0%
2 School Tours	0	0.0%	0	0.0%
Non-mandatory Travel Only	136,425	40.2%	138,126	40.7%
Stay at Home/Out of Area/ External Travel Only	57,963	17.1%	52,080	15.4%
<b>Total</b>	<b>339,193</b>	<b>100%</b>	<b>339,193</b>	<b>100%</b>

**Table 4.8 Regional Comparison of Daily Activity Patterns**  
*Adult Student*

Daily Activity Pattern Type	Expanded household <u>survey data</u>		<u>Model Results</u>	
	Count	Percentage	Count	Percentage
1 Work Tour, No Stops	19,694	7.6%	19,331	7.4%
1 Work Tour, with Stops	19,338	7.4%	19,091	7.3%

2 Work Tours, No Stops		0.0%	0	0.0%
2 Work Tours, Stops on One		0.0%	0	0.0%
2 Work Tours, Stops on Both	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, No Stops	4,111	1.6%	3,989	1.5%
1 Univ. Tour/1 Work Tour, Stops on Work Tour	2,076	0.8%	2,064	0.8%
1 School Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 Univ. Tour	92,835	35.7%	91,869	35.2%
2 Univ. Tours	4,127	1.6%	4,131	1.6%
1 School Tour	0	0.0%	0	0.0%
2 School Tours	0	0.0%	0	0.0%
Non-mandatory Travel Only	75,908	29.2%	76,026	29.1%
Stay at Home/Out of Area/ External Travel Only	41,699	16.1%	44,832	17.2%
<b>Total</b>	<b>259,788</b>	<b>100%</b>	<b>261,333</b>	<b>100%</b>

**Table 4.9. Regional Comparison of Daily Activity Patterns**  
*Senior*

Daily Activity Pattern Type	Expanded household <u>survey data</u>		<u>Model Results</u>	
	Count	Percentage	Count	Percentage
1 Work Tour, No Stops	8,322	1.8%	8,452	1.8%
1 Work Tour, with Stops	6,859	1.5%	8,067	1.8%
2 Work Tours, No Stops	838	0.2%	0	0.0%
2 Work Tours, Stops on One	570	0.1%	0	0.0%
2 Work Tours, Stops on Both	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 Univ. Tour	1,775	0.4%	791	0.2%
2 Univ. Tours	0	0.0%	0	0.0%
1 School Tour	0	0.0%	0	0.0%
2 School Tours	0	0.0%	0	0.0%
Non-mandatory Travel Only	272,942	59.7%	263,859	57.7%
Stay at Home/Out of Area/ External Travel Only	165,985	36.3%	176,122	38.5%
<b>Total</b>	<b>457,291</b>	<b>100%</b>	<b>457,291</b>	<b>100%</b>

**Table 4.10 Regional Comparison of Daily Activity Patterns**  
*Non-working Adult*

Daily Activity Pattern Type	Expanded household survey data		Model Results	
	Count	Percentage	Count	Percentage
1 Work Tour, No Stops	393	0.1%	467	0.1%
1 Work Tour, with Stops	357	0.1%	472	0.1%
2 Work Tours, No Stops	0	0.0%	0	0.0%
2 Work Tours, Stops on One	0	0.0%	0	0.0%
2 Work Tours, Stops on Both	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 Univ. Tour	587	0.1%	369	0.1%
2 Univ. Tours	0	0.0%	0	0.0%
1 School Tour	0	0.0%	0	0.0%
2 School Tours	0	0.0%	0	0.0%
Non-mandatory Travel Only	398,552	68.4%	397,174	68.2%
Stay at Home/Out of Area/ External Travel Only	182,503	31.3%	183,910	31.6%
<b>Total</b>	<b>582,392</b>	<b>100%</b>	<b>582,392</b>	<b>100%</b>

**Table 4.11 Regional Comparison of Daily Activity Patterns**  
*Child Age Less than 5*

Daily Activity Pattern Type	Expanded household survey data		Model Results	
	Count	Percentage	Count	Percentage
1 Work Tour, No Stops	0	0.0%	0	0.0%
1 Work Tour, with Stops	0	0.0%	0	0.0%
2 Work Tours, No Stops	0	0.0%	0	0.0%
2 Work Tours, Stops on One	0	0.0%	0	0.0%
2 Work Tours, Stops on Both	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 Univ. Tour	0	0.0%	0	0.0%
2 Univ. Tours	0	0.0%	0	0.0%
1 School Tour	125,930	37.3%	133,788	39.6%

2 School Tours	456	0.1%	0	0.0%
Non-mandatory Travel Only	126,069	37.3%	132,190	39.1%
Stay at Home/Out of Area/ External Travel Only	85,476	25.3%	71,953	21.3%
<b>Total</b>	<b>337,931</b>	<b>100%</b>	<b>337,931</b>	<b>100%</b>

**Table 4.12 Regional Comparison of Daily Activity Patterns**  
*Child Age 5-15*

Daily Activity Pattern Type	Expanded household <u>survey data</u>		<u>Model Results</u>	
	Count	Percentage	Count	Percentage
1 Work Tour, No Stops	365	0.0%	0	0.0%
1 Work Tour, with Stops	167	0.0%	0	0.0%
2 Work Tours, No Stops	0	0.0%	0	0.0%
2 Work Tours, Stops on One	0	0.0%	0	0.0%
2 Work Tours, Stops on Both	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, No Stops	701	0.1%	0	0.0%
1 School Tour/1 Work Tour, Stops on Work Tour	457	0.1%	0	0.0%
1 Univ. Tour	0	0.0%	0	0.0%
2 Univ. Tours	0	0.0%	0	0.0%
1 School Tour	513,502	70.1%	531,673	72.6%
2 School Tours	10,047	1.4%	10,717	1.5%
Non-mandatory Travel Only	125,639	17.2%	117,309	16.0%
Stay at Home/Out of Area/ External Travel Only	81,658	11.1%	72,837	9.9%
<b>Total</b>	<b>732,536</b>	<b>100%</b>	<b>732,536</b>	<b>100%</b>

**Table 4.13 Regional Comparison of Daily Activity Patterns**  
*Child Age 16-17*

Daily Activity Pattern Type	Expanded household <u>survey data</u>		<u>Model Results</u>	
	Count	Percentage	Count	Percentage
1 Work Tour, No Stops	4,783	3.6%	5,042	3.8%
1 Work Tour, with Stops	1,123	0.8%	1,654	1.2%
2 Work Tours, No Stops	0	0.0%	0	0.0%
2 Work Tours, Stops on One	0	0.0%	0	0.0%
2 Work Tours, Stops on Both	0	0.0%	0	0.0%
1 Univ. Tour/1 Work Tour, No Stops	0	0.0%	0	0.0%

1 Univ. Tour/1 Work Tour, Stops on Work Tour	0	0.0%	0	0.0%
1 School Tour/1 Work Tour, No Stops	4,699	3.5%	5,789	4.3%
1 School Tour/1 Work Tour, Stops on Work Tour	924	0.7%	934	0.7%
1 Univ. Tour	0	0.0%	0	0.0%
2 Univ. Tours	0	0.0%	0	0.0%
1 School Tour	87,307	64.9%	86,685	64.5%
2 School Tours	3,080	2.3%	2,601	1.9%
Non-mandatory Travel Only	17,482	13.0%	17,152	12.8%
Stay at Home/Out of Area/ External Travel Only	15,052	11.2%	14,592	10.9%
<b>Total</b>	<b>134,449</b>	<b>100%</b>	<b>134,449</b>	<b>100%</b>

### School Escorting Model

For each child traveling to school, the school escorting model determines whether he or she is escorted by another household member to or from school, and, if so, which household member does the escorting, and whether that household member escorts the student as part of a mandatory tour (for example, on the way to or from work). The Excel file that summarizes the results of the school escorting model is *SchoolEscort.xlsm*.

The five alternatives for each student are:

- Outbound mandatory – Escorting to school as part of a mandatory tour
- Outbound standalone – Escorting to school as part of a standalone tour
- Return mandatory – Escorting from school as part of a mandatory tour
- Return standalone – Escorting from school as part of a standalone tour
- None – Student is not escorted.

The more detailed comparisons in the Excel files show the following results:

- The survey data show that very little school escorting occurs in zero car households. The model results reflect this unsurprising result.
- Both the survey data and model results show that most escorts are full-time workers or non-working adults.
- The household survey data show that 72 percent of escorts are female; the model results show a lower percentage of female escorts (62 percent).
- Generally, fewer children from higher income households are escorted, especially for the youngest children.

**Table 4.14 Regional Comparison of School Escorting Alternatives**

Escort Type	Child Age	Expanded household survey data		Model Results		Percentage Point Difference (Model - Survey)
		Count	Percentage	Count	Percentage	
Outbound Mandatory	< 5 Years	53,716	20.9%	54,958	20.5%	-0.4%
Outbound Standalone	< 5 Years	39,172	15.2%	38,501	14.4%	-0.9%
Return Mandatory	< 5 Years	49,848	19.4%	51,756	19.3%	-0.1%
Return Standalone	< 5 Years	39,941	15.5%	41,492	15.5%	0.0%
None	< 5 Years	74,263	28.9%	80,869	30.2%	1.3%
<b>Total</b>	< 5 Years	<b>256,940</b>		<b>267,576</b>		
Outbound Mandatory	5-15 Years	78,304	7.3%	115,881	10.5%	3.2%
Outbound Standalone	5-15 Years	114,090	10.6%	84,261	7.6%	-3.0%
Return Mandatory	5-15 Years	48,057	4.5%	78,485	7.1%	2.6%
Return Standalone	5-15 Years	99,897	9.3%	75,770	6.8%	-2.4%
None	5-15 Years	734,989	68.3%	751,817	68.0%	-0.4%
<b>Total</b>	5-15 Years	<b>1,075,337</b>		<b>1,106,214</b>		
Outbound Mandatory	16+ Years	8,528	4.3%	6,736	3.4%	-0.9%
Outbound Standalone	16+ Years	11,927	6.0%	21,528	10.9%	4.9%
Return Mandatory	16+ Years	3,840	1.9%	5,327	2.7%	0.8%
Return Standalone	16+ Years	6,166	3.1%	3,982	2.0%	-1.1%
None	16+ Years	167,920	84.6%	159,647	80.9%	-3.7%
<b>Total</b>	16+ Years	<b>198,381</b>		<b>197,220</b>		
Outbound Mandatory	All	140,548	9.2%	177,575	11.3%	2.1%
Outbound Standalone	All	165,189	10.8%	144,290	9.2%	-1.6%
Return Mandatory	All	101,746	6.6%	135,568	8.6%	2.0%
Return Standalone	All	146,004	9.5%	121,244	7.7%	-1.8%
None	All	977,171	63.8%	992,333	63.2%	-0.7%
<b>Total</b>	<b>All</b>	<b>1,530,658</b>		<b>1,571,010</b>		

**Table 4.15 Regional Comparison of School Escorting Types – Household Survey vs. Model Results**

Escort Person Type	Expanded Household Survey Data		Model Results		Percentage Point Difference (Model - Survey)
	Count	Percentage	Count	Percentage	
Adult Student	22,542	5.0%	6,629	1.5%	-3.6%
Full-time Worker	274,394	61.4%	308,295	68.2%	6.8%
Part-time Worker	53,548	12.0%	65,243	14.4%	2.5%
Non-working Adult	90,952	20.3%	67,654	15.0%	-5.4%
Senior	5,603	1.3%	4,347	1.0%	-0.3%
<b>Total</b>	<b>447,038</b>		<b>452,168</b>		

## Joint Travel Model

The fully joint tour models include a **generation** model, which simulates the number (zero, one, or two) and purpose (meal, shopping, personal business or social-recreation) of fully joint tours made by each household; and **participation** model, which determines which household members participate in each simulated joint tour. The Excel file that summarizes the results of the fully joint tour models is *JointTour Gen & Part.xlsm*. The household survey data set shows an average of 0.255 fully joint tours per household while the model results show 0.262 joint tours per household.

The more detailed comparisons in the Excel files show the following results:

- The survey data set shows varying rates of joint tours per household by county, with the lowest rates in Baltimore City and Washington, D.C. This is not surprising since average household size is lower in these cities than in the rest of the model region. These two municipalities also have the lowest joint tour rates in the model results, though the model somewhat overestimates the joint tour rate in Baltimore City.
- Among households making joint tours, there is no discernable pattern of the number of joint tours made by income level. The model somewhat overestimates the number of joint tours for the lowest income group.
- Among households making joint tours, zero vehicle households make fewer joint tours, though the model somewhat overestimates the joint tour rate for these households. The model also underestimates the joint tour rate for households with three or more vehicles.
- The distributions of joint tours by purpose and by party size (two, three or three-plus) are similar for the survey data set and the model results. The cross-classifications of tour purpose by party size also match well, with the largest differences appearing for the combinations with the lowest incidence in the survey data set (generally, the three-plus person tours).

## Individual Non-Mandatory Tour Generation Model

The individual non-mandatory tour generation model simulates the number (zero, one, two, or three) and purpose (meal, shopping, personal business, escorting or social-recreation) of non-mandatory tours made by each person in the synthetic population for whom a mandatory or non-mandatory daily activity pattern has been simulated. (At least one non-mandatory tour must be simulated for persons with non-mandatory patterns). The Excel file that summarizes the results of the non-mandatory tour generation model is *INMTourGeneration.xlsm*. The number of modeled, non-mandatory tours per person is slightly lower (about four percent) than the number of such tours in the expanded household survey data set.

The more detailed comparisons in the Excel files show the following results:

- The modeled percentages of non-mandatory tours by purpose match the percentages from the survey data set almost exactly.
- Compared to the survey data set, the model slightly overestimates the number of non-mandatory tours for seniors and slightly underestimates for children,

adult students and adult non-workers. The model percentages of non-mandatory tours by purpose for each person type also match the percentages from the survey data set almost exactly.

- Most of the underestimation of non-mandatory tours by the model is for males. The expanded show very similar percentages of tours by purpose for males and females. The slight differences between genders in the data (for example, slightly higher percentages of shopping and escort tours for females) are reflected in the model.
- There are few relatively minor differences in the model results by income level, household size and vehicle availability, compared to the survey data set. Perhaps most notable is that the model does not pick up the differences by household size for individual meal tours—the survey data show that the larger the household, the lower the incidence of individual meal tours.

### **Work-based Subtour Generation Model**

The work-based subtour generation model simulates the number (zero, one, or two) and purpose (work, meal, shopping, personal business, escorting or social-recreation) of work-based subtours made by persons making work tours. The Excel file that summarizes the results of the work-based subtour generation model is *WBTourGeneration.xlsx*. The number of modeled work-based subtours per work tour is about the same as the number of such subtours in the expanded household survey data set (0.154 observed versus 0.155 modeled).

The more detailed comparisons in the Excel files show the following results:

- The modeled percentages of work-based subtours by purpose closely match the percentages from the survey data set.
- The model results show that those who take non-auto modes to work make fewer subtours than those who take auto modes, with those who walk or bike to work making fewer subtours than those who take transit. This seems logical, in that individuals taking non-auto modes typically lack access to cars for making work-based subtours. However, the survey data set generally shows the opposite pattern. The model results slightly overestimate the number of subtours by workers who use auto modes to work and slightly underestimate the number of subtours by workers who use non-auto modes to travel to work.
- The survey data set shows that males make more work-based subtours than females, and the model results match this result.
- The household survey data show that the rate of making work-based subtours increases with income level. The model data show this pattern for all travelers, but at a much more moderate rate of increase.

#### **4.2.1.3 Tour Level Choice Models**

##### **Tour Destination Choice Models**

The tour destination choice models simulate the location of the primary activity of each tour. There are Excel files with detailed results for various aggregate activity purposes:

- Work (not to regular workplace) – *Tour Dest Work.xlsx*
- University - *Tour Dest Uni.xlsx*
- Fully joint – *Tour Dest Joint.xlsx*
- Individual non-mandatory (except escort tours) - *Tour Dest INM.xlsx*
- Work-based subtours – *WB Tour Dest.xlsx*

Each spreadsheet file includes histograms comparing the tour length frequency distributions, by both time and distance, for the corresponding activity purpose.

**Table 4.16 Coincidence Ratios for Tour Length Frequency Distributions**

<b>Tour Purpose</b>	<b>Coincidence Ratio</b>	
	<b>Time</b>	<b>Distance</b>
Work (including tours to regular workplace)	86%	80%
University	91%	76%
Joint	90%	85%
Individual Non-mandatory	89%	89%
Work-based Subtours	98%	77%

For each tour purpose, the following comparisons between the observed (expanded household survey) data and model results are included in the Excel files:

- Average tour length (time and distance) by:
  - Tour activity (meal, shop, personal business or social-recreation) - *for non-mandatory tours only*
  - Income level
  - Area type at home (or workplace for work-based subtours) and at the primary activity location
  - Person type – *except joint tours*
  - Number of household vehicles – *except work-based subtours*
  - Parent tour mode – *work-based subtours only*
- Percentage of intrazonal tours (primary activity location zone is the same as the home zone, or work zone for work-based subtours) by:
  - Area type

- Person type – *work tours only*
- Number of household vehicles – *except work-based subtours*
- Parent tour mode – *work-based subtours only*

In the model results, the average tour lengths show a logical progression with tour lengths increasing as income increases; the survey data do not show this pattern for all tour purposes. Another difference is that the model shows more intrazonal tours for less dense areas while the survey data do not show this pattern.

**Table 4.17 Average Trip Length Comparisons by Tour Purpose**

<b>Tour Purpose</b>	<b><u>Observed</u></b>		<b><u>Model</u></b>	
	<b>Time (min)</b>	<b>Distance (miles)</b>	<b>Time (min)</b>	<b>Distance (miles)</b>
Work	23.0	12.0	23.1	11.3
University	17.2	9.0	17.2	8.3
Joint	12.8	7.1	14.4	7.2
Individual Non-mandatory	11.0	5.7	11.9	5.6
Work-based Subtours	5.7	3.3	6.0	2.5

**Table 4.18 Comparison of Intrazonal Percentages**

<b>Tour Purpose</b>	<b>Time</b>	<b>Distance</b>
Work (including tours to regular workplace)	2%	3%
University	3%	4%
Joint	7%	8%
Individual Non-mandatory	11%	13%
Work-based Subtours	18%	18%

### **Tour Time-of-Day Choice Models**

The tour time-of-day choice models simulate the start and end times, in half-hour increments, of the primary activity of each tour. There are Excel files with detailed results for various aggregate activity purposes:

- Mandatory (work, school and university) - *TOD\_Mand.xlsm*
- Joint - *TOD\_Joint.xlsm*
- Individual non-mandatory - *TOD\_NM.xlsm*
- Work-based subtours - *TOD\_WB.xlsm*

Each spreadsheet presents histograms comparing the distributions of activity arrival and departures for the corresponding activity purpose.

**Table 4.19 Coincidence Ratios for Time-of-Day Distributions**

<b>Tour Purpose</b>	<b><u>Coincidence Ratio</u></b>	
	<b>Arrival</b>	<b>Departure</b>
Work	85%	83%
School	81%	86%
University	61%	51%
Joint	83%	78%
Individual Non-mandatory	83%	83%
Work-based Subtours	90%	91%

**Table 4.20 Modeled and Observed Activity Durations by Purpose**

<b>Tour Purpose</b>	<b><u>Duration (hours)</u></b>	
	<b>Survey</b>	<b>Model</b>
Work	7.1	7.4
School	7.0	7.2
University	4.5	5.0
Joint	1.9	1.7
Meal	0.9	1.5
Shopping	0.9	1.0
Personal Business	1.8	1.4

Social-recreation	1.0	1.8
Escort	0.1	0.2
Work-based Subtours	0.8	0.7

Each spreadsheet also compares the average activity duration in hours by the following segmentations:

- Income level
- Person type – *except joint tours*
- Gender – *except joint tours*
- Specific activity purpose (e.g., meal, shopping) – *joint, non-mandatory, and work-based*

In most cases, the modeled and survey activity durations are within 10 percent or within 10 minutes of one another, when there are sufficient observations in the segment. Some exceptions include the following:

- Modeled activity durations for work tours are about 15 percent high for part time workers and about 25 percent high for adult students. Modeled activity durations for work tours are about 15 percent high for the \$15,000 to \$30,000 income group and about 10 percent high for the \$30,000 to \$50,000 income group.
- Modeled activity durations for non-mandatory tours are high for most segments, since the average activity duration is high by about 15 minutes.
- Several of the university tour segments have greater differences between the observed data and model results; this is due to the relatively low number of these tours.

### **Tour Mode Choice Models**

The tour mode choice models simulate the main mode of each tour. -There are Excel files with detailed validation results for various aggregate activity purposes:

- Work - *TourModeChoice\_Work.xlsx*
- School - *TourModeChoice\_Sch.xlsx*
- University - *TourModeChoice\_Uni.xlsx*
- Joint - *TourModeChoice\_Joint.xlsx*
- Individual non-mandatory (except escort tours) - *TourModeChoice\_INM.xlsx*
- Escort - *TourModeChoice\_Escort.xlsx*
- Work-based subtours - *TourModeChoice\_WB.xlsx*

**Table 4.21 Regional Modeled and Observed Tour Mode Shares by Purpose**

Tour Mode	<u>Work</u>		<u>School</u>		<u>University</u>		<u>Individual Non- mandatory</u>	
	Survey	Model	Survey	Model	Survey	Model	Survey	Model
Drive Alone	59.0%	58.6%	2.0%	1.9%	50.9%	49.5%	46.9%	47.4%
Shared Ride 2	14.4%	14.9%	19.5%	7.1%	11.6%	11.5%	24.3%	23.5%
Shared Ride 3+	7.7%	8.2%	38.0%	53.3%	11.4%	11.3%	12.2%	11.7%
Transit-Walk Access	9.1%	8.6%	4.5%	4.4%	13.3%	14.3%	6.0%	5.8%
Transit-Auto Access	6.6%	6.4%	0.3%	0.3%	3.9%	3.9%	0.9%	0.9%
Walk	2.3%	2.5%	7.7%	6.3%	6.3%	7.6%	8.9%	10.1%
Bike	0.8%	0.8%	0.4%	0.4%	1.7%	1.9%	0.6%	0.6%
School Bus			27.7%	26.3%			0.2%	

Tour Mode	<u>Escort</u>		<u>Joint</u>		<u>Work-based Subtours</u>		<u>ALL TRIPS</u>	
	Survey	Model	Survey	Model	Survey	Model	Survey	Model
Drive Alone					44.5%	46.1%	38.9%	38.4%
Shared Ride 2	45.7%	46.9%	46.7%	47.2%	12.9%	12.7%	22.4%	20.9%
Shared Ride 3+	42.7%	40.5%	43.0%	43.5%	7.2%	7.0%	17.9%	20.1%
Transit-Walk Access			2.9%	2.6%	2.0%	2.2%	6.1%	5.9%
Transit-Auto Access			0.4%	0.0%	0.2%	0.0%	2.8%	2.7%
Walk	11.7%	12.5%	6.7%	6.4%	32.8%	31.7%	7.8%	8.0%
Bike			0.3%	0.3%	0.2%		0.6%	0.6%
School Bus					0.2%		3.5%	3.5%

For each tour purpose, the spreadsheet files show the following comparisons between the survey and modeled tour mode shares:

- Area type at home (or workplace for work-based subtours) and at the primary activity location
- Distance range
- Transit in-vehicle time ranges (walk and auto access)
- Household size and income level
- Vehicles less than, equal to, or greater than number of workers/drivers
- Age and gender

In nearly all cases, the mode shares from the model closely matched those from the survey data. Some of the key results, which are true for both the observed and model data, include the following:

- Not surprisingly, transit and non-motorized mode shares increase as the area becomes more densely developed, while auto mode shares decrease. This trend noted in the expanded survey data is also seen in the model, although the rate of changes among area types is more moderate. (It should be noted that except for work and individual non-mandatory tours, the number of survey observations is fairly small for the most urban area types.)
- Transit-walk access mode shares decrease with increased distance; the opposite holds for transit-auto access shares (nearly all transit-auto access tours are for work or university purposes). The model captures these trends better for the walk access tours. Non-motorized trips, naturally, are nearly all short distance, and the model results reflect this.
- Transit mode shares to all counties are low—from zero to two percent—for all tour purposes, with the exception of the three Maryland counties in the MWCOC region, where the transit shares are a bit higher. Transit shares to the cities of Baltimore and Washington are substantially higher. The model results reflect these trends.
- Transit-walk access shares decrease with increasing income levels for all tour purposes, and the model results accurately reflect this trend. For work tours, transit-auto access shares increase with increasing income levels, and the model results accurately reflect this as well.
- Generally, transit shares decrease with increasing household size, and the model accurately reflects this trend.
- Not surprisingly, transit and non-motorized mode shares are much higher in households with fewer vehicles than workers, or fewer vehicles than drivers, and are even higher in households with zero vehicles. The model reflects these trends accurately.
- Auto shares, especially drive alone, decrease while transit shares decrease with increasing age.
- For work tours, transit and shared ride mode shares are higher for females; drive-alone and bike mode shares are higher for males.

#### **4.2.1.4 Stop Generation Models**

The stop generation models simulate the number and purpose of stops made on each tour. Separate models were estimated for each tour purpose. There are Excel files with detailed results for various aggregate activity purposes:

- Mandatory (work, school and university) - *Stops\_Mand.xlsx*
- Joint - *Stops\_Joint.xlsx*
- Individual non-mandatory - *Stops\_INM.xlsx*
- Work-based subtours - *Stops\_WB.xlsx*

As these tables show, the model results are close to the observed results from the expanded household survey data set.

**Table 4.22 Regional Modeled and Observed Shares of Half Tours by Number of Stops by Purpose**

Stops	<u>Work</u>		<u>School/ University</u>		<u>Individual Non- mandatory</u>	
	Survey	Model	Survey	Model	Survey	Model
0 stops	50%	48%	78%	80%	55%	60%
1 stop	35%	37%	15%	13%	30%	27%
2 stops	10%	9%	5%	5%	9%	8%
3 stops	6%	7%	2%	2%	6%	5%

Stops	<u>Joint</u>		<u>Work-based Subtours</u>		<u>All Half Tours</u>	
	Survey	Model	Survey	Model	Survey	Model
0 stops	51%	58%	80%	86%	57%	60%
1 stop	34%	30%	17%	13%	29%	28%
2 stops	10%	9%	3%	2%	9%	8%
3 stops	4%	3%	0%	0%	5%	5%

\*\*Note: Model is constrained to produce only 1 or 2 stops per half tour on work based subtours.

**Table 4.23 Observed and Modeled Average Number of Stops per Half Tour**

Tour Purpose	<u>Outbound Half Tour</u>		<u>Return Half Tour</u>	
	Survey	Model	Survey	Model
Work	0.24	0.28	0.47	0.46
School/University	0.08	0.08	0.24	0.22
Individual Non-mandatory	0.30	0.26	0.37	0.32
Joint	0.32	0.24	0.36	0.34
Work-based Subtours	0.10	0.07	0.17	0.13
<b>Total – All Tours</b>	<b>0.21</b>	<b>0.25</b>	<b>0.33</b>	<b>0.34</b>

For each tour purpose, the spreadsheet files show the following comparisons between survey and modeled stops:

- Income level
- Person type – *except joint tours*
- Age and gender – *except joint tours*

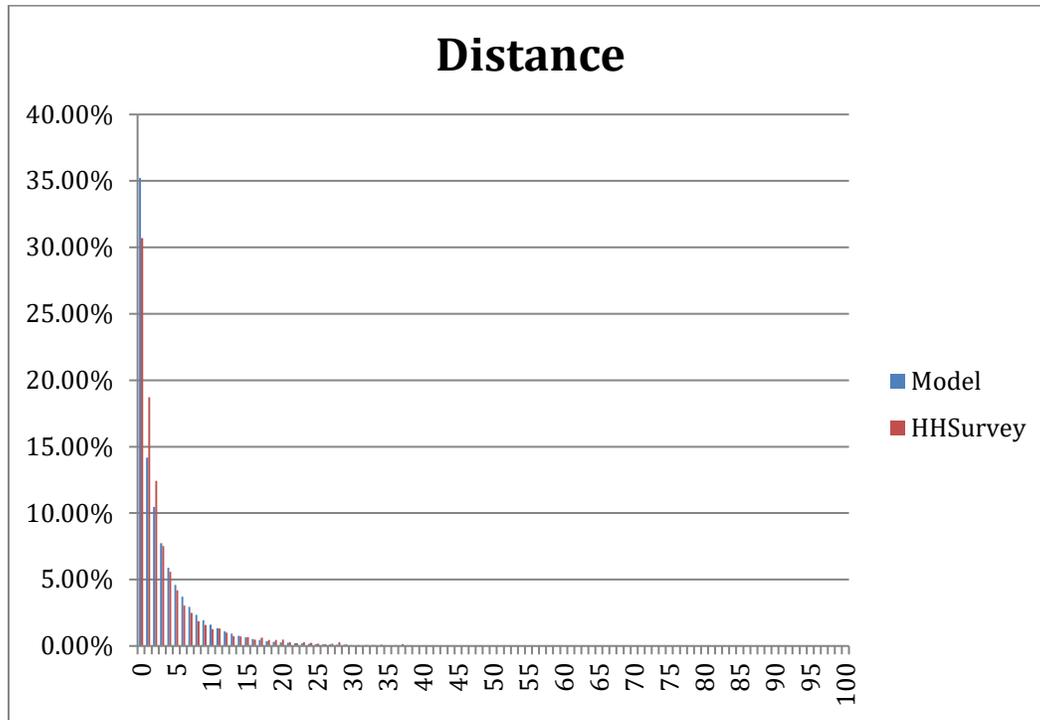
The more detailed comparisons in the Excel files show some differences between the model results and the expanded survey data (some due to small sample sizes for certain segments), but overall the model results reflect the observed data fairly accurately.

#### **4.2.1.5 Stop/Trip Level Choice Models**

##### **Stop Destination Choice Models**

The stop destination choice model simulates the locations of all intermediate stops between the home (or workplace, for work-based subtours) and primary activity location on tours. The Excel file that summarizes the results of this model is *StopDestChoice.xlsm*.

While there are some differences in the distributions, the fits are good; the coincidence ratio is 85 percent. The average trip distances are 4.1 miles (observed) and 4.4 miles (modeled).



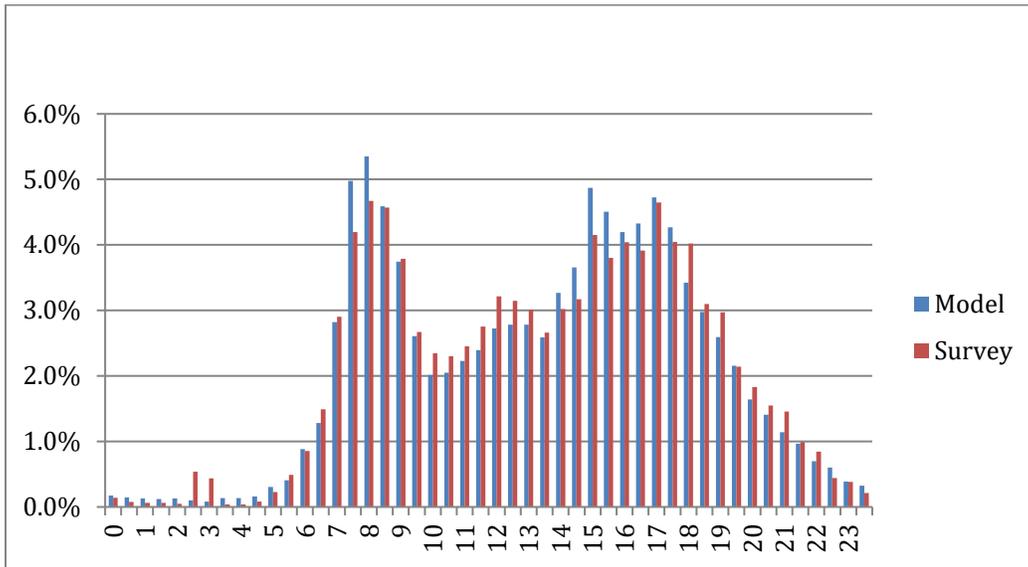
**Figure 4.6 Trip length frequency distribution for stops (miles)**

The spreadsheet file also provides comparisons between the observed data from the household survey and the model results for the average trip distances segmented by stop purpose, household income level, tour mode, area type at home and primary destination, and tour purpose. These comparisons show a good match, with 36 of the 54 segments having modeled trip lengths within one half mile of the observed trip length and 47 of the segments having modeled trip lengths within one mile of observed. The largest overestimates by the model are for two of the midrange area types and for school stops, while the largest underestimates are for rural area type and for work and university stops.

##### **Stop Time of Day Choice Models**

The stop time of day choice model simulates the times (at the half-hour level) of all intermediate stops between the home (or workplace, for work-based subtours) and primary activity location on

tours. The Excel file that summarizes the results of this model (as well as the trip mode choice model, described in the next section) is *TripModeTODChoice.xlsx*.



**Figure 4.7 Trip length frequency distribution for stops (miles)**

**Trip Mode Choice Model**

The trip mode choice model simulates the mode for each trip that is part of a tour, conditional on the simulated tour mode. The Excel file *TripModeTODChoice.xlsx* also summarizes the results of the trip mode choice model. The following is a summary of trip mode shares by tour mode:

**Drive alone:**

- Survey: drive alone – 99%, walk 1%
- Model: drive alone – 98%, walk 2%

**Tour mode shared ride 2:**

- Survey: drive alone – 30%, shared ride 2 – 68%, walk 2%
- Model: drive alone – 31%, shared ride 2 – 67%, walk 2%

**Tour mode shared ride 3+:**

- Survey: drive alone – 18%, shared ride 2 – 18%, shared ride 3+ – 62%, walk – 2%
- Model: drive alone – 15%, shared ride 2 – 21%, shared ride 3+ – 63%, walk – 1%

**Tour mode transit-walk access:**

- Survey: drive alone – 1%, shared ride 2 – 5%, shared ride 3+ – 3%, transit-walk access – 68%, walk – 23%, bike – 1%
- Model: drive alone – 6%, shared ride 2 – 8%, shared ride 3+ – 4%, transit-walk access – 51%, walk – 29%, bike – 3%

**Tour mode transit-auto access:**

- Survey: drive alone – 9%, shared ride 2 – 9%, shared ride 3+ – 5%, transit-walk access – 5%, transit-auto access – 65%, walk – 7%
- Model: drive alone – 6%, shared ride 2 – 29%, shared ride 3+ – 18%, transit-walk access – 2%, transit-auto access – 39%, walk – 6%

**Tour mode school bus:**

- Survey: dhared ride 2 – 4%, shared ride 3+ – 6%, walk – 2%, school bus – 88%
- Model: dhared ride 2 – 5%, shared ride 3+ – 8%, walk – 3%, school bus – 84%

Note that all trips on walk and bicycle tours have the same trip mode as the tour mode.

The spreadsheet file also shows the following comparisons between survey and modeled trip mode shares:

- Area type at home (or workplace for work-based subtrips) and at the primary activity location
- Distance range
- Transit in-vehicle time ranges (walk access)
- Household size and income level
- Vehicles less than, equal to, or greater than number of workers/drivers
- Age and gender

In most cases, the mode shares from the model accurately matched those from the survey data.

#### **4.2.2 Highway and Transit Assignment**

Since the highway and transit assignment processes are essentially the same static, aggregate process used in BMC's previous (trip-based) model, the checks are like those performed for the validation of the previous model. They consist mainly of comparisons of model results to observed data, including traffic and transit ridership counts. Highway assignment checks include:

- Volume/vehicle-miles traveled (VMT) by facility type
- Volume/VMT by area type

- Volume/VMT by county
- Volume/VMT by volume level
- Volume/VMT by time of day
- Volume/count ratio on key routes
- Sum of volumes on screenlines/cutlines

Transit assignment checks include:

- Boardings by service category (Metrobus local, Metrobus park-and-ride, Metrorail)
- Boardings by service category and geographic orientation, defined as follows:
  - Local-radial
  - Local-crosstown
  - Local-circulator
  - Local-limited
  - Local-shuttle
  - Park-and-Ride-CBD
  - Park-and-Ride-secondary
  - Metrorail
- Boardings per linked trip (transfer rate)
- Boardings by route
- Boardings by Metrorail station

### **4.2.3 Sensitivity Testing**

One goal of activity-based models is an increased sensitivity to model inputs. Sensitivity testing involves adjusting key factors in the model and observing the effects on forecasted travel. These adjustments can be made to model parameter values (e.g., the mode choice cost coefficient) and to model inputs (e.g., land use variables, socioeconomic conditions, fuel costs, etc.).

The following sensitivity tests were performed:

- Aging population showing more retirees
- Brownfield development
- Time of day switching due to congestion

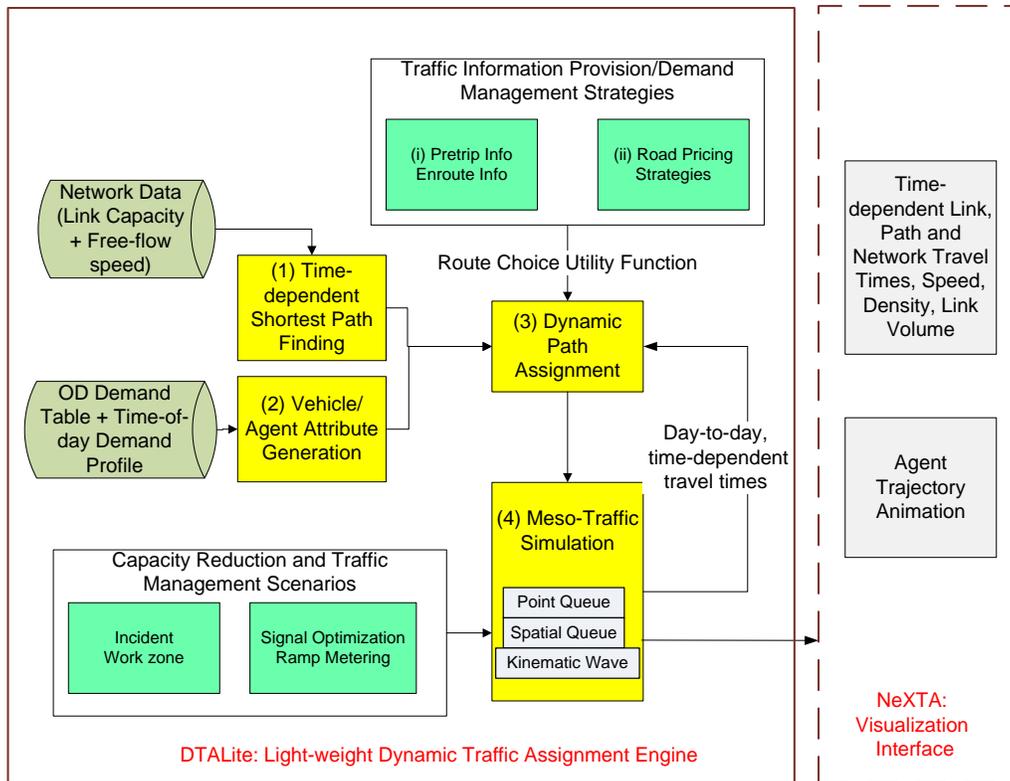
## 4.3 DTALite with InSITE ABM

DTALite, an open-source light-weight, mesoscopic DTA simulation package, in conjunction with the Network eXplorer for Traffic Analysis (NeXTA) graphic user interface (GUI), has been developed to provide transportation planners, engineers and researchers with a theoretically rigorous and computationally efficient traffic network modeling tool. This fully functional, open-source dynamic traffic assignment model can be downloaded from <https://code.google.com/p/nexta/>. In general, the software suite of DTALite + NeXTA seeks to:

- (1) Provide an open-source code base to enable transportation researchers and software developers to expand its range of capabilities to various traffic management applications.
- (2) Present results to other users by visualizing time-varying traffic flow dynamics and traveler route choice behavior in an integrated 2D/3D environment.
- (3) Provide a free, educational tool for students to understand the complex decision-making process in transportation planning and optimization processes.

Additionally, DTALite adopts a new software architecture and algorithm design to facilitate the most efficient use of emergent parallel (multi-core) processing techniques, and exploit the unprecedented parallel computing power newly available on both laptops and desktops.

The overall structure of DTALite, illustrated in Figure 4.8, integrates the four major modeling components, highlighted in yellow.



**Figure 4.8 DTALite software system architecture with key modeling components**

DTALite’s four major modeling components include:

- (1) Time-dependent shortest path finding, based on a node-link network structure.
- (2) Vehicle/agent attribute generation, which combines an origin-destination demand matrix with additional time-of-day departure time profile to generate trips.
- (3) Dynamic path assignment module, which considers major factors affecting agents’ route choice or departure time choice behavior, such as (i) different types of traveler information supply strategies (e.g., historical, pre-trip and/or en-route information, and variable message signs), and (ii) road pricing strategies where economic values are converted to generalized travel time.
- (4) A class of queue-based traffic flow models that can accept essential road capacity reduction or enhancement measures, such as work zones, incidents and ramp meters. The queue-based traffic simulation model in DTALite only requires basic link capacity and free-flow speed for operation, which are readily available from static traffic assignment models. By using simple input parameters, in addition to possible connections with common signal data interfaces, the proposed simulation package may enable state DOTs and regional MPOs to rapidly apply advanced DTA

methodologies for large-scale regional networks, subareas or corridors. Additionally, the modularized system design may help serve future needs by simplifying the process for transportation researchers and software developers to continue to build upon and expand its range of capabilities.

The traffic assignment and simulation modules are fully integrated and iterated to either capture day-to-day user response or find steady-state equilibrium conditions. Within this simulation-assignment framework, the rich set of output data include traffic measures of effectiveness (MOEs) at different spatial and temporal scales, ranging from network, corridor-level and specific links. Typical speed, volume and density measures, and agent-based trajectories can be visualized through the NEXTA user interface.

To allow consistent origin-destination demand estimation (ODME) for fast model calibration, DTALite embeds a path, flow-based optimization model that utilizes sensor data (i.e., observed link flows and densities) and target OD demands to estimate a set of path flow volumes. This approach combines origin-destination estimation/ adjustment with traffic assignment seamlessly. Under this ODME model, DTALite first runs  $K$  assignment iterations (e.g., 40 iterations) to generate likely paths. It then performs another  $K'$  iterations (e.g., 100 iterations) with ODME enabled, to provide the final solution as a set of path-flow patterns satisfying “tolled user equilibrium.”

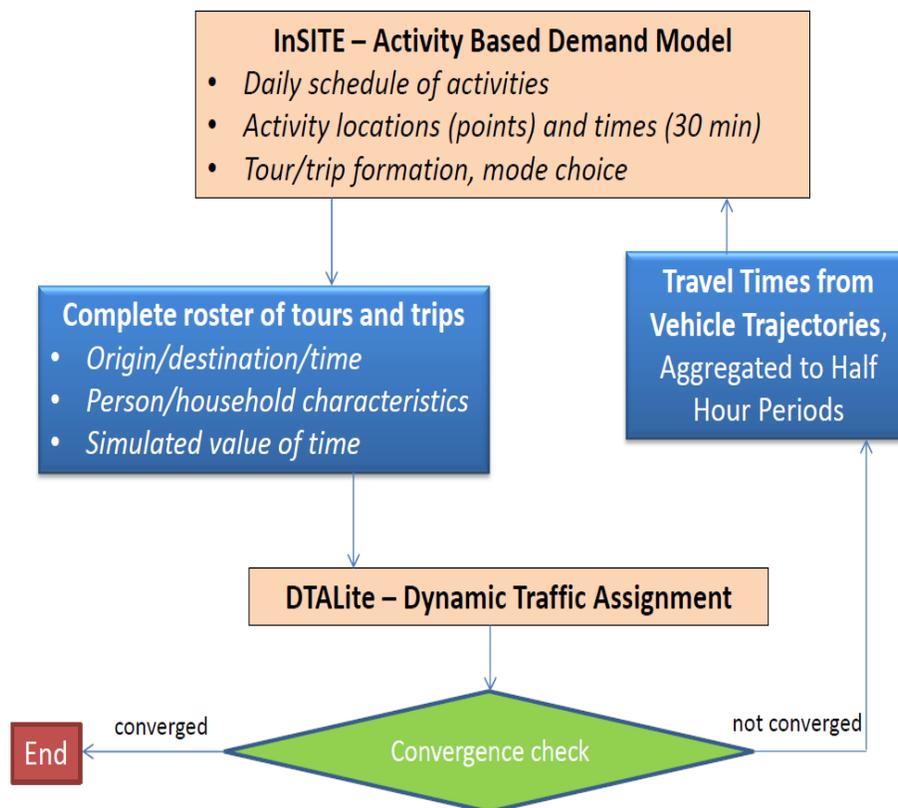
Based on the design structure and queue-based mesoscopic traffic simulation model, DTALite has considerable potential for generalizing the modeling framework into the field of real-time traffic state estimation and prediction. In future research, different ways will be examined for calibrating the maximum queue discharge rates, utilizing end-to-end travel times, and considering an agent-based learning framework to fully consider behavioral heterogeneity.

## 4.4 Integrated InSITE-DTALite – Sequential Integration

The primary objective of the model integration is to produce a working model of the Baltimore region that is disaggregate for auto travel on both the demand and supply sides. This model is intended to be “application ready,” meaning that it has been validated and includes all functionality required for typical planning analyses. Therefore, the integration pursued between InSITE and DTALite falls in the *sequential integration paradigm*.

In this paradigm, the travel demand model (InSITE) and the DTA model (DTALite) are run in iterations, with a single “big loop” consisting of an iteration of the travel demand model followed by a run of the DTA model. The travel demand information (e.g., agents with their activities and travel decisions, as well as characteristics) is passed to the DTA model within each big loop, and travel time information resulting from the dynamic assignment is passed back to the travel demand model for use as input (e.g., update activity pattern choices) in the next big loop. Note that DTA models themselves are run iteratively, and so there are several “small loops” of the DTA run within each “big loop.” A number of “big loops” are run until a measure of convergence is achieved,

where the change in travel times (or some other measure) from one big loop to the next is within a specific tolerance. The data exchange between InSITE and DTALite is such that an agent is preserved during the exchange of data from InSITE to DTALite and vice versa. Thus, the integrated InSITE-DTALite model is fully disaggregated. There is also the *tight integration paradigm*, where both the travel demand model and DTA model share information continuously (i.e., the analyst specifies the temporal resolution for the exchanges) within each big loop. This paradigm is not considered for the integrated model InSITE-DTALite.



**Figure 4.9 Sequential integration process**

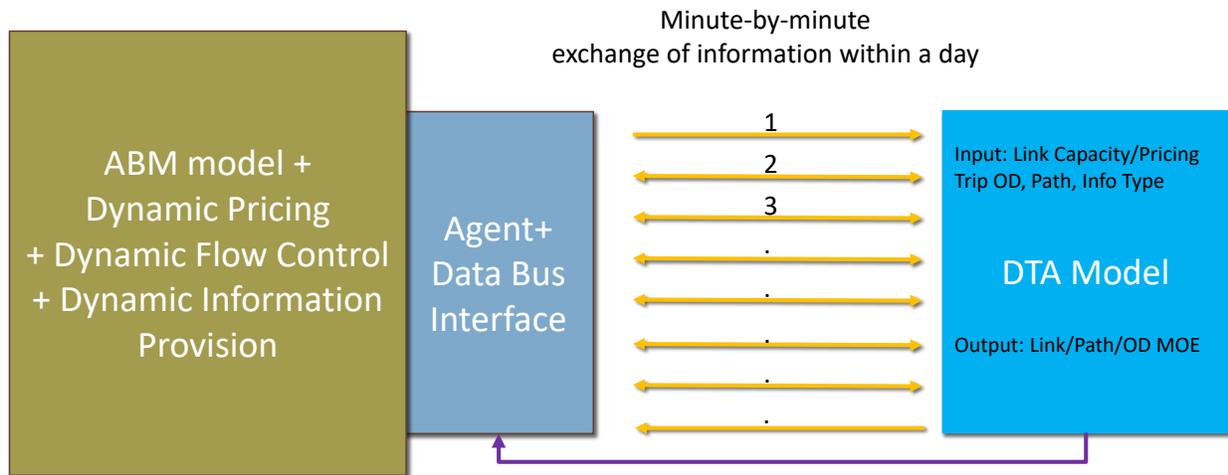
#### **4.4.1 Information to be Exchanged Between Models**

##### **From InSITE to DTALite**

Many traffic analysis tools (e.g., VISSIM, TransModeler, Aimsun, Paramics) have proprietary components that could create impediments for exchanging data across independent software packages, which require the development and application of individual utility functions. While the actual solution will be derived during the project execution, our proposed solution to this challenge is to develop an open data format that allows data conversion and communication utilities from each tool formats to the proposed Agent+ data bus. Although the framework can be implemented

without an Agent+ data bus or an open data communication format standard on a data hub, only ad-hoc linkages across specific software packages can be established. This could fail to allow comprehensive and streamlined analysis of the complex transportation system environment to support a wide range of activity-based model/DTA scenario evaluations.

The research team has already implemented an open-source Agent+ data hub that allows a day-by-day and minute-by-minute integration of ABM and DTA models. New development efforts are needed to implement the data bus for an on-line data communication environment, but can be implemented with minimal risks given the team’s experience in developing the data hub.



**Figure 4.10 Data flow chart of Agent+ data bus interface for integrating DTA and activity-based models.**

**Table 4.24. Example Data Flow of Individual DTA or Activity-based Model Module within Agent+ Data Bus**

Package	Software Packages Connected to Data Bus	Receive Input from Data Bus	Provide Output to Data Bus
Activity-based Model	CSV file reading and writing utility	Day-to-day or min-by-min traffic conditions at link/path/OD levels	Min-by-min traffic request for possible departure time, destination, and mode change
DTA	CSV file reading and writing utility or data receiving and aggregation package	Traffic network, vehicle origin and destination, existing traveled path, real-	Information for upstream vehicles to change routes, updated travel time and queue

time traffic data  
stream

warnings, traffic  
conditions at  
link/path/OD levels

---

### Relationship between data hub and data bus modules

- The **data hub** hosts essential network, demand and signal control data for multi-resolution modeling. The real-world traffic measurements from private data vendors and public agencies can also be archived in the databases.
- The **data bus** aims to enable tight interconnections between three critical entities: data, models, and simulators of decision makers. The data bus provides excellent data interchange interoperability, which holds a key means of accelerating the integration of DTA and activity-based model distributed simulation. In the long run, it allows simulation models to be executed through multiple computation instances in a virtual cloud computing center or in a parallel computing environment.

**Need for synchronizing simulation clock** Multiple modules (e.g., demand simulator, network simulator) within the multi-resolution analysis framework execute in different processors and communicate with each other asynchronously. Different simulation modules can run with different execution cycles. Each step size is determined by the corresponding simulation needs (e.g., six seconds for mesoscopic traffic simulator vs. one minute for travel demand simulator), computational resource constraints, as well as input/output dependency between different models. The proposed data bus concept further highlights the needs for standardized data format across multi-resolution simulation, a tight coupling with activity-based model and DTA simulation, and clearly defined rules for performing multi-horizon simulation/prediction for applications such as dynamic flow control, dynamic pricing and dynamic information provision. In the MITAMS test bed with multiple simulators, it is important to ensure all the computer applications are synchronized with a common simulation clock, while certain high-fidelity simulation systems may be even slower than real-time (e.g., one simulated second is equal to ten real-world seconds).

### Synchronizing simulation clock by using scheduled files as software file token

In general, a security token (or sometimes a hardware token, authentication token, USB token, cryptographic token, software token, virtual token or key fob) may be a physical device that an authorized user of computer services is given to ease authentication. In this work, files are used as software tokens. -This is illustrated in Figure 4.11. -Each DTA/activity-based model process must have the file token to move forward, perform simulation and generate the next token required by the sequential process. The file tokens have fixed names to be better synchronized across processes.

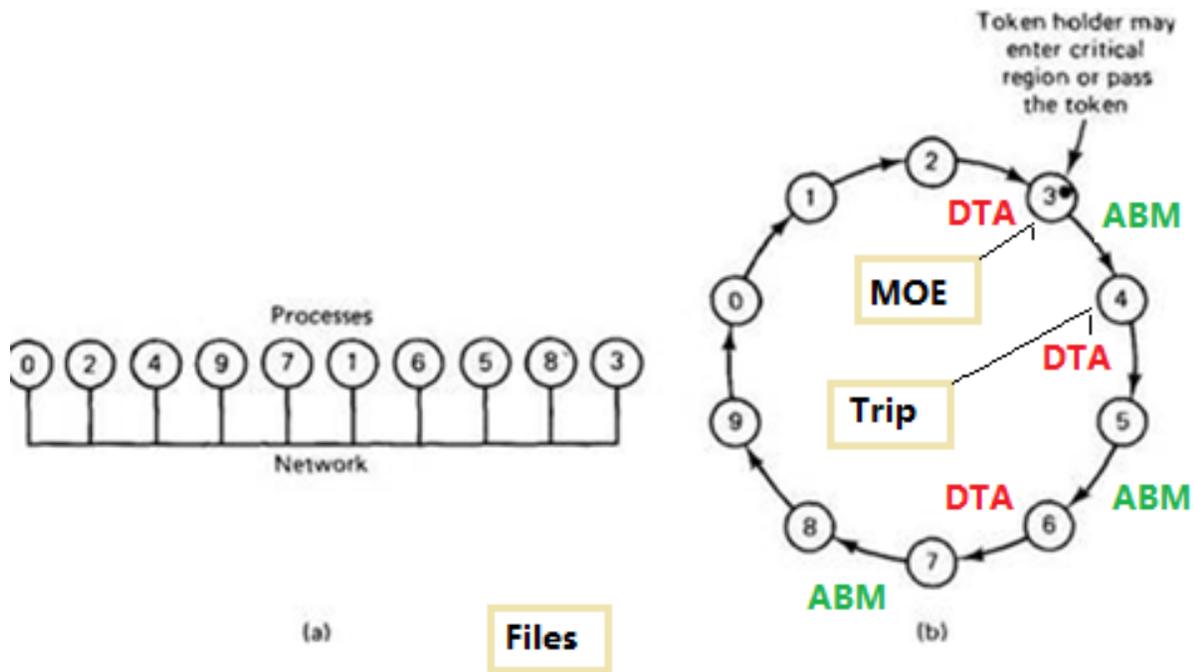


Figure 4.11 Illustration of tokens in a distributed computing environment.

### Data flow chart for each process in the integrated framework

The read input schedule file to follow the token reading, output frequency and file name specification:

```

While (for each x second scheduled)
{
    If ( input file token is ready)
    {
        Perform simulation/analysis
        Generate output file token (for next process)
    }
}

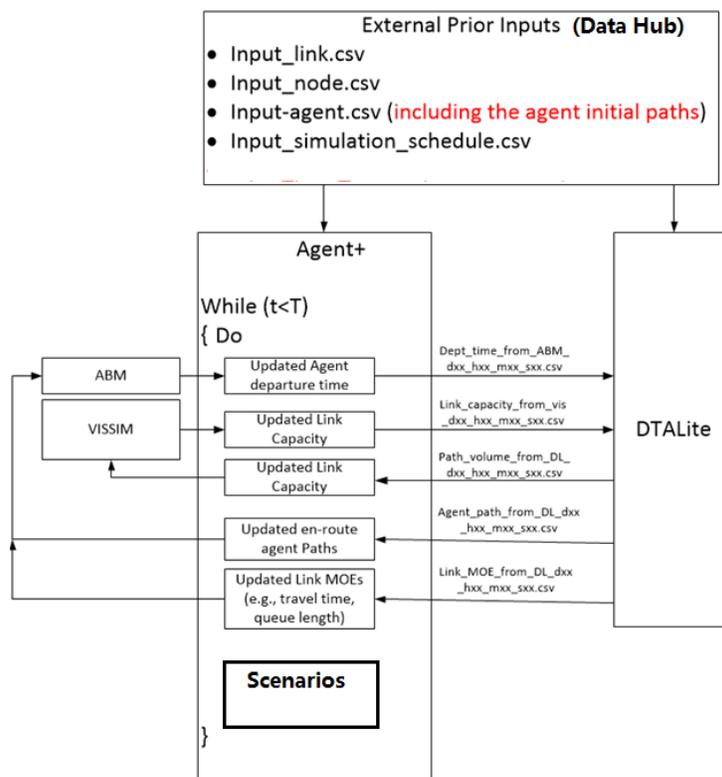
```

Specifically, in the proposed Agent+ modeling framework, the data are first specified updating frequency and file name convention in a file called “*Input\_simulation\_schedule.csv*” for output Link/Path/OD MOE files and for input attributes and agent routing attributes. In this example, the input and output data streams are defined with respect to the DTA module.

attribute	start_time	end_time	RT_Output_LinkMOE	RT_Output_PathMOE	RT_Output_ODMOE	RT_Output_Current_Agent	RT_Input_LinkAttribute	RT_Input_Update_Agent
time_interval	0	300	60	60	300	6	0	300

**Figure 4.12** Sample content of *Input\_simulation\_schedule.csv* file.

By detecting if the required input files are available for every time interval (say, one minute), the Agent+ modeling framework can use the scheduled input and output files accordingly as tokens for synchronizing the simulation clock with the data bus's master clock. See Figure 4.13 for an example.



*Data File Descriptions*

**Figure 4.13** Implemented example of Agent+ framework with connections to data hub.

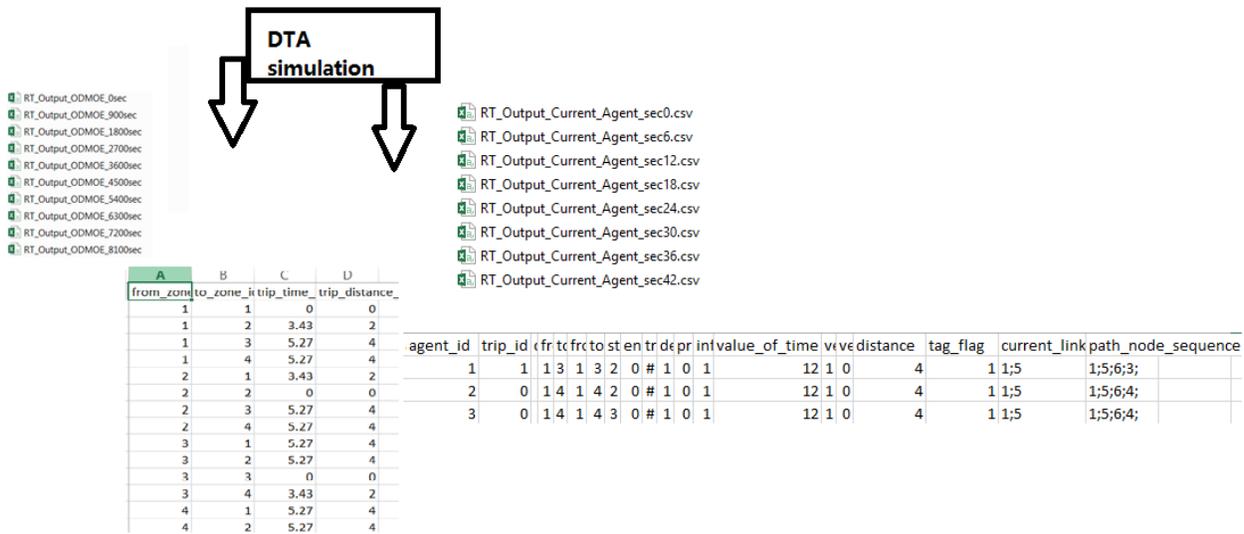


Figure 4.14 Sample files from DTALite.

#### Data Block 1: MOE output files from DTA

- *RT\_output\_linkMOE\_secXX.csv*

Attributes: from\_node, to\_node, travel time, speed and timestamp at 60-second resolution.

- *RT\_output\_ODMOE\_secXX.csv*

Attributes: zone number, travel time, travel distance and number of agents every 10 or 15 minutes.

- *RT\_output\_pathMOE\_secXX.csv*

Major attribute is the node sequence for each path at 60-second resolution.

#### Data Block 2: Current agent list output from DTA, at current time or after a whole day

- *RT\_Output\_Current\_Agent\_secXX.csv*

Main 15 attributes for agents currently in the traffic simulator: agent\_id, detour node sequence, pricing type and value of time, as shown in Figure 4.15.

- *RT\_Output\_End-of-Day\_Agent\_dayXX.csv*

Attributes: Experienced travel times and path for each agent from the previous day.

agent_id	trip_id	of	fr	to	fr	to	st	en	tr	de	pr	in	value_of_time	ve	distance	tag_flag	current_link	path_node_sequence
1	1	1	3	1	3	2	0	#	1	0	1		12	1	0	4	1 1;5	1;5;6;3;
2	0	1	4	1	4	2	0	#	1	0	1		12	1	0	4	1 1;5	1;5;6;4;
3	0	1	4	1	4	3	0	#	1	0	1		12	1	0	4	1 1;5	1;5;6;4;

**Figure 4.15 Sample attribute list about current agents in DTA simulator**

### From DTALite to InSITE

The main data to be passed from DTALite will be related to highway travel times. The way static highway assignments will be executed in the original version of InSITE is discussed below to provide an understanding of how travel time inputs are fed back from highway assignments in an iterative process.

InSITE uses 48 aggregate time periods of 30 minutes in length. In its original formulation, the static highway assignment in InSITE is expected to be performed for eight time periods that range in length from 1 hour during peak periods, to 11 or 12 hours for the overnight period. All of the periods are groups of 30 minute periods, and so the trip tables used in assignments are aggregated from the half-hour periods.

For the first iteration of the demand model components in InSITE, travel time inputs are “skimmed” from the highway network (and the transit network) for four broadly defined periods: a.m. peak, mid-day, p.m. peak, and off-peak. After the highway assignment is run in the first iteration, travel times are available for each assignment period, and the loaded network resulting from the highway assignment for each of the eight periods will be skimmed. The travel time inputs used for each of the half-hour periods for the demand model will come from the skims for the assignment period where the half-hour period falls.

The skimming process builds paths between all pairs of zones, using the volume on each network link to estimate an average travel time for the link for the time period. This is done using a simple formula relating the link travel time to its free flow time, capacity and volume, as loaded during the highway assignment for the period. Since any highway link could be part of a path, travel time estimates are computed for every link.

In the integrated model, the DTA will replace the static highway assignment process. The DTA simulates highway travel dynamically over the entire day from beginning to end (because InSITE defines the beginning and end of the day at 3:00 a.m., this time will signify the start and end of the day for the DTA as well). There are no time periods used for assignments that are longer than the half-hour periods used in InSITE; therefore, DTALite will not produce travel times for longer periods. After the initial iteration, the travel time inputs to InSITE can be unique for each half-hour period.

Unlike the static process, the DTA does not produce a loaded network with volumes representing the entire period. With a DTA model, the network travel times are constantly changing as the simulated vehicles move through the network. Since the demand model uses aggregate time

periods, however, travel times between zones are needed as inputs for the demand model components.

It will therefore be necessary to create link travel times for each half-hour period by averaging the times experienced by the vehicles in the DTA during the period. To do this, it may be possible to adapt the process developed in the SHRP C10B project (see Cambridge Systematics, Inc., et al, 2014, Section 2.3.1). The feedback process employed in the C10B model combines information from the DTA for all relevant trajectories within a time period (half-hour or broad period) to estimate an average time that can be used for input to the activity-based model.

## 4.5 Convergence Paradigm in Sequential Integration

The convergence paradigm adopted for the sequential integration is based on the concept of activity-schedule gaps and is explained as follows. Imagine you have an activity-schedule sequence for an agent as the one shown in the figure:



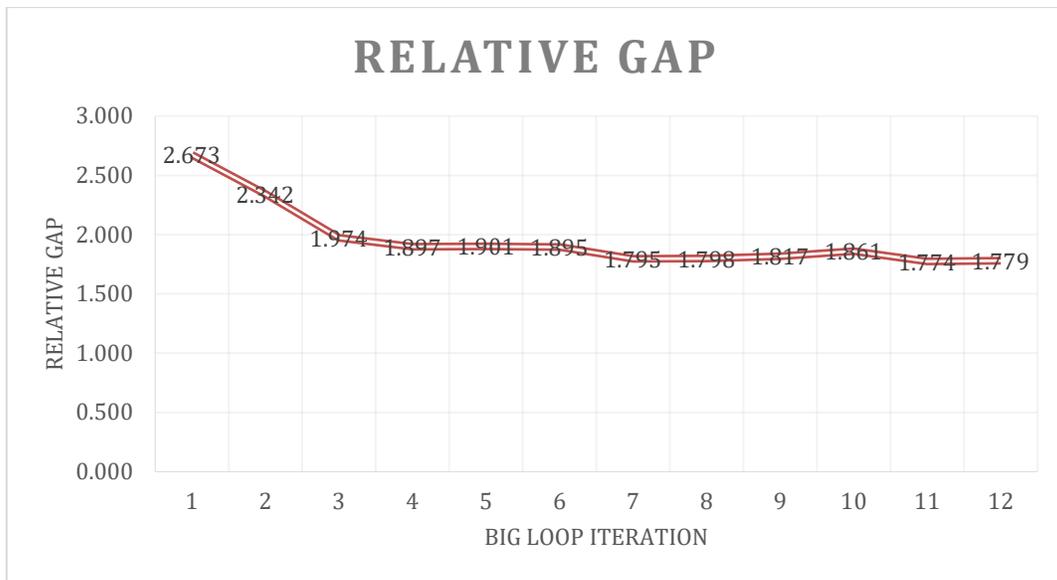
- Primary represents the primary activity of the travel activity, for example, work.
- Stops represents secondary activities, such as shopping and eating out.
- Trv represents traveling by a given mode (car, transit, etc.).
- The  $t_s$  represent departure times and arrival times to the activities.

For an agent, InSITE ABM will use the static assignment skims (in the big loop iteration 1) or the travel time skims from the previous iteration, generated after applying the method of successive averages (MSA) to the DTALite travel time skims in order to simulate the travel decisions of the agent. These simulation results are then put together in an agent list of trips and provided to DTALite to perform DTA. The DTA results will generate new travel time skims for the current iteration and MSA will be used to average them.

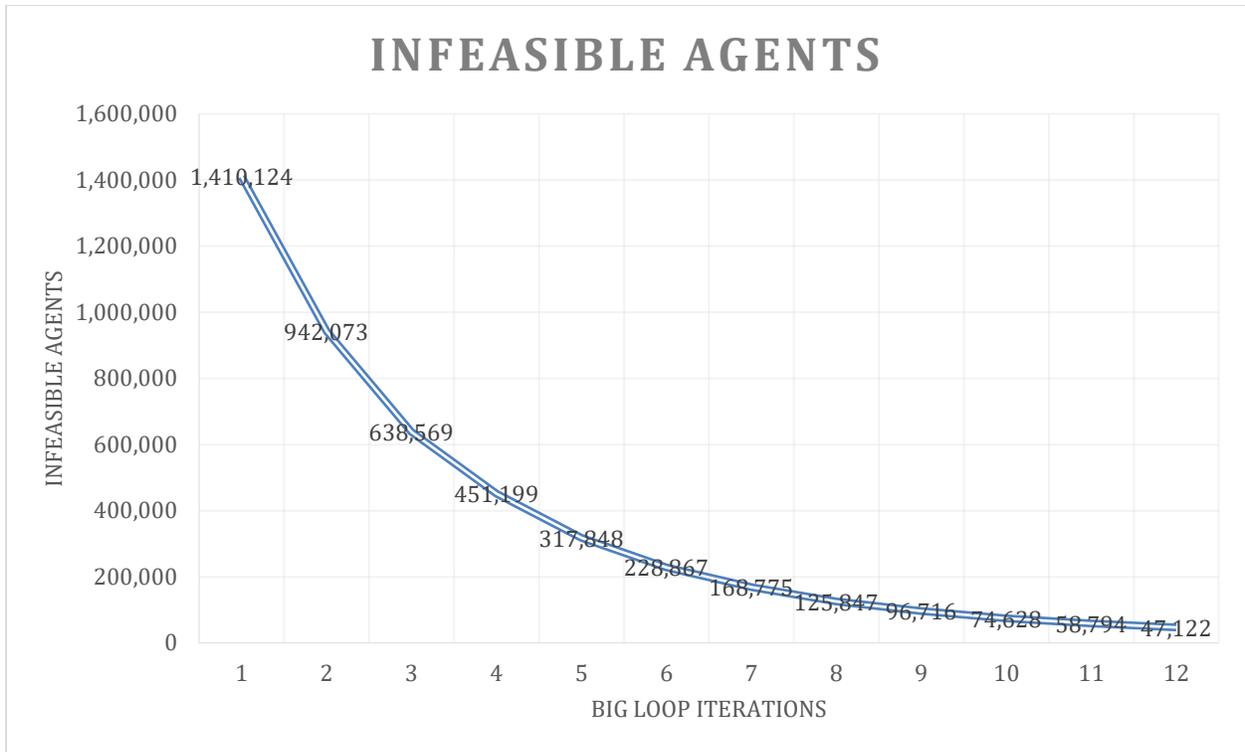
Convergence will be based on measuring scheduling gaps, which are defined as the changes in travel times between the previous skims and the current travel time skim for the same Trv element. Furthermore, agents traveling where the current travel time skims reduce their activity durations to durations less than the minimum activity durations allowed will be identified. The minimum activity durations are obtained from the Household Travel Survey. These agents are flagged as infeasible and are re-simulated by InSITE ABM in the next big loop iteration. Agents not flagged as infeasible are not re-simulated by InSITE ABM in the next big loop iteration. The activity schedules of both agents re-simulated and agents not re-simulated are provided to DTALite in the

next big loop iteration to perform DTA. It should be noted that an agent not re-simulated in one big loop iteration could be re-simulated in another big-loop iteration. In general, the number of infeasible agents will decrease with probability one as more big loop iterations are completed, if the travel times converge with probability one to a constant value compatible with the minimum activity durations.

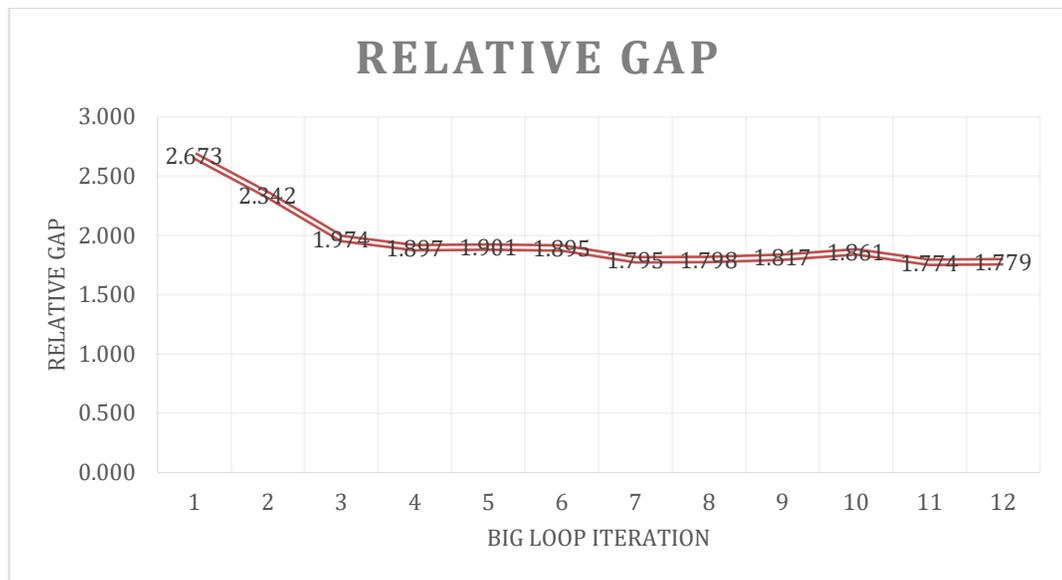
DTALite simulations across big loop iterations will become consistent as the number of infeasible agents decreases to a value less than 50,000 agents. In other words, link volumes and OD volumes generated by DTALite across big iterations will become approximately the same for all links. This is a result of DTALite being deterministic and only InSITE ABM being stochastic. The stochasticity in the travel decisions will not significantly affect DTALite, as only the travel decisions of the shrinking infeasible agents change across big loop iterations, and the travel decisions of the feasible agents become constant across iterations. However, the number of infeasible agents is stochastic; it is likely that it will not reach zero and will likely fluctuate (increase and/or decrease across iterations eventually). DTALite results will still be approximately consistent across big loop iterations if the number of infeasible agents is significantly small, compared to the total number of agents. The relationship between number of infeasible agents and number of big loop iterations is shown in figure.



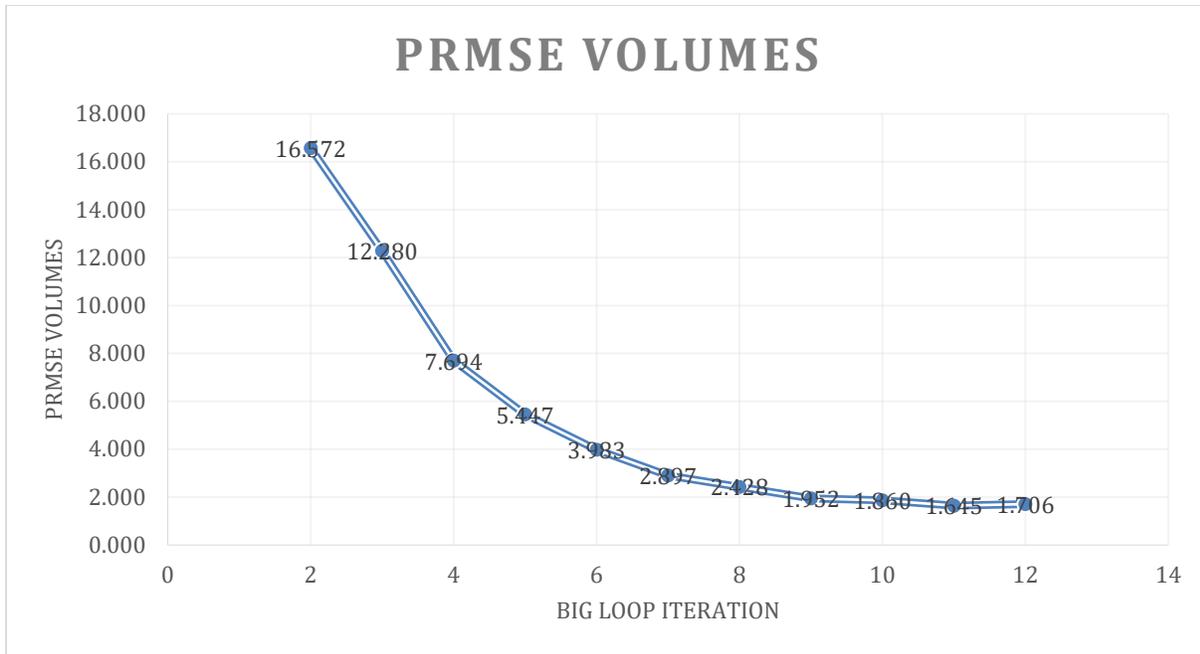
**Figure 4.16 Relative gap DTALite vs big loop iterations**



**Figure 4.17** Number of infeasible agents vs big loop iterations



**Figure 4.18** Relative gap vs big loop iterations.



**Figure 4.19 PRMSE Volumes vs Big loop iteration**

As the number of infeasible agents decreases, the volumes start to become stable. However, some fluctuation is observed as the number of infeasible agents starts to fluctuate instead of further decreasing.

## 4.6 Integrated InSITE-DTALite - Validation

Validation is a critical element in ensuring that the new model is robust and that agencies can use the finished product. Establishing a validation plan early in the model development process helps ensure that appropriate validation data are assembled or collected, and that adequate time and resources are devoted to the validation process.

Two separate types of tests, validation tests and sensitivity tests, are described in this section. For clarification, the following definitions will be employed in this project:

- **Calibration** (not covered in this section) is the adjustment of model parameters to instruct models to replicate observed data for a base (calibration) year, or otherwise produce more reasonable results.
- **Validation** is the application of the calibrated models and comparison of the results against data not used for estimation or calibration. Validation data may include additional data collected for the same year as the estimation or calibration of the model, or data collected for an alternative year.

- **Sensitivity testing** is the application of the models and model set using alternative input data or assumptions. While sensitivity testing may be performed for individual model components through the estimation of the elasticities and cross-elasticities of model coefficients, the focus of sensitivity testing for this project will be the application of the entire model set, using an alternative forecast or backcast year, and alternative transportation system or policy perturbations. The focus of the sensitivity testing will be the plausibility and reasonability of the model results.

#### **4.6.1 Summary of Validation Data**

- Directional hourly count data of 15 vehicle classes by day of week is available at 1122 locations within the BMC network.
- Directional hourly speed data by day of week is available at 52 locations within BMC network.
- The Maryland Mobility Report, provided to by the BMC. We have coded the bottleneck locations in the network based on this report. We will combine this information with information from Google Maps to identify bottlenecks.

#### **4.6.2 Calibration and Validation Process of Integrated InSITE-DTALite**

For the model calibration, a sequential calibration approach was adopted. This is primarily because the run-times of the integrated model are between 1.5 to 1.8 days per big loop iteration (the main load residing in the InSITE ABM as it is responsible for about 60 percent of the run-time). This run-time is prohibitive for a simultaneous calibration of both the InSITE ABM and DTALite. Therefore, InSITE ABM was calibrated separately with the Household Travel Survey data; DTALite, using the demand from InSITE ABM, was calibrated separately using the base year data.

For the validation, two aspects of traffic assignment should be considered: the accuracy of the process (i.e., model fit) and the inherent variability in the assignment process. It should also be noted that the InSITE ABM, as discussed previously, was extensively validated against the Household Travel Survey data.

The DTALite trip assignment process introduces a time dimension to the assignment process. Specifically, DTALite explicitly accounts for the fact that time is required for a vehicle to traverse the transportation network, from his or her origin to destination. This is different from the static equilibrium assignment process used with the initial InSITE development that, in effect, models vehicles as simultaneously being on all links connecting the origin and destination.

The time dimension in DTALite makes it possible to identify bottlenecks, queue locations and lengths, and congestion delay in ways not possible with static equilibrium or normal transit assignment processes. While the modeling of bottlenecks and queue lengths adds more realism to the trip assignment process, it also introduces more complexity in the model validation process. Specifically, bottlenecks and queues do not form in the same way every day. There is randomness to daily traffic and transit volumes, resulting travel times, and other impacts such as congestion delay. Identifying these issues in observed data may be a difficult task.

Observed traffic data will form the basis of integrated model validation comparisons. A structured, two-level approach will be used:

- The Level 1 validation is comparable to traffic assignment validations typically performed for regional travel models. Generally, the Level 1 tests include aggregate (groups of links) comparisons of model volumes to observed data for the model base year.
- The Level 2 validation is comparing the traffic assignment results from InSITE ABM static assignment, with the DTA assignment results for volumes and VMT.

#### **4.6.2.1 Trip Assignment Validation**

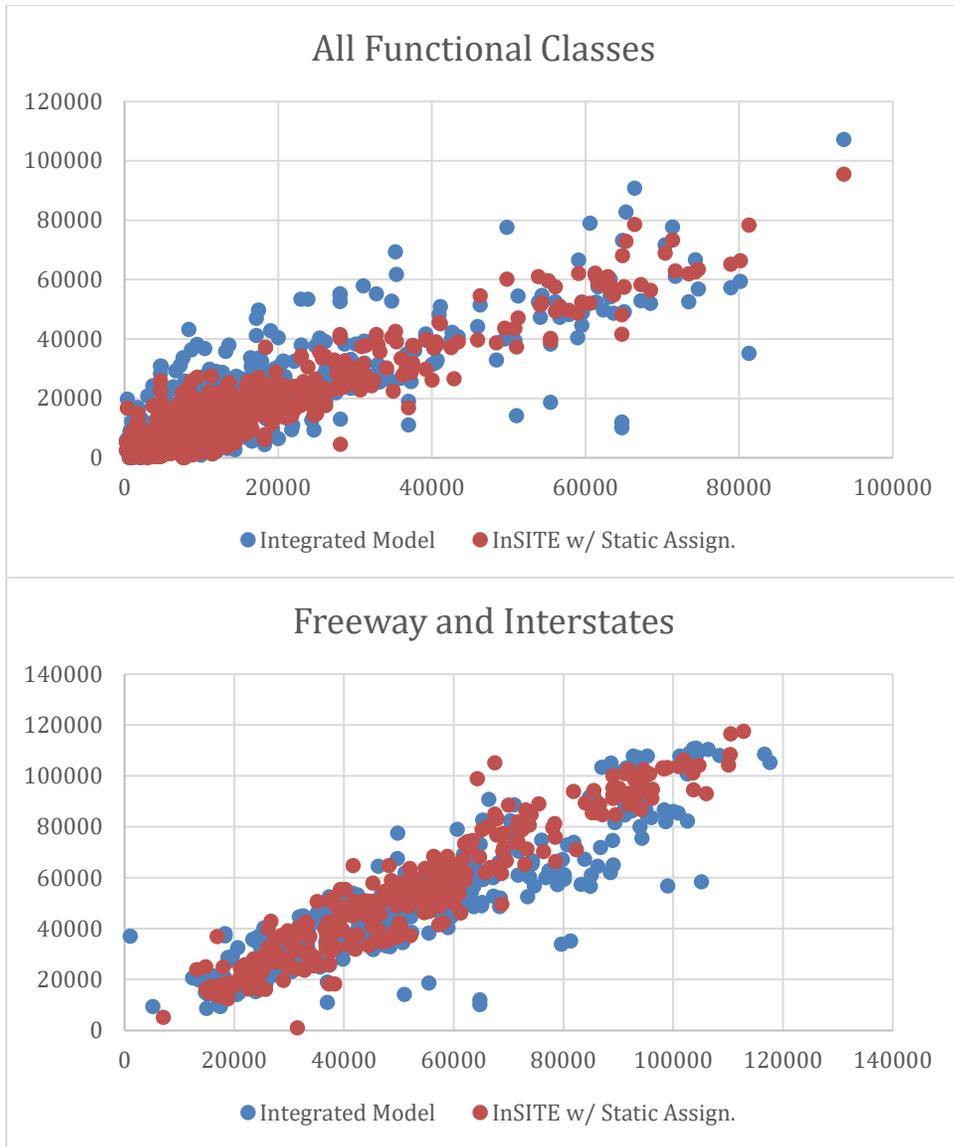
The primary focus of the Trip Assignment Validation will be the ability of the Integrated Model assignment to reproduce observed daily traffic volumes and vehicle-miles traveled (VMT). This validation can be considered a model system validation, since it will be impacted by the travel models embodied in both InSITE and DTALite.

**Table 4.25 Volumes by Functional Class and Area Type**

Stratification	Percent Difference
Functional Class	
Freeways/Expressways	±7%
Principal Arterials	±10%
Minor Arterials	±15%
Collectors	±20%
All Links	

**Table 4.26 Level 1 Validation Tests (Aggregated to Daily)**

Validation Focus	Validation Measures	Expected Outcomes
Individual link volumes	<ul style="list-style-type: none"> <li>• Root mean squared error (RMSE) and percent RMSE by: <ul style="list-style-type: none"> <li>– Functional class</li> </ul> </li> <li>• Anomalous links: <ul style="list-style-type: none"> <li>– Links with 0 volumes</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Visual inspection for large errors in modeled link volumes or for general trends in errors.</li> <li>• %RMSE by volume group from guidelines for several static equilibrium traffic assignment model validation guideline documents (for daily traffic volumes). Results from DTALite should, at a minimum, satisfy similar guidelines.</li> </ul>



**Figure 4.20 Observed daily volumes vs. model daily volumes.**

**Table 4.27 PRSME of Volume by Functional Classes**

<b>Function Class</b>	<b>InSITE ABM w/ Static Assign. (PRMSE)</b>	<b>Integrated Model (PRMSE)</b>
<b>Interstate</b>	-2.18%	-3.69%
<b>Freeway</b>	-6.28%	-3.24%
<b>Primary Arterial</b>	1.65%	5.78%
<b>Minor Arterial</b>	-10.96%	4.06%
<b>Collector &amp; Others</b>	-11.91%	-2.65%
<b>Total</b>	-3.53%	0.84%

The results presented in the figures and table highlight the performance of the integrated model relative to the InSITE ABM with static assignment. The results of the integrated model are more accurate to the observed daily count data, even when looking at specific functional classes, with some exceptions. However, additional validation is required to further assess performance of the integrated model.

## 4.7 Integrated InSITE-DTALite - Application

For the application, Baltimore Metropolitan Council (BMC) wanted to test a land use change scenario in the Brownfield area at Port Covington. This scenario is also for the base year 2012.

For the scenario, the changes are briefly summarized as follows:

- A new population was synthesized to account for the increased number of workers in the Brownfield area.

**Table 4.28 Average Percentage Change of Travel Time by Functional Classes**

<b>Functional Class</b>	<b>Base vs. Brownfield</b>
<b>Interstate</b>	0.09%
<b>Freeway</b>	-0.31%
<b>Primary Arterial</b>	1.05%
<b>Minor Arterial</b>	3.33%
<b>Collector &amp; Others</b>	0.97%
<b>Total</b>	1.46%

- Land use changes were conducted to spur new jobs and related new opportunities stemming from the redevelopment of the Brownfield area.

Preliminary results compared to the base scenario are presented as follows:

The results indicate that most of the increase is in the arterials (both primary and secondary) in terms of travel time, particularly those arterials closest to the Port Covington redevelopment. The interstates will also see an increase, especially those close to Port Covington. Further analyses are required to understand the system's response to the redevelopment.

## 5. Moving towards Tighter Integration of Dynamic Activity-Travel Demand and Network Assignment Models

In this project, the integration between InSITE and DTALite falls in the sequential integration paradigm. This chapter depicts how tighter integration can be accomplished. The tighter integration paradigm described in this chapter uses a continuous time activity-based model (openAMOS) and DTALite within the SimTRAVEL framework. Enhancements to InSITE would have to be made to facilitate the type of tight integration presented in the chapter. Making such enhancements to InSITE and developing a tight InSITE-DTALite integrated modeling platform constitutes a major future research direction.

### 5.1 Introduction

The contemporary technological landscape has made information available on a number of platforms, creating opportunities to influence, manage, and enhance transportation network performance with respect to a number of possible criteria (e.g., delay, congestion, mode share, energy and emissions footprint) and under a wide variety of scenarios (e.g., regular traffic, work zones, evacuation situation, special events). In cities around the world, transportation and mobility professionals are exploring ways to leverage the power of technology, connectivity and communication capabilities. Tools are being developed to manage congestion, influence travel choices, and optimize transportation system performance, using analytics derived from big data streams obtained from a variety of sensors, embedded throughout the network. With the rapid development of connected and automated vehicles, and the availability of transportation-related apps that provide mobility options and traffic conditions in real time, the role of technology in shaping and managing the transportation system will only increase in the years ahead.

The U.S. Department of Transportation (USDOT) has several intelligent transportation system initiatives underway intended to provide jurisdictions around the country a suite of tools, strategies and methods for enhancing system performance. The Applications for the Environment: Real-Time Information Synthesis (AERIS) Program, for example, included a suite of strategies to reduce the energy and emissions footprint of travel in regions through the use of connected vehicle technologies, including vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications (USDOT, 2016a). More recently, there has been a growing level of emphasis on the deployment of dynamic mobility applications (DMA) and active transportation and demand management (ATDM) strategies. DMA include a number of technologies aimed at enhancing network operations for a variety of system users, including auto users, transit riders, and freight transport providers (USDOT, 2016b). ATDM strategies include a suite of approaches to proactively manage travel demand (e.g., dynamic pricing schemes and the provision of information to system users), traffic flow (e.g., dynamic speed limits and adaptive ramp metering), and parking operations (e.g., dynamic parking pricing and dynamic way-finding) (FHWA, 2012).

With rapid advances in technology and communication systems, the line that has traditionally separated transportation planning from traffic operations is blurred. An examination of the strategies and techniques in the DMA and ATDM toolboxes clearly illustrate that transportation professionals are moving towards erasing the line that divides, and more closely integrate planning and operations. Several DMA and ATDM strategies are aimed at influencing and managing travel demand (e.g., pricing and traveler information) while others aim to improve network operations (e.g., intelligent traffic signal control systems and adaptive ramp metering). The notion of “planning for operations” or “operations for planning” is motivating the seamless blending of planning strategies and operational measures to optimize system performance.

This environment calls for the development of analysis, modeling and simulation tools capable of reflecting the dynamic interactions between planning strategies and network operations. In the past, travel demand model systems have been limited in their ability to consider such dynamic interactions. On the one hand, travel demand models purport to simulate choices related to activity generation, destination choice, mode and route choices as a function of socio-economic, demographic, built environment, and network attributes. On the other hand, transportation network models and traffic microsimulation models have been used to simulate traffic flow dynamics at a fine level of spatial, temporal and network detail. These two modeling enterprises have largely developed independent of one another; connections between them have been achieved mostly through loose sequential coupling of demand and network model systems, incapable of reflecting the impacts of real-time operational strategies on the full slate of travel choices and network dynamics. Mobile technologies now make it possible for travelers (on all modes in both personal and freight transportation domains) to adjust travel choices in response to prevailing or predicted network conditions. Unfortunately, integrated models of travel demand and network dynamics capable of fully accounting for such real-time phenomena have not yet matured to a stage where they can be implemented in practice.

In light of the need for robust and behaviorally sensitive integrated transport demand and network models, this chapter presents a framework that reflects travel demand and network dynamics arising from real-time and proactive dynamic mobility/traffic management strategies. The framework can be enhanced through a series of levels, with each level more complex than the previous, but with the ability to account for a larger set of dynamic behaviors. The chapter presents various levels of dynamic integrated transport demand and network modeling to demonstrate the efficacy of the approach.

The scope of the framework presented here is limited to the integration of activity-based travel microsimulation model systems of travel demand with dynamic traffic assignment models of network dynamics. These model systems are used more frequently to simulate the impacts of dynamic mobility management strategies, because of their ability to accommodate time-space interactions at the level of the individual traveler (or vehicle). Incorporating interactions between travel demand choices and network dynamics within the four-step travel demand modeling framework is challenging, due to the inability to track individuals through space and time. In addition, the use of activity-based microsimulation models in conjunction with dynamic traffic assignment models allows the recognition of behavioral heterogeneity in the population (for example, not everybody has the same value of time, inclination to use and follow traveler information, or level of schedule flexibility). Thus, the framework presented in this chapter is limited to the integration of activity-based microsimulation models of travel demand with dynamic

traffic assignment models. The further integration of traffic microsimulation models is also beyond the scope of this discussion.

### **5.1.1 Integrated models of transport demand and network dynamics**

Significant advances have been made in the development of integrated models that link state-of-the-art, activity-based travel demand models (ABMs) with dynamic traffic assignment (DTA) models. While some of these models operate in a sequential coupling framework where exchange of information between the model systems occurs at the end of a full iteration of each model system, others employ a tighter integration where information is exchanged on a more continuous basis. Lin et al (2008) present the conceptual framework for an integrated ABM-DTA model system in which the model components are loosely coupled together in a sequential fashion; the travel times from the DTA model are fed back as input to the ABM, only at the end of a full model iteration. In other words, within-day or within-tour adjustments to travel choices (in response to network conditions) cannot be modeled using their framework. Hao et al. (2010) present the results from an integration of the TASHA ABM with MATSim, which constitutes another example of an integrated model system where the components are connected in a similar sequential paradigm. Bekhor et al (2011) discuss the results from an integrated model effort, where the Tel-Aviv ABM was coupled with MATSim. For purposes of integration, the time-of-day choice was disaggregated into fifteen distinct time periods, while considering activity duration constraints. Goulias et al (2012) present the results from an integration of the CEMDAP ABM with a multi-period static traffic assignment model and report on-going efforts for integration with a DTA model. All of the model systems discussed above are based on a loosely coupled or sequential integration paradigm where the ABM and DTA models communicate and exchange data only at the end of a full iteration. This sequential information exchange process renders the use of such models to realistically represent the impacts of network events or disruptions on activity-travel patterns quite challenging. In recognition of this limitation with sequentially integrated model systems, a stream of integrated models that adopt a much tighter integration framework have been developed.

Balmer et al (2009) present results from an operational prototype of the MATSim integrated model. MATSim is comprised of several modular components that work in concert to link the activity-schedules obtained from an ABM with network phenomena captured by a DTA model. Pendyala et al (2012) present the first fully operational prototype of an integrated model called SimTRAVEL, in which the ABM and the DTA model systems constantly communicate and exchange data with each other on a minute-by-minute basis, thus rendering the integration process truly dynamic. Auld et al (2016) propose a conceptual framework and present results from a test implementation of POLARIS. POLARIS also executes a continuous exchange of information between the ABM and DTA components. Such a tight integration allows for realistic representation of impacts of network dynamics—and real-time information systems and proactive traffic management strategies—on traveler behavior.

Real-time information systems (RTIS) or advanced travel information systems (ATIS) may impact various dimensions of travel behavior, including destination, time-of-day, route, and mode choices. Several studies have attempted to model the influence of RTIS/ATIS on travel behavior. In an empirical study using an interactive simulator, Srinivasan and Mahmassani (2003) concluded that route-switching behavior is significantly influenced by the nature, timeliness, quality and extent of real-time information. Sun et al (2005) developed a model that employed Bayesian

updating approaches to account for the impact of travel information on activity-travel decisions. Jang et al (2013) proposed a utility maximization-based activity rescheduling model to allow adjustments in a traveler's daily activity-travel schedule in response to time pressure or time surplus. Ouyang et al (2011) introduced an activity-based network user equilibrium algorithm for modeling activity-travel patterns under network congestions. Paz and Peeta (2009) proposed paradigms that enhance the effectiveness of real-time traffic routing strategies by explicitly accounting for driver's likely response to path recommendations using an information-based network control approach.

Considerable progress has been made recently in modeling/representing the influence of RTIS/ATIS on travel demand using integrated microsimulation model systems. Frei et al (2014) integrated demand models with weather-responsive network management methodologies and assessed the behavioral responses of travelers along several dimensions in response to weather-responsive traffic management strategies (e.g., pre-trip information and rescheduled work and school hours). Zhang et al (2015) proposed a routing and guidance strategy (network assignment model) that can be integrated within an overall agent-based simulation model system. Their study involved modeling the impact of real-time network information on evacuation scenarios, but their framework has a scope for wider application. Xiong et al (2016) developed a model of en-route diversion by integrating an agent-based (microscopic) en-route diversion model with a (mesoscopic) DTA model. They conducted a simulation of the impacts of variable message signs, providing real-time information on a sub-network for the Washington/Baltimore region.

Bustillos et al (2011) studied the effects that pre-trip and en-route traveler information have on route choice. They developed models of congestion-responsive rerouting, pre-trip information-responsive rerouting, and en-route information-responsive rerouting, and then integrated these models into a DTA model. Konduri et al (2013) modeled the impact of network disruptions on activity-travel choice patterns under varying levels of traveler information provision using SimTRAVEL (Pendyala et al, 2012). Mitsakis et al (2015) proposed the development of an integrated framework for modeling the impacts of real-time advanced traveler information systems and applied the model system to a large-scale network for Thessaloniki, Greece. POLARIS (Auld et al, 2016) is another example of a model system capable of simulating the impacts of real-time information systems, by incorporating such behaviors as trip re-planning and en-route switching. Bellemans et al (2010) present the implementation framework of FEATHERS, a comprehensive model system capable of modeling travel scheduling decisions, within-day rescheduling, and learning processes of individual agents.

Examples from the literature cited above illustrate the depth of interest in the development of integrated microsimulation models of activity-travel demand and network dynamics. Efforts to develop such models continue to be rather fragmented, and there is no comprehensive modeling and simulation framework that can accommodate and reflect the varied impacts of the full suite of dynamic mobility management strategies. Seeking to fill this critical gap in knowledge, this chapter presents and describes the various levels of a comprehensive conceptual framework for integrated microsimulation modeling of transport systems that builds on the SimTRAVEL structure documented previously (Pendyala et al, 2012).

## 5.2 Levels of an Integrated Transport Modeling Framework

This section presents conceptual designs of increasing levels of dynamic integration between an activity-based travel demand model and a dynamic traffic assignment model. It seeks to incorporate the simulation of the impacts of dynamic mobility management strategies on the activity-travel choices of agents in the system. The conceptual designs presented here build on the SimTRAVEL framework.

### 5.2.1 Level Zero: Sequential Integration

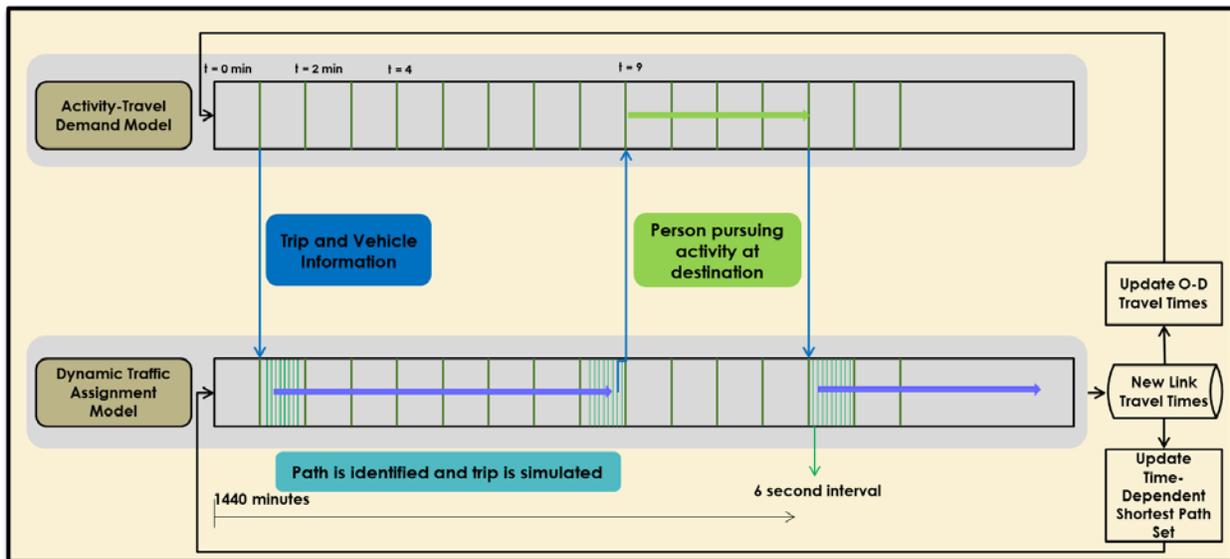
The simplest conceptual design of an integrated model system involves the loose sequential coupling of the two major model components: ABM and DTA models. In the sequential model design, the activity-based model system and the dynamic traffic assignment models are run sequentially in their entirety in a looping fashion. The ABM is run for a 24-hour period, for example, to simulate daily activity-travel patterns for all agents in the synthetic population of the region. Travel time, cost and other skim (network attribute) matrices are derived from the transport network files. The entire trip list generated by the ABM is then input into the DTA model. The DTA model routes and simulates all of the trips (or tours, in some instances) on the network. Within the DTA model, time-dependent, shortest paths are computed and trips are routed on these paths. Iterations internal to the DTA model may be invoked to achieve dynamic user equilibrium, where trips between origins and destinations are routed along  $k$  time-dependent shortest paths to reflect traveler route choice behavior. When travelers on the network can no longer improve their travel times by switching routes, as evidenced by the computation of a gap function (Chiu et al, 2011), the model outputs assignment completes, combining time-of-day with a new set of skim matrices that can be input into the ABM for the next iteration of the integrated model system. The ABM uses these time-dependent skim matrices to generate a new daily activity-travel pattern for each agent in the population and the new set of trips or tours are input to the DTA model. The DTA model assigns the trips and writes out an updated set of skim matrices. This process is continued until there is convergence in the OD matrices output by the model system from one iteration to the next. Other mechanisms to monitor and test convergence include a check of consistency between time-dependent skim matrices output by the DTA model (and those used to simulate activity-travel demand choices in the ABM), and the use of a difference function that demonstrates stability in link flows and end-to-end travel times across successive iterations of the integrated model system. This sequential paradigm is the simplest level of model integration and has been documented in the literature (e.g., Lin et al, 2008). However, this model integration platform is not responsive to network dynamics and real-time information that may be available to a traveler. If there is an instantaneous network disruption (due to an incident) that lasts for a finite period of time and results in a substantial increase in travel times along certain links of the network, travelers may divert to alternative routes or adjust their activity-travel schedule to accommodate the unexpected disruption. However, in the sequential paradigm, there is no ability to simulate such behavioral adaptation mechanisms.

### 5.2.2 Level One: Dynamic Integration – No Real-time Information

The first level of dynamic integration corresponds to the SimTRAVEL framework, which is documented in Pendyala et al (2012). SimTRAVEL refers to a Simulator of Transport, Routes,

Activities, Vehicles, Emissions, and Land and was developed as part of an early Exploratory Advanced Research Program (EARP) project of the U.S. Department of Transportation. The framework is inspired by earlier work presented by Kitamura et al (2008), who envisioned a tightly integrated modeling platform to reflect behavioral dynamics both on the demand side and network side. The framework is briefly presented here; more details are available in Pendyala et al (2012).

The original SimTRAVEL framework is shown in Figure 5.1. In this framework, the ABM and DTA models are run in parallel with data exchange occurring along the continuous time axis at a time-step specified by the analyst. Data may be exchanged at the resolution of one minute or more aggregate time-steps, depending on the size of the network and the number of trips being simulated, to manage computational complexity. An initial set of coarse time-of-day period-specific network skim matrices are obtained from an existing, four-step travel demand model or a set of bootstrap runs, in which the ABM and DTA models are run in sequence (Level 0). The dynamic integrated platform is then launched with the ABM, simulating activity-travel choices at the resolution of one minute. In each minute of the day, the ABM provides a list of trips to the DTA model that are departing within that minute. The DTA model then routes and simulates the trips on the network along time-dependent, shortest paths that are identified based on link travel times computed and saved from a prior iteration of the model system. In each minute of the simulation day, travelers will arrive at their destinations, and the DTA model will send the set of trips (travelers) that have arrived to the ABM so that the ABM can determine the subsequent activity-travel choices based on destination arrival time. When trips (travelers) are navigating the network, the ABM is not concerned about them, even though the DTA model updates traveler (vehicle) position on the network in six-second time steps; the ABM is concerned with simulating subsequent activity-travel choices only after the traveler has arrived at the intended destination. In simulating activity-travel choices (such as activity duration, next activity to be pursued, destination and mode choice), the ABM explicitly recognizes and considers time-space prism constraints associated with mandatory or fixed activities; the need to take care of household members and obligations; modal availability; and history of activity-travel engagement throughout the day up to the decision point in question. If a traveler is delayed on the network (e.g., congestion or an incident) and arrives later than expected at the destination, then the ABM will simulate subsequent activity-travel choices accordingly (recognizing the delayed arrival and the time-space prism constraints). By ensuring that activity-travel demand choices are exactly in sync with *actual* travel times experienced by the travelers on the network, the dynamic integrated platform ensures that there are no overlaps and gaps in the simulation of daily activity-travel patterns and schedules.

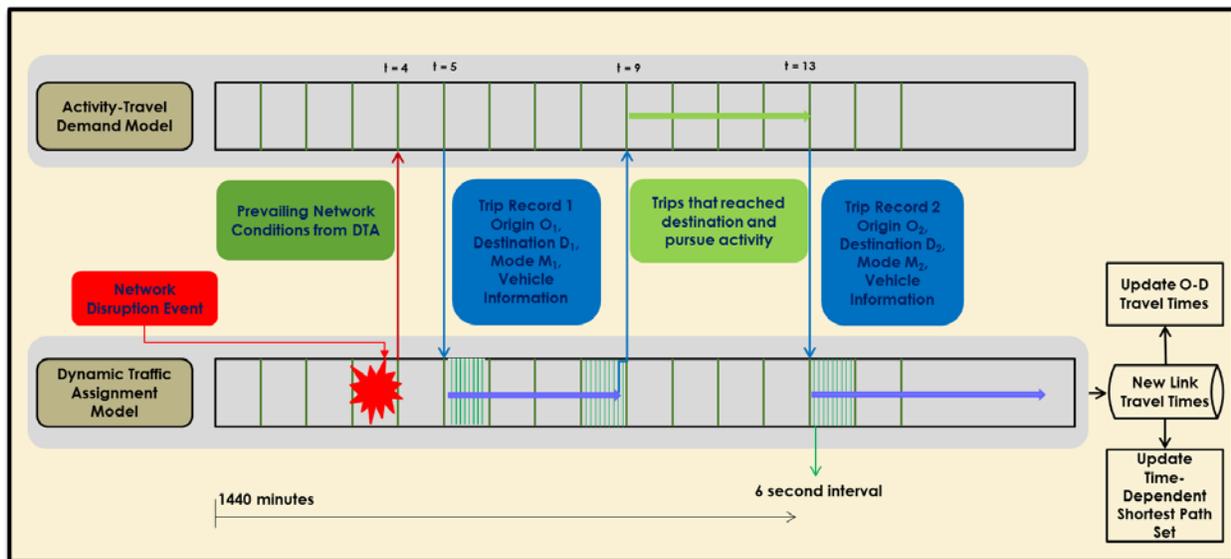


**Figure 5.1 SimTRAVEL integrated modeling framework.**

In the original SimTRAVEL integrated modeling framework, travelers do not have information about prevailing network conditions when making decisions about activity-travel schedules, routes, modes and destinations. The DTA model outputs time-dependent skim matrices at the end of each iteration. These time-dependent skim matrices are used by the ABM in the next iteration to simulate activity-travel choices along the continuous time axis. Once again, the ABM and DTA models exchange data continuously, but all routes and activity-travel decisions are based on travel times and costs output, which are saved from the prior iteration of the model run. In other words, the iterative process mimics a day-to-day learning and information updating process, where travelers experience the network on one day, update their memory regarding experienced travel times, and then make modifications to their activity-travel choices in subsequent days based on the accumulated memory of experiences. The process continues until convergence is achieved, wherein OD matrices show no appreciable change from one iteration to the next on the ABM side, and the gap function shows no further decrease on the DTA model side. The model is said to have converged when both criteria are met. This tight dynamic integration framework can be further extended to incorporate the simulation of the impacts of dynamic mobility management and real-time information strategies.

### **5.2.3 Level Two: Dynamic Integration – Only Pre-trip Information**

In the next level of dynamic model integration, the SimTRAVEL platform is enhanced to accommodate the effects of the availability of pre-trip information about prevailing network conditions. However, travelers do not have access to real-time travel information in the middle of a journey; they are able to make activity-travel choices that are network-sensitive based on information they access prior to embarking on a trip. For example, suppose there is a network disruption (planned or unplanned) that affects prevailing travel times. A traveler will be able to access pre-trip, real-time information about prevailing network conditions and make activity-travel choices (activity type, mode, destination, and accompaniment) – subject to time-space prism constraints – that take such information into account. A simplified representation of this conceptual design is shown in Figure 5.2.



**Figure 5.2 Dynamic SimTRAVEL framework – Level 2 pre-trip Information.**

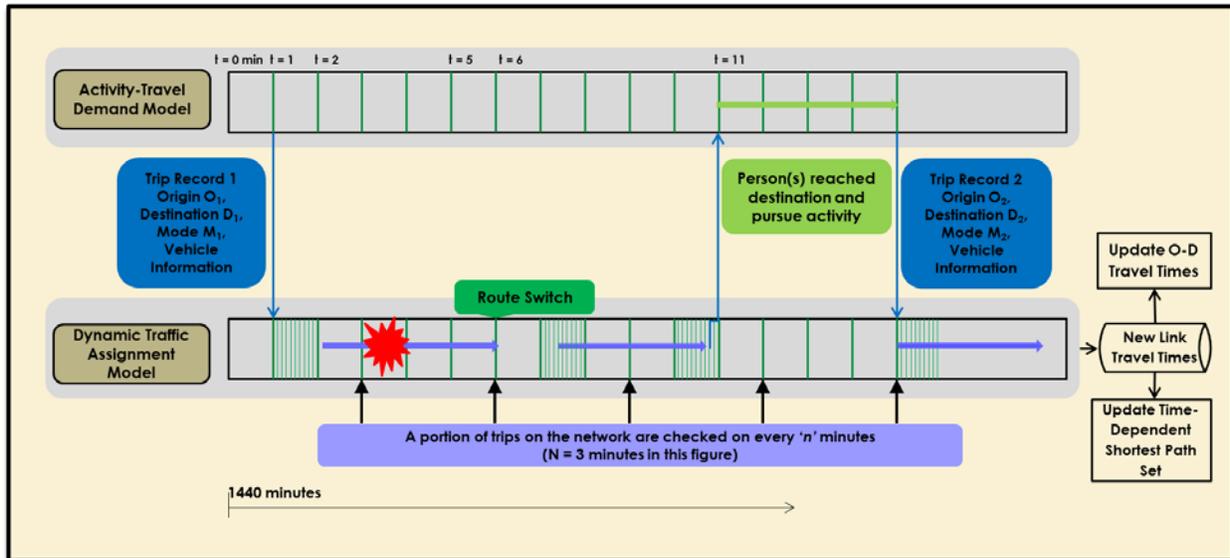
To illustrate this level of the framework, consider a network disruption event as shown in Figure 5.2. If an individual has pre-trip information, then that individual should be provided data about prevailing network conditions and travel times rather than travel time matrices from the prior iteration. Those with no pre-trip information, on the other hand, will rely on historical (prior iteration) accumulation of travel time information to make activity-travel choices. Therefore, there are different traveler segments in the population: those who have pre-trip information and use prevailing network travel time information; and those who do not have pre-trip information and use network travel time information saved from prior iterations. In this conceptual design, the DTA model outputs network travel times at regular intervals (as small as 1 minute, but a 15 minute time step may be sufficient to reflect lag in information availability). Travelers with pre-trip information consult these travel time matrices (reflecting prevailing network conditions) when making activity-travel choices. The simulation continues throughout the day, with information-equipped travelers constantly utilizing real-time information on prevailing network conditions as they execute their activity-travel patterns and schedules. Travelers with no pre-trip information rely on accumulated history of travel time information in decision-making, and may therefore encounter severe congestion and delayed arrival times as a result of the disruption. The ABM simulates their activity-travel choices at each decision point, recognizing their situation and time-space prism constraints. Presumably, those who have access to pre-trip information about prevailing network conditions can avoid the disruption effects (to the extent possible) and do not experience delays and setbacks as severe as those who do not have access to such information.

It should be noted that Figure 5.2 does not represent an iterative process wherein the model components are run repeatedly to convergence. Whether or not the model system needs to be run in an iterative process may depend on the specific network disruption context being considered. For example, if the network disruption occurs because of an unexpected accident or spill, it may be conjectured that the network will experience disequilibrium conditions where travelers experience chaos and delay. In such a context, the intent of the analysis is to determine the impacts of an unexpected network disruption on activity-travel patterns, delays, network conditions, and travel times and speeds. A single iteration of the integrated model run should provide this

information. On the other hand, suppose the disruption is a planned network event, such as a work zone that is going to be present for several months. In such a context, it may be conjectured that travelers will learn from one day to the next, and use both accumulated history of experienced travel times as well as prevailing network conditions (if they have pre-trip information) to make activity-travel choices. As individuals learn how the network is performing in the presence of the work zone, they may use updated information about network travel times to alter their activity-travel patterns and schedules; the iterative process will capture such a learning process, and analysts may be interested in determining the “new normal” into which traffic patterns will settle over a number of days. The feedback loops are removed in Figure 5.2 (as well as figures representing subsequent levels of integration) to reflect the former (unexpected) network disruption state.

#### **5.2.4 Level Three: Dynamic Integration – Pre-trip and En-route Information with Route Diversion Only**

The next level of dynamic integration adds the ability to account for the effects of real-time, en-route traveler information on route choice and route diversion that may occur in response to network dynamics. A graphic representation of this level of integration is shown in Figure 5.3. This level of the SimTRAVEL framework continues to account for the effects of pre-trip information that individuals may have about prevailing network conditions. Thus, the ABM in the SimTRAVEL framework simulates activity-travel choices in a network-responsive manner for those with pre-trip information, but continues to use accumulated history of travel time information from prior iterations for those with no information. In addition, individuals who have en-route travel information can adjust their route in response to network travel times. At regular time intervals (specified by the analyst), the DTA model will check on the status of a portion of travelers in the network who have real-time en-route travel information. The DTA model will determine whether the prevailing travel time to destination satisfies the time-space prism constraints. If time-space prism constraints are not violated, then the individual (traveler or vehicle) continues on the same path until the next check is done  $n$  minutes later. If a traveler is on a path that is disrupted or experiencing unusual congestion, then the DTA model will search for alternate, time-dependent, shortest paths based on prevailing link travel times and re-route the traveler to the best alternative. If there is no path that allows a traveler to reach the destination without violating the time-space prism constraints, then the best feasible path is chosen and a violation of the time-space prism constraints is permitted (e.g., being late for work, school, child pick-up, or a business meeting).



**Figure 5.3 Dynamic SimTRAVEL framework – Level 3 pre-trip and en-route information with route diversion only.**

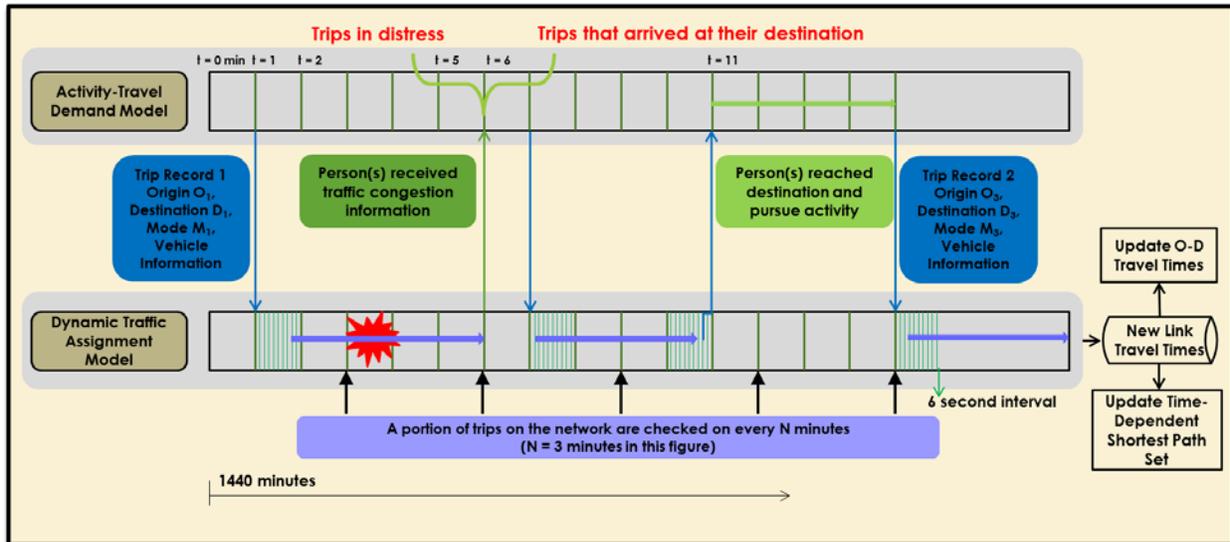
In this configuration, the traveler is not allowed to abandon or re-schedule the activity, switch to a different destination, or make any other adjustments to the activity-travel demand decisions that generated the original trip. Presumably, an individual with information is making well-informed travel choices, and then attempts to avoid the disruption using en-route traveler information that facilitates route diversion. Note that this level of model integration also supports a scenario where individuals do not access pre-trip information, but have access to en-route information (for instance, via variable message signs) to exercise route diversion. In this case, the traveler uses historical network travel time skims from prior iterations to make activity-travel choices, and then adjusts route on the network in response to en-route traveler information.

### 5.2.5 Level Four and Beyond – Pre-trip and En-route Information with Full Activity-Travel Choice Adjustments

The final level of the SimTRAVEL framework is where travelers have access to both pre-trip and en-route travel information, *and* can adjust the full complement of activity-travel choices in response to the information they access. Similar to the previous levels, individuals with pre-trip information about prevailing network conditions will rely on such information to make activity-travel choices in the ABM, as opposed to individuals with no information who use the historical accumulation of travel time matrices from prior iterations to make activity-travel choices. Individuals with en-route travel information can now adjust activity type choice, destination choice, mode choice, and accompaniment choice *in addition to* route choice (which was accommodated in the prior level).

Just like the prior level, individuals on the network with en-route travel information are checked every  $n$  minutes to determine if the traveler should be diverted to an alternate path. In general, it is assumed that individuals will first try to simply change path and find an alternative route to their intended destination and activity to minimize disruptions to the activity schedule. In previous work, Ye and Pendyala (2007) found that individuals first exercise choices that are least constrained and minimize disruptions before modifying more constrained choices. If the traveler

should be diverted (because the current path is not tenable given time-space prism constraints), then the best feasible alternate path is identified by the DTA model with a view to facilitate route switch. If, however, the best feasible alternative path is also deemed unacceptable (because a substantial violation of the time-space prism constraints would occur), then the individual is tagged as a traveler in *distress*. In each one-minute time step of the simulation, the DTA model now returns to the ABM two sets of trips, or, the set of trips that have routinely arrived at their destination and the set of trips that are in distress. The latter set of trips correspond to travelers who are stuck on the network and have no feasible alternative path that can get them to their intended destination on time (without violating time-space prism constraints).



**Figure 5.4 Dynamic SimTRAVEL Framework – Level 4 Pre-trip and En-route Information with Full Activity-Travel Choice Adjustments**

For the travelers that are tagged as being in distress, the ABM will determine if alterations to their travel choices are in order. Travelers may choose to switch to an alternate destination if the activity in question is discretionary in nature and has flexibility with respect to location. After choosing a new destination (based on prevailing network conditions), the trip is sent back to the DTA model (in the next one-minute time step) for routing and simulation on the network. If the activity in question does not have location flexibility or there is no alternate destination that meets time-space prism constraints, then the individual is assumed to switch modes or abandon the activity altogether. Mode switching is unlikely to occur mid-journey, so this particular behavioral adaptation choice is suppressed in the current framework. However, with the increasing popularity of ride-sourcing and vehicle-sharing services, it is plausible that mid-journey mode-switching will become more feasible. In addition, in locations where park-n-ride lots are available for fixed and dedicated guideway transit services, an individual may divert to the nearest park-n-ride facility and hop on the train or dedicated busway for the remainder of the trip. Finally, an individual may abandon the activity in question and proceed to the next fixed activity defined by the time-space prism constraints, switch to a different (discretionary) activity type which would entail choosing a new destination, or simply return home (e.g., work or study from home). After the ABM simulates the choice adjustment made by the traveler, the information is sent to the DTA model and the DTA model routes and simulates the traveler on the network, starting from the point where the traveler

was originally tagged as being in distress. An abandoned activity may be pursued later in the day depending on time-space prism constraints. Because the activity-based model recognizes history dependency in simulating activity-travel choices, an activity that was abandoned at one point in the day may be performed later in the day (because the model recognizes that the activity was not performed earlier).

It is plausible to envision levels of model integration that go beyond the ability to change route and activity-travel demand choices in response to real-time travel information facilitated by connectivity, communications and technology. A further level of the model integration framework would accommodate re-allocation of tasks and activities among household members. Therefore, when a member of the household abandons a certain activity, the activity may be re-assigned to another household member who has an open time-space prism that could accommodate the activity. However, this reassignment of tasks within the household can only occur for specific types of activities and for specific individuals (for example, a work activity cannot be reassigned, and small children cannot go to the bank). The incorporation of a full-fledged array of intra-household interactions with rule-based constraints remains a fruitful area of further research. Finally, the ultimate level of model integration would incorporate interactions across households or colleagues, where some tasks are reassigned to non-household-members.

## 5.3 Case Study

This section presents results of a case study application that demonstrates the efficacy of the framework, and the detailed levels of integration that capture the effects of dynamic mobility management strategies and real-time traveler information on activity-travel patterns and network dynamics. In implementing the framework, it is assumed that any and all information about network conditions arising from the introduction of dynamic mobility management strategies or the occurrence of a network disruption can be translated into equivalent time and cost units, or mobility options (alternative routes, destinations, and modes). Thus, if there are dynamic pricing strategies, they will result in changes in travel times and costs. A new intelligent traffic signal timing algorithm will affect travel times. A dynamic parking management system will impact terminal times or parking costs. Therefore, each DMA or ATDM strategy that may be implemented to improve system performance is represented in terms of the time and cost differences that the strategy engenders for different travelers. With increasing connectivity in vehicles, it is likely that travelers will avail of information to learn about alternative destination or modal options (e.g., ride-sourcing) that they would not have considered otherwise. In other words, individuals may be able to take more risks and try new possibilities when armed with information.

All travelers in the synthetic population are tagged with an information-type flag: individuals with no information, pre-trip information only, en-route information only, and both pre-trip and en-route information. These four information groups will exercise activity-travel decisions and route choices based on different sources of information. The DTA model will provide information about network conditions, both historical (saved from prior iterations) and prevailing (based on current traffic), and users will use the appropriate information depending on their information type status.

Two specific model systems were used for this illustrative demonstration of the SimTRAVEL framework. The activity-based travel demand model system is openAMOS, an open-source activity mobility simulator (SimTRAVEL, 2016). This ABM provides a platform for

microsimulating activity-travel patterns of individuals in a synthetic population, while specifically accounting for time-space prism constraints, coupling constraints, and modal availability constraints. Unlike many tour-based models that have been implemented in practice (e.g., Outwater and Charlton, 2006), openAMOS is a continuous-time evolutionary activity-based model, where the activity-travel pattern evolves over the course of a day as individuals make decisions regarding activity engagement at various decision-points and subject to a variety of constraints.

In the openAMOS modeling framework, time-space prism constraints are modeled first to develop an activity skeleton for individuals in a household. Then, beginning at the start of the day, individuals pursue various secondary, flexible and discretionary activities within available time-space prisms, while adhering to various constraints that the skeleton imposes. This evolutionary activity-based model design is ideal for tight coupling with a DTA model, where the ABM and DTA models exchange information about trips and travelers on a continuous (one-minute time step) basis. The DTA model used in this illustrative exercise is DTALite (Zhou and Taylor, 2014), a dynamic traffic assignment model capable of routing and simulating trips through a network while recognizing the time-dependent nature of shortest paths. The model system is able to output information about prevailing and historical network conditions (e.g., travel times) in the form of time-dependent matrices that travelers are assumed to use when making activity-travel choices. Although this exercise utilizes these specific model components, the SimTRAVEL framework may be implemented using any ABM and DTA models that can be tightly coupled together and engineered to exchange data at a fine temporal resolution. By tightly integrating the two models so that specific travelers can be tracked through both the ABM and DTA models, the SimTRAVEL framework constitutes a true agent-based simulation model system.

In this illustrative example, the first three levels of the SimTRAVEL framework are demonstrated using the small Sioux Falls network that includes 24 traffic analysis zones (TAZs) and 76 network links (at 60 mph free-flow speed). To ensure that significant traffic volumes and congestion levels would be realized on the network, a synthetic population of 24,853 persons residing in 8,259 households was generated using the population synthesizer embedded within openAMOS, based on assumed control distributions on various socio-economic variables (derived from census data). The SimTRAVEL framework was applied to a number of network scenarios to demonstrate the ability of the framework to reflect the impacts of network dynamics and traveler information on activity-travel patterns and system outcomes. The scenarios implemented are as follows:

- Baseline Scenario: No network disruption
- Scenario 1: Network disruption with no traveler information (SimTRAVEL Level 1)
- Scenario 2: Network disruption with pre-trip traveler information only (SimTRAVEL Level 2)
  - Individuals in 50 percent of households have pre-trip information; others have no information
- Scenario 3: Network disruption with pre-trip and en-route traveler information allowing only route changes in the middle of a journey (SimTRAVEL Level 3)
  - Individuals in 50 percent of households have pre-trip and en-route information; others have absolutely no information

The only scenario that was not implemented is that where individuals are tagged as being in distress if they cannot identify a feasible alternative path, and are then sent to the activity-based model to determine if changes in activity-travel demand choices are warranted. In the SimTRAVEL framework, it is possible to see how individual travelers behave in the presence of information and when confronted by a network disruption. Such disaggregate output can be very valuable in identifying groups that may be disproportionately adversely affected by a certain disruption. The disaggregate individual-level activity-travel pattern outputs can also help identify the types of adjustments that individuals are making in response to network disruptions and traveler information provision. The network disruption involved constraining three of the 76 links during key periods of the day. Heavily traveled links were selected for disruption. One link was assumed to have a speed of 10 mph during 8 a.m. to 9 a.m., another link was assumed to have this speed between 12 noon and 1 p.m., and a third link was assumed to experience 10 mph speed during 5 p.m. to 6 p.m. Appropriate capacity restrictions were also imposed on the links during these time periods. DTALite identified alternate paths for individuals who had en-route travel information and provided skim matrices with information about prevailing network conditions, so that individuals with pre-trip information could make informed activity-travel demand choices.

The first column shows travel characteristics for the baseline case where no disruption is introduced in the network. Average trip duration is 9.26 minutes per trip and average trip length is 8.23 miles. Given the high free-flow speed of the links in the network, these statistics are consistent with expectations. In scenario 1, where a severe disruption is introduced, average trip duration increases as expected. In the absence of any information, and in view of the disruption, DTALite routes individuals on somewhat more circuitous paths leading to an increase in trip length as well. The trip length and trip duration are both higher during the disruption period when compared with the overall daily average. This finding is consistent with expectations, as travelers are likely to experience disruptions and delays in the absence of any information.

Results of Scenario 2 clearly demonstrate the benefit of providing information about prevailing network conditions to travelers. Armed with real-time, pre-trip information, travelers can make more informed decisions with respect to activity-travel schedules and location choices. As a result, the average trip duration and average trip length for travelers with pre-trip information is considerably lower than corresponding values for individuals without any traveler information. Travelers with pre-trip information are better off during the disruption period as well. Compared to the case where nobody had any information, travelers with no information are better off in the scenario when individuals in one-half of the households are equipped with information. As the individuals with information make more informed decisions, travelers with no information benefit as well due to reduced congestion and delays.

Finally, results of Scenario 3 demonstrate the case where individuals in one-half of the households have pre-trip and en-route travel information about prevailing network conditions. It is interesting to note that those with information are better off than those without information (similar to Scenario 2). However, these individuals (with both pre-trip and en-route information) experience a degradation in average trip duration and trip length in comparison to the case where they had only access to pre-trip information. The differences are modest, but very systematic, suggesting that these differences are not attributable to stochastic randomness inherent to microsimulation model systems.

**Table 5.1 Results of Illustrative Demonstration of SimTRAVEL Integrated Modeling Framework**

Travel Characteristic	Baseline: No disruption	Scenario 1: Disruption  100% No Information	Scenario 2: Disruption		Scenario 3: Disruption	
			~50% No Information	~50% Pre-trip Information	50% No Information	50% Pre-trip + En-route Information
Total trips	115683	115625	59216	56393	59060	56404
Total number of persons	24853	24853	12729	12124	12729	12124
Trip rate	4.655	4.652	4.652	4.651	4.640	4.652
Total travel time	1071410	1079195	553535	523511	552639	525363
Total travel distance	952195.75	959231.75	491991.25	464131.75	491307.75	466323.25
Average travel time (min)	9.26	9.33	9.35	9.28	9.36	9.31
Average travel distance (mile)	8.23	8.30	8.31	8.23	8.32	8.27
<b>During disruption time interval</b>						
Total travel time	232011	236504	204531	202808	205787	203622
Total travel distance	205343.25	209824	181202.25	179587	182555.5	180569.25
Average travel time (min)	9.26	9.40	9.33	9.30	9.37	9.34
Average travel distance (mile)	8.19	8.34	8.26	8.23	8.31	8.28

A further examination of the phenomena at play suggested that re-routing travelers resulted in longer trip lengths, as DTALite attempted to re-route travelers while mid-journey, resulting in more circuitous paths and increased trip durations. In the case where travelers were armed only with pre-trip information, they made informed decisions and stayed on the shortest path that had been designated at the time that they embarked on the trip. In the case where travelers adjust their route during the course of the journey, two factors are found to contribute to increased travel times and distances. First, travelers switch to a more circuitous path in an attempt to divert and abide by time-space prism constraints. Second, the alternative paths themselves experience greater congestion as individuals are diverted by DTALite. In other words, it appears that travelers are being switched to alternative paths even when the path will eventually prove to be detrimental (as more travelers switch to the alternative path).

To address this issue, multiple iterations internal to DTALite need to be executed so that travelers are diverted to alternative paths in a more optimal fashion. In this exercise, multiple iterations of DTALite were not performed, resulting in a diversion pattern that was suboptimal. The enumerated alternative paths were circuitous and became congested due to excessive diversion, resulting in greater delays for those with both pre-trip and en-route information. This illustrative demonstration shows that the SimTRAVEL framework is able to capture dynamics of activity-travel choices and network attributes, and reflect the impacts of real-time traveler information and dynamic mobility management strategies.

## 5.4 Conclusions

The advent of the information era has ushered in a new suite of dynamic mobility management strategies for proactively managing travel demand, freight and passenger travel choices, and network performance. The suite of strategies included under the Dynamic Mobility Applications (DMA) and Active Transportation and Demand Management (ATDM) umbrella include measures that affect both travel demand choices (activity generation, destination choice, and mode choice, for example) as well as network attributes (e.g., in terms of travel time, cost, and intersection delay). As the profession seeks to integrate planning and operations through the deployment of proactive dynamic mobility management strategies, there is a growing need for integrated model systems that are capable of adequately reflecting behavioral dynamics and network dynamics, and the continuous interaction that exists between them.

This chapter presents various levels of a dynamic integrated modeling framework where an activity-based travel microsimulation model is tightly coupled with a dynamic traffic assignment model. In the proposed framework, the two model components are able to communicate with one another along the continuous time axis, so that trips or tours generated by the activity-based model are routed and simulated on the network as they happen, and activity-travel demand choices made by travelers recognize actual experienced network conditions as travelers go about their daily lives. Through such a tight coupling, it is possible for the framework to simulate the impacts of real-time information provision and operational strategies aimed at influencing traveler behavior and optimizing network performance.

Various levels of the integrated modeling framework, referred to as SimTRAVEL, are described in detail. The increasingly detailed and capable levels of the framework accommodate scenarios where a portion of the travelers are equipped with pre-trip information or en-route travel information (or both) about prevailing network conditions at all times. When travelers are armed with such information, they may switch routes or adjust travel demand choices in response to network disruptions and traveler recommendations offered by various connected platforms. In the most detailed level of integration, travelers will be able to adjust their route or adjust activity-travel demand choices mid- journey in response to network conditions.

The efficacy of the integrated model system is demonstrated through an illustrative case study of the small Sioux Falls network, albeit with a sizeable population to realize network congestion. It is found that the model is capable of simulating network performance under alternative information scenarios and provide results that are consistent with expectations with improvements in performance seen in the information provision scenarios. The scenario where travelers had both pre-trip and en-route information showed that they experienced a degradation in performance (travel time and trip length), compared to the scenario where they had only pre-trip information. It appears that travelers are being diverted to longer, alternative paths in the en-route information scenario, resulting in congestion on the alternative paths; this calls for the execution of route switching patterns in the network model, such that diversion is done in a more user optimal fashion. This would require that the dynamic traffic assignment model be run through multiple iterations until an equilibrium (in terms of route switching) is achieved.

The efficacy of running such large scale integrated model systems for large networks is yet to be demonstrated and the validity of the results furnished by such large scale integrated models (in response to exogenously specified scenarios) remains largely unconfirmed. When integrated model systems are applied in the context of emerging travel demand management and technology-based scenarios, they generally provide indications that are consistent with the patterns of behavior found in the survey data sets that informed the model specifications. While this is reasonable as a point of departure, there is still a severe paucity of information on how individuals actually respond to dynamic mobility management strategies when implemented in the field. The predictions from SimTRAVEL or any integrated transport model system need to be validated against real-world data before they can be applied for “planning for operations.” In addition, references such as the TCRP Report 95 Collection (TCRP, 2016) would provide data about traveler response to system changes that can be used to validate predictions from novel integrated model systems.

## 6. OUTREACH AND TECHNOLOGY

This chapter presents the outreach plan and technology transfer activities proposed as part of this SHRP2 C10 project. The main objective of the outreach and technology transfer activities is to provide effective training of agency users for the practice-ready integrated models (InSITE ABM-DTALite and SILK AgBM-DTALite) and related simulation tools (DTALite and InSITE ABM). The secondary objective is the development of workshops with long-term value to the partner agencies: Maryland State Highway Administration (MD SHA) and the Baltimore Metropolitan Council (BMC).

The planned activities are grouped into four categories of workshops based on the simulation tools as follows and are described subsequently:

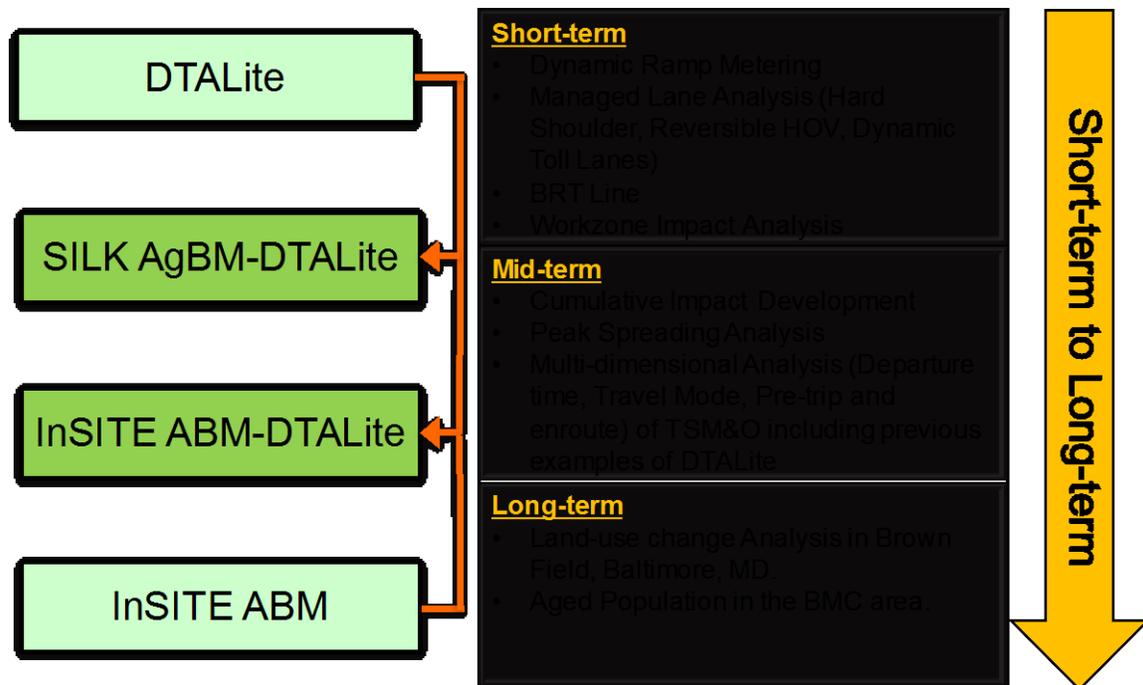
- (1) **DTALite**
  - a. DTA Simulation for planned applications
  - b. Calibration and Validation of DTA Models.
- (2) **InSITE ABM**
  - a. Simulation for planned applications
- (3) **SILK AgBM-DTALite**
  - a. Simulation for planned applications
- (4) **InSITE ABM-DTALite**
  - a. Simulation for planned applications

DTALite is the common link between the integrated models, and thus will receive more attention through two workshops. The first workshop will focus on simulation for planned applications similar to the other workshops for the other models, with a second workshop focusing exclusively on calibration and validation of DTA models. Workshops for the InSITE ABM and each of the integrated models are also planned with a focus on simulation for planned applications only; however, there will be discussions about validation of these models that build upon the DTALite workshops.

### 6.1 Planned Workshops

#### 6.1.1 Planned Activities

The planned workshops are divided into four groups as presented in Figure 6.1: DTALite, InSITE ABM, and the integrated models (SILK AgBM-DTALite and InSITE ABM-DTALite). The order of the workshops in terms of the timeline is: DTALite, SILK AgBM-DTALite, InSITE ABM, and InSITE ABM-DTALite. The actual dates and times of the workshops will be decided in coordination with the partner agencies.



**Figure 6.1 Groups of workshops**

### **DTALite**

DTALite, an open-source light-weight, mesoscopic DTA simulation package, in conjunction with the Network eXplorer for Traffic Analysis (NeXTA) graphic user interface (GUI), has been developed to provide transportation planners, engineers and researchers with a theoretically rigorous and computationally efficient traffic network modeling tool.

DTALite is the common intersection of the two integrated models; therefore, the first workshops will focus on DTALite. The main activities of these workshops are briefly described in Table 6.1. The goal is to prepare the users for the following tasks:

- Preparation of inputs for applications
- Running the model and understanding the output
- Calibration and validation of the DTA models

The main applications of interest are transportation system management and operations (TSM&O), and several case studies outlined in Figure 6.1 will be prepared as examples for the users.

### **InSITE ABM**

InSITE is an activity-based model system composed of interconnected, discrete choice models representing choices at distinct dimensions (e.g., travel mode, destination) that focus on decisions related to daily activity and mobility for a typical weekday. InSITE adopts the day activity-schedule approach, where a day activity schedule is defined through the concepts of activity pattern and activity schedule. The activity pattern defines the participation in activities as primary and secondary. Primary activities are the anchors of a

tour. For example, a home-to-work trip and a work-to-home trip represent a tour, with work as primary activity. Secondary activities are intermediate stops within a particular tour (i.e., stopping for shopping at the work-to-home half tour). The activity schedule adds detailed information to the activity pattern about tours, such as timing, travel mode, destination of primary activity, and also the stops for secondary activities within the tours.

As part of their activities with the Baltimore Metropolitan Council (BMC), Cambridge Systematics has already conducted several workshops to instruct and transfer InSITE ABM to BMC. They have also prepared several documents to support BMC in learning how to use this model. The team plans to leverage this previous work to hold a workshop about using the InSITE ABM for applications.

The main activities of this workshop are briefly described in Table 6.1. The goal is to prepare the users for the following tasks:

- Preparation of inputs for applications
- Preparing configuration files for applications
- Running the model and understanding the output

The main applications of interest are long-term transportation planning projects; several case studies outlined in Figure 6.1 will be prepared as examples for the users.

An important challenge of the InSITE ABM is understanding the wealth of output of this model, so the team will also work on preparing visualization tools to help the agencies gain a good understanding of how to use the output for their own analyses.

### DTALite and SILK AgBM-DTALite Case Studies

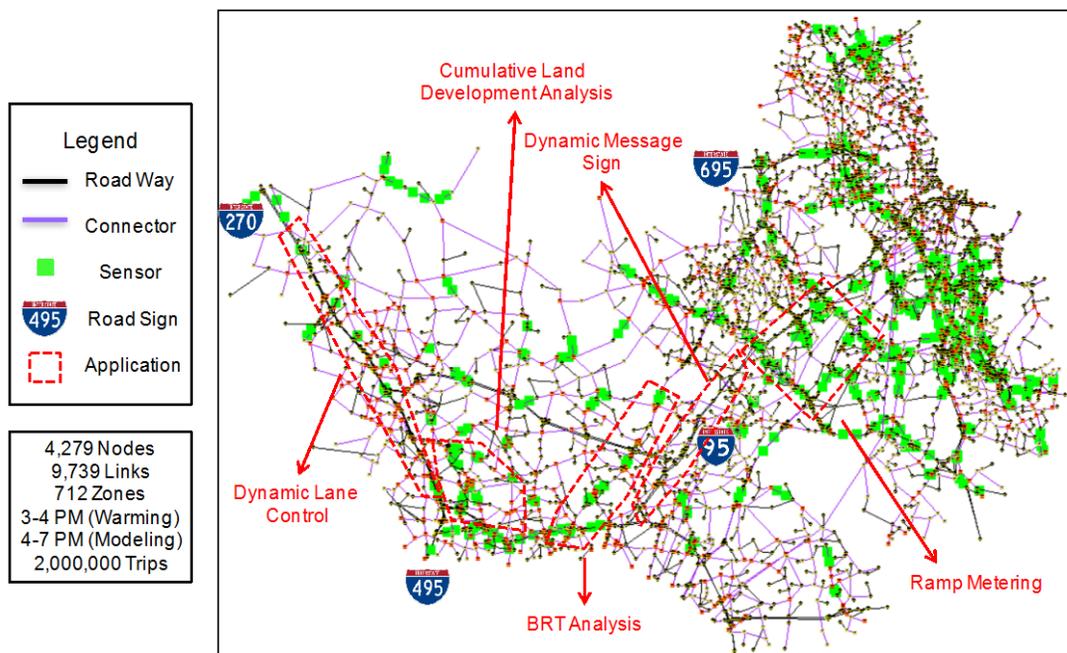


Figure 6.2 Coverage of DTALite and SILK AgBM-DTALite

### **SILK AgBM-DTALite**

SILK-AgBM is an agent-based travel behavioral model system that emphasizes the role of Searching, Information, Learning and Knowledge (SILK). It can be efficiently interfaced with a dynamic traffic assignment model. SILK AgBM does not only rely on rules extracted from data mining to explain behavior, but also theorizes a multidimensional knowledge updating, search start/stop criteria and search/decision heuristics. These behavioral modules are formulated, and empirically modeled and integrated, in a unified and coherent system. Procedural and dynamic behavioral rules are estimated to represent rich behavior in travel mode, departure time, route and en-route diversion choices. They can also be enhanced to potentially accommodate real-time decision making on the part of the traveler and the driver response to network conditions/congestion and real-time traveler information (in addition to a host of other dynamic mobility applications and active traffic and demand management strategies). SILK-AgBM is integrated with a dynamic traffic simulation tool named DTALite. Among a number of DTA simulators and software packages, DTALite is an open-source mesoscopic (i.e., a resolution level that lies between macroscopic and microscopic) DTA simulation package.

The SILK AgBM-DTALite model (and also the DTALite model) has a wide coverage, as shown in Figure 6.2, and a significant number of completed applications as presented in Figure 6.1 and 6.2.

The team will follow an incremental approach to transfer the SILK AgBM-DTALite model. The team will start the applications with one behavioral dimension (e.g., departure time choice), and the application with more than one behavioral dimension (travel mode, pre-trip and en-route choices) will be transferred later within the workshop.

The main activities of this workshop are briefly described in Table 6.1. The goal is to prepare the users for the following tasks:

- Preparation of inputs for applications
- Configuration of the model to use multiple behavioral dimensions
- Running the model and understanding the output

The main applications of interest are transportation system management and operations (TSM&O); several case studies outlined in Figure 6.1 will be prepared as examples for the users.

### **InSITE ABM-DTALite**

The integration pursued between InSITE and DTALite falls in the sequential integration paradigm. In this paradigm, the travel demand model (InSITE) and the DTA model (DTALite) are run in iterations, with a single “big loop” consisting of an iteration of the travel demand model followed by a run of the DTA model. The travel demand information (e.g., agents with their activities and travel decisions, as well as characteristics) is passed to the DTA model within each big loop; travel time information resulting from the dynamic assignment is passed back to the travel demand model for use as input (e.g., update activity

pattern choices) in the next big loop. Note that DTA models themselves are run iteratively, and so there are several “small loops” of the DTA run within each “big loop.” A number of “big loops” are run until a measure of convergence is achieved, where the change in travel times (or some other measure) from one big loop to the next is within a specific tolerance. The data exchange between InSITE and DTALite is such that an agent is preserved during the exchange of data from InSITE to DTALite and vice versa. Thus, the integrated InSITE-DTALite model is fully disaggregated.

The InSITE ABM-DTALite is a straightforward integration of InSITE ABM and DTALite; its outputs are the same outputs of InSITE ABM and DTALite, with the addition of a convergence measure output for the integrated model. Therefore, we will leverage the previous workshop materials of DTALite and InSITE ABM.

The main activities of the workshop for this integrated model are briefly described in Table 6.1. The goal is to prepare the users for the following tasks:

- Preparation of inputs for applications
- Preparing configuration files for applications
- Understanding the convergence measure for the integrated model
- Running the model and understanding the output

The main applications of interest are long-term transportation planning and similar applications to the InSITE ABM, but focusing on applications where DTA will provide additional information unavailable through a static traffic assignment model.

**TABLE 6.1 Summary of the Activities for the Workshops**

<b>DTALite</b> (2 Workshops)	<ul style="list-style-type: none"> <li>• Preparing Inputs - Using NeXTA</li> <li>• TSM&amp;O Case Studies – Model Runs &amp; Output</li> <li>• Calibration &amp; Validation</li> </ul>
<b>SILK AgBM-DTALite</b> (1 Workshop)	<ul style="list-style-type: none"> <li>• Preparing Inputs for SILK AgBM</li> <li>• TSM&amp;O and additional Case Studies – Model Runs &amp; Output</li> </ul>
<b>InSITE ABM</b> (1 Workshop)	<ul style="list-style-type: none"> <li>• Preparing Inputs for InSITE ABM</li> <li>• Preparation of Configuration Files for Scenarios</li> <li>• Understanding of the output of the ABM</li> </ul>
<b>InSITE ABM-DTALite</b> (1 Workshop)	<ul style="list-style-type: none"> <li>• Linking the InSITE ABM output to the DTALite output for analysis</li> <li>• Model Runs with the integrated model</li> </ul>

### 6.1.2 Format of the Planned Workshops

The format of the planned workshops is as follows:

- **Recorded training sessions:** all the workshops will be professionally recorded by the audio-visual facilities of the University of Maryland, College Park. The goal is to provide video training materials that will offer continuous staff training for the agencies.
- **Hands-on training with a selected group of agency participants:** all of the workshops will be fully hand-on, meaning that the participants will learn how to use the tools with the help of an instructor during the workshops. In addition, the number of agency participants will be limited for more effective training.
- **Incremental training approach:** the workshops will build the knowledge of the participants by starting from the most important component of integrated models such as DTALite, and also by starting from the most straightforward case studies.
- **Software and case Studies provided in advance:** the MITAMS team will transfer the software materials and case studies to the agencies prior to the start of the workshops.
- **On-site support to flatten the learning curve:** the MITAMS team will have one staff member on-site once a week at each agency to help support the newly trained staff in using the simulation tools.

## 6.2 Supporting Documents

The MITAMS team will provide a User's Guide for each of the simulation tools to provide step-by-step guidance to the agency staff. The goal is to provide a go-to resource that, together with the recorded video material, will support the agencies in integrating these simulation tools as part of their internal procedures.

In Table 6.2, a summary of the supporting documents for the workshops is presented.

**TABLE 6.2. Supporting Documents for the Simulation Tools**

<p><b>DTALite</b></p>	<ul style="list-style-type: none"> <li>• Leveraging and adapting DTALite User' Guide</li> <li>• Adding specific support for the case studies</li> </ul>
<p><b>SILK AgBM-DTALite</b></p>	<ul style="list-style-type: none"> <li>• User's Guide for the AgBM component and exchanges with DTALite</li> <li>• Detailed explanation for the case studies</li> </ul>
<p><b>InSITE ABM</b></p>	<ul style="list-style-type: none"> <li>• Leveraging the InSITE ABM's User Guide</li> <li>• Providing additional support for understanding the case studies.</li> </ul>
<p><b>InSITE ABM-DTALite</b></p>	<ul style="list-style-type: none"> <li>• User's Guide on running the integrated model</li> <li>• Detailed explanation for the case studies.</li> </ul>

## 7. Conclusions and Future Research

In summary, the Maryland team has developed two integrated models: BMC InSITE-DTALite and SILK AgBM-DTALite. The first is the result of integrating a dynamic traffic assignment (DTA) tool based on an existing DTALite model that covers the whole State of Maryland, and a state-of-the-art activity-based model (ABM), InSITE, (developed by Cambridge Systematics for the BMC metropolitan area). This integrated model covers most of the urbanized areas in Maryland and it will support long-range planning, transportation conformity analysis, investment programming, project prioritization, equity analysis, and corridor/project planning at both SHA and Baltimore- Washington MPOs. The second is an agent-based microsimulation travel demand model (AgBM), named SILK (for its emphasis on Search Information, Learning, and Knowledge in the travel decision-making process). SILK is funded by the FHWA Exploratory Advanced Research Program and developed by the University of Maryland, for TSM&O, active traffic management, integrated corridor management, and pricing analysis along selected corridors in Maryland. This subarea/corridor AgBM-DTA model will be integrated into the Reliability Roadmap and TSM&O Strategic Plan for sustainable future applications at SHA, MPOs, and MPO member jurisdictions.

### **BMC InSITE-DTALite**

This integrated model was developed as part of the MITAMS C10 project and has close ties with two agencies: the Maryland State Highway Administration (MD SHA) and the Baltimore Metropolitan Council (BMC). This integrated model is composed of the InSITE Activity-based Model (ABM) system, and the mesoscopic traffic simulator, DTALite. InSITE is an activity-based model system composed of interconnected, discrete choice models representing choices at distinct dimensions (e.g., travel mode, destination) that focus on decisions related to daily activity and mobility for a typical weekday. InSITE adopts the day activity-schedule approach, where a day activity schedule is defined through the concepts of activity pattern and activity schedule. The activity pattern defines the participation in activities as primary and secondary. DTALite is an open-source, lightweight, mesoscopic dynamic traffic assignment (DTA) simulation package that interoperates with the Network eXplorer for Traffic Analysis (NeXTA) graphic user interface (GUI). It has been developed to provide transportation planners, engineers and researchers with a theoretically rigorous and computationally efficient traffic network modeling tool. DTALite adopts multiprocessor programming to exploit the parallel computing power available on both laptops and desktops.

The primary objective of the model integration is to produce a working model of the Baltimore region that is disaggregate for auto travel on both the demand and supply sides. This model is intended to be “application ready,” meaning that it has been validated and includes all functionality required for typical planning analyses. Therefore, the integration pursued between InSITE and DTALite falls in the sequential integration paradigm. In this paradigm, the travel demand model (InSITE) and the DTA model (DTALite) are run in iterations, with a single “big loop” consisting of an iteration of the travel demand model, followed by a run of the DTA model. The travel demand information (e.g., agents with

their activities and travel decisions, as well as characteristics) is passed to the DTA model within each big loop, and travel time information resulting from the dynamic assignment is passed back to the travel demand model for use as input (e.g., update activity pattern choices) in the next big loop. Note that DTA models themselves are run iteratively, and so there are several “small loops” of the DTA run within each “big loop.” A number of “big loops” are run until a measure of convergence is achieved, where the change in travel times (or some other measure) from one big loop to the next is within a specific tolerance. The data exchange between InSITE and DTALite is such that an agent is preserved during the exchange of data from InSITE to DTALite and vice versa. Thus, the integrated InSITE-DTALite model is fully disaggregated.

## **Key Innovations**

### **Use of Simulated Values of Time:**

DTALite receives the agent-specific vector of VOTs from InSITE, in addition to other additional inputs. DTALite incorporates the vector of VOTs of the agent according to the activity purpose to be simulated, and to its respective schedule decisions (i.e., travel mode, origin and destination, timing decisions and others). Moreover, once the VOTs are assigned to the activities of an agent, the VOTs are used as input in the route choice model. The model is composed of path deterministic utility functions (i.e., each path available to the agent has a generalized travel time function), and these functions receive as input the assigned VOT from the vector of VOTs of an agent. The path deterministic utility functions determine the path choice. Furthermore, the route choice model adds the path decisions to each of the activities and their respective schedule decisions of each agent. This is done as the simulation progresses, and as the day activity schedule (i.e., list of activities and their respective schedule decisions) is realized for each agent. The collection (path choices by all the agents) of the path deterministic utility functions is used to calculate the deterministic dynamic user equilibrium.

### **Intra-Household Interactions:**

InSITE allows the explicit modeling of the intra-household interactions through the day activity pattern model, joint tour generation model, and the joint tour models for destination, time of day, and travel mode choices. The day activity pattern model accounts for the interactions by applying the model strategically, so that the day pattern outcomes of a synthetic agent are conditional on the day activity pattern outcomes of previously simulated household members. In addition, the outcomes of previously simulated household members appear as explanatory variables for other household members. For example, children are simulated first as they are the most travel-dependent individuals.

## **Key Challenges**

### **Computationally expensive:**

The integrated model takes about 1.5 to 1.8 days (InSITE ABM represents about 60 percent of the load) per big loop iteration. This means that 12 big loop iterations require about 18 days in the best of circumstances, and that simultaneous calibration of demand and supply is infeasible due to the runtime. Also, validation of the integrated model was delayed due to the long runtimes. Moving forward, running the model will require redeveloping the InSITE ABM and adjusting DTALite to reduce the overall runtime of the integrated model.

### **Convergence Paradigm:**

A straightforward convergence paradigm was implemented based on behavioral assumptions of the travelers. This convergence paradigm may be updated in order to consider more realistic behavioral assumptions.

### **SILK AgBM-DTALite**

A typical activity-based model plans traveler activity by assuming that travelers make rational choices based on perfect information. That model ceases whenever the assignment reaches equilibrium or the maximum number of iterations. Satisficing behavior rules are often used in the computational process models. However, which alternatives travelers will consider and when they become satisfied are controversial. Another limitation of existing approaches is the computational efficiency and convergence issue. The calculation of dynamic user equilibrium takes many iterations and is difficult to achieve in a DTA model, especially when the size of the network is large. Adding a set of choice models onto DTA models will further increase its complexity.

### **Key Innovations**

Fully recognizing these limitations, the team is motivated to develop an innovative integrated modeling system that can address these issues. In this study, an agent-based travel behavioral model, SILK-AgBM, is integrated with a dynamic traffic simulation model named DTALite. The innovations are three-fold:

- (1) SILK-AgBM behaviorally theorizes multidimensional learning, information, knowledge updating and search start/stop criteria, to explain travelers' multidimensional travel behavior. These components are formulated or empirically modeled and integrated in a unified and coherent approach.
- (2) The integrated model employs behavioral user equilibrium (BUE) as an alternative to the conventional user equilibrium. The search start/stop criterion are defined in a way that the modeling process should eventually lead to a steady state where every traveler in the system will stop seeking behavioral changes. This BUE definition is potentially useful in large-scale model implementations because of its multidimensional behavioral foundation and its guaranteed convergence.
- (3) AgBM-DTALite reduces the computational burden of running large-scale integrated models. The production rules in AgBM and the queue-based traffic simulation models in DTALite are computationally inexpensive. BUE also ensures a faster model convergence. In addition, the integrated model adopts a new software architecture and algorithm, design to facilitate the most efficient use of emergent parallel processing techniques.

### **Key Challenges**

#### **Extending Multi-Dimensional Behavior:**

Long-term travel behavior is not represented in the SILK AgBM-DTALite; it may be a venue for future research in order to enrich the behavioral representation of this integrated model.

## 8. REFERENCE

- Amini, N., Aydos, C., Wijayarathna, K., Grzybowska, H., & Waller, T. (2015). A network wide evaluation of the SCATS ramp metering system using microsimulation. In Australasian Transport Research Forum (ATRF), 37th, 2015, Sydney, New South Wales, Australia.
- Arentze, T. & Timmermans, H. (2003). Modeling learning and adaptation processes in activity-travel choice - A framework and numerical experiment. *Transportation*, 30(1), 37-62.
- Arentze, T. & Timmermans, H. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B*. 38(7), 613-633.
- Arentze, T. and H. Timmermans (2002). Albatross: a learning based transportation oriented simulation system. TU, Faculteit Bouwkunde.
- Arentze, T. and H. Timmermans (2003). Modeling learning and adaptation processes in activitytravel choice a framework and numerical experiment. *Transportation* 30(1), 37-62.
- Arentze, T., Hofman, F. & Timmermans, H. (2004). Predicting multi-faceted activity-travel adjustment strategies in response to possible congestion pricing using an internet-based stated adaptation experiment, *Transport Policy* 11(1), 31-41.
- Auld, J. & Mohammadian, A. (2012). Activity planning processes in the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Research Part A* 46(8), 1386-1403.
- Auld, J., Hope, M., Ley, H., Sokolov, V., Xu, B., & Zhang, K. (2016). POLARIS: Agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations. *Transportation Research Part C: Emerging Technologies*, 64, 101-116.
- Balmer, M., Rieser, M., Meister, K., Charypar, D., Lefebvre, N., Nagel, K., & Axhausen, K. (2009). MATSim-T: Architecture and simulation times. *Multi-agent Systems for Traffic and Transportation Engineering*, 57-78.
- Bekhor, S., Dobler, C., & Axhausen, K. (2011). Integration of activity-based and agent-based models: Case of Tel Aviv, Israel. *Transportation Research Record: Journal of the Transportation Research Board*, 2255, 38-47.
- Bellemans, T., Kochan, B., Janssens, D., Wets, G., Arentze, T., & Timmermans, H. (2010). Implementation framework and development trajectory of FEATHERS activity-based simulation platform. *Transportation Research Record: Journal of the Transportation Research Board*, 2175, 111-119.
- Bennett, P. N. (2003). Using asymmetric distributions to improve text classifier probability estimates. In: the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, 2003. ACM, 111-118.

- Börjesson, M., Eliasson, J., Hugosson, M. B., and Brundell-Freij, K. (2012). The Stockholm congestion charges - 5 years on. Effects, acceptability and lessons learnt. *Transport Policy* 20, 1-12.
- Bowman, J. L. & Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A* 35(1), 1-28.
- Brier, G. W. (1950). "Verification of forecasts expressed in terms of probability", *Monthly weather review*. 78(1), 1-3.
- Bustillos, B. I., Chiu, Y. C., & Papayannoulis, V. (2011). Pre-trip and en-route multi-modal travel decisions considering habitual travel times under unexpected incident scenarios. In *14th International IEEE Conference on Intelligent Transportation Systems* (pp. 594-599), IEEE.
- Cambridge Systematics, Inc. (2016). BMC InSITE Activity-Based Model Estimation. Prepared for Baltimore Metropolitan Council.
- Cendrowska, J. (1987). PRISM: An algorithm for inducing modular rules. *International Journal of Man-Machine Studies*, 27, 349-370.
- Chen X., Zhu Z., He X. and Zhang L.\* (2015) Surrogate-based optimization for solving mixed integer network design problem. *Transportation Research Record*, accepted.
- Chiu, Y. C., Bottom, J., Mahut, M., Paz, A., Balakrishna, R., Waller, T., & Hicks, J. (2011). Dynamic traffic assignment: A primer. *Transportation Research E-Circular (E-C153)*.
- Cohen, W. W. (1995). Fast effective rule induction. In: Prieditis, A., and Russell, S. (eds.) *Twelfth International Conference on Machine Learning*. Morgan Kaufmann Publishers.
- DeGroot, M. H. and Fienberg, S. E. (1983). "The comparison and evaluation of forecasters", *The statistician*. 12-22.
- Dia, H. (2002). An agent-based approach to modelling driver route choice behaviour under the influence of real-time information. *Transportation Research Part C: Emerging Technologies*, 10 (5):331–349.
- Durkin, J. (1994). *Expert Systems: Design and Development*, New York, Maxwell.
- FHWA. (2012). Office of Operations: ATDM Program Brief. Retrieved July 26, 2016, from <http://ops.fhwa.dot.gov/publications/fhwahop12032/index.htm>.
- Flötteröd, G., Bierlaire, M. & Nagel, K. (2011). Bayesian demand calibration for dynamic traffic simulations. *Transportation Science*, 45(4), 541-561.
- Frank, E. and Witten, I. H. (1998). Generating accurate rule sets without global optimization. In: Shavlik, J. (ed.) *Fifteenth International Conference on Machine Learning*. Morgan Kaufmann Publishers.
- Frei, A., Mahmassani, H., Zockaie, A., & Frei, C. (2014). Integrating behavioral models in network operations: Evaluating traveler information and demand management for weather-related events. *Transportation Research Record: Journal of the Transportation Research Board*, 2434, 80-88.

- Gill, J. (2002). *Bayesian Methods: A Social and Behavioral Science Approach*. CRC Press. 2002.
- Good, I. J. (1952). "Rational decisions", *Journal of the Royal Statistical Society. Series B (Methodological)*. 107-114.
- Goulias, K. G., Bhat, C. R., Pendyala, R. M., Chen, Y., Paleti, R., Konduri, K. C., Lei, T., Tang, D., Yoon, S. Y., Huang, G., & Hu, H. H. (2012). Simulator of activities, greenhouse emissions, networks, and travel (SimAGENT) in Southern California. In 91st Annual Meeting of the Transportation Research Board, Washington, DC.
- Hao, J., Hatzopoulou, M., & Miller, E. (2010). Integrating an activity-based travel demand model with dynamic traffic assignment and emission models: Implementation in the Greater Toronto, Canada, Area. *Transportation Research Record: Journal of the Transportation Research Board*, 2176, 1-13.
- Herbert, D. (1999). *Adaptive Learning by Genetic Algorithms: Analytical Results and Applications to Economic Models*. Springer-Verlag, New York.
- Humphreys, J.S., and Whitelaw, J.S. (1979). Immigrants in an unfamiliar environment: Locational decision-making under constrained circumstances. *Geografiska Annaler* 61B(1): 8-18.
- Jang, Y., Chiu, Y. C., & Zheng, H. (2013). Modeling within-day activity rescheduling decisions under time-varying network conditions. In S. V. Ukkusuri, and K. Obay (Eds.), *Advances in dynamic network modeling in complex transportation systems* (pp. 225-244). New York, NY: Springer.
- Khademi, E., Arentze, T. A., and Timmermans, H. J. P. (2012). Designing stated adaptation experiments for changes to activity-travel repertoires: approach in the context of pricing policies. In *Proceeding of European Transport Conference*, Glasgow, Scotland.
- Kitamura, R., Kikuchi, A., & Pendyala, R. M. (2008, June). Integrated, dynamic activity-network simulator: Current state and future directions of PCATS-DEBNetS. In 2nd TRB Conference on Innovations in Travel Modeling, Portland, OR, June (pp. 22-24).
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International joint Conference on artificial intelligence*, volume 14, pages 1137-1145. LAWRENCE ERLBAUM ASSOCIATES LTD.
- Kohavi, R. and H. J. George. (1997). Wrappers for Feature Subset Selection. *Artificial Intelligence*, Vol. 97, No. 1, 1997, pp. 273-324.
- Konduri, K. and R. Pendyala (2015). Updates to the BMC Population Synthesis Model: Incorporating Controls at Multiple Geographic Resolutions, Final Report.
- Konduri, K. C., Pendyala, R. M., You, D., Chiu, Y. C., Hickman, M. D., Noh, H., Waddell, P., Gardner, B., & Wang, L. (2013). Network-sensitive transport modeling framework for evaluating impacts of network disruptions on traveler choices under

- varying levels of user information provision. In Transportation Research Board 92nd Annual Meeting (No. 13-3366).
- Konduri, K., Ning, Y., Pendyala, R. & Zhou, X. (2014). Modeling behavioral response to real time traveler information using an integrated transport model framework. Presentation at 2014 Innovations in Travel Modeling, Baltimore MD, April 29 – May 2, 2014.
- Lee, M., Zhu, Z., Xiong, C. and Zhang, L. (2017). Integrating dynamic traffic assignment with ramp metering control for large-scale corridor management. Presentation at the 2017 TRB Annual Meeting, 17-05671. Washington D.C., Jan 8-12, 2017.
- Lin, D. Y., Eluru, N., Waller, S., & Bhat, C. R. (2008). Integration of activity-based modeling and dynamic traffic assignment. *Transportation Research Record: Journal of the Transportation Research Board*, 2076, 52-61.
- Mitsakis, E., Grau, J. M. S., Stamos, I., & Aifadopoulou, G. (2015). An integrated framework for embedding large-scale dynamic traffic assignment models in advanced traveler information systems. *Transportation Planning and Technology*, 38(8), 866-877.
- Outwater M. L., and Charlton, B. (2006). The San Francisco model in practice: Validation, testing, and application. Presented at the 2006 Innovations in Travel Modelling Conference, Austin, TX.
- Ouyang, L., Lam, W., Li, Z., & Huang, D. (2011). Network user equilibrium model for scheduling daily activity travel patterns in congested networks. *Transportation Research Record: Journal of the Transportation Research Board*, 2254, 131-139.
- Paz, A., & Peeta, S. (2009). Paradigms to deploy a behavior-consistent approach for information-based real-time traffic routing. *Networks and Spatial Economics*, 9(2), 217-241.
- Pendyala, R. M., Kitamura, R., and Prasuna Reddy, D. (1998). Application of an activity-based travel demand model incorporating a rule-based algorithm. *Environment and Planning B* 25, 753-772.
- Pendyala, R. M., Konduri, K. C., Chiu, Y. C., Hickman, M. D., Noh, H., Waddell, P., Wang, L., You, D., & Gardner, B. (2012). Integrated land use-transport model system with dynamic time-dependent activity-travel microsimulation. *Transportation Research Record: Journal of the Transportation Research Board*, 2303, 19-27.
- Pinjari, A. R., Pendyala, R. M., Bhat, C. R., and Waddell, P. A. (2011). Modeling the choice continuum: An integrated model of residential location, auto ownership, bicycle ownership, and commute mode choice decisions. *Transportation* 38(6), 933-958.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1, 81-106.

- Rossetti, R., Bordini, R., Bazzan, A., Bampi, S., Liu, R. and Van Vliet, D. (2002). Using bdi agents to improve driver modelling in a commuter scenario. *Transportation Research Part C: Emerging Technologies*, 10(5):373–398.
- Rothschild, M. (1974). Searching for the lowest price when the distribution of prices is unknown. *Journal of Political Economy* 82, 689-711.
- Salvini, P., and Miller, E. J. (2005). ILUTE: An operational prototype of a comprehensive microsimulation model of urban systems. *Networks and Spatial Economics* 5(2), 217-234.
- SimTRAVEL. (2016). OpenAMOS Modeling Framework. Retrieved July 26, 2016, from <http://simtravel.wikispaces.asu.edu/OpenAMOS+Modeling+Framework>.
- Srinivasan, K. K., & Mahmassani, H. S. (2003). Analyzing heterogeneity and unobserved structural effects in route-switching behavior under ATIS: A dynamic kernel logit formulation. *Transportation Research Part B: Methodological*, 37(9), 793-814.
- Sun, Z., Arentze, T., & Timmermans, H. (2005). Modeling the impact of travel information on activity-travel rescheduling decisions under conditions of travel time uncertainty. *Transportation Research Record: Journal of the Transportation Research Board*, 1926, 79-87.
- TCRP. (2016). TCRP Report 95 Collection: Traveler Response to Transportation System Changes Handbook. Retrieved July 26, 2016, from <http://www.trb.org/Publications/TCRPReport95.aspx>.
- USDOT. (2016b). Intelligent Transportation Systems Joint Program Office: Dynamic Mobility Applications. Retrieved July 26, 2016, from [http://www.its.dot.gov/research\\_archives/dma/index.htm](http://www.its.dot.gov/research_archives/dma/index.htm).
- Vovsha, P., Donnelly, R., and Gupta, S. (2008). Network equilibrium with activity-based microsimulation models: the New York experience. *Transportation Research Record* 2054, 102-109.
- Wahba, M., and Shalaby, A. (2011). Large-scale application of MILATRAS: Case study of the Toronto transit network. *Transportation* 38(6), 889-908.
- Xie, C., J. Lu, and E. Parkany. (2003). Work Travel Mode Choice Modeling with Data Mining: Decision Trees and Neural Networks. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1854, TRB, National Research Council, Washington, D.C., 2003, pp. 50-61.
- Xiong C. and L. Zhang (2013). A descriptive Bayesian approach to modeling and calibrating drivers' en-route diversion behavior. *IEEE Transactions on Intelligent Transportation Systems*. 14(4). 1817-1824.
- Xiong, C. and Zhang, L. (2013). Deciding whether and how to improve statewide travel demand models based on transportation planning application needs. *Transportation Planning and Technology*, 36(3), 244-266.
- Xiong, C. and Zhang, L. (201b). Positive model of departure time choice under road pricing and uncertainty. *Transportation Research Record* 2345, 117-125.

- Xiong, C., and Zhang, L. (2014). Dynamic travel mode searching and switching analysis considering hidden modal preference and behavioral decision processes. Presented at Transportation Research Board 93rd Annual Meeting, Washington D.C.
- Xiong, C., Chen, X., & Zhang, L. (2015). Multidimensional Travel Decision-Making: Descriptive Behavioral Theory and Agent-Based Models. *Bounded Rational Choice Behaviour: Applications in Transport*, 213.
- Xiong, C., Chen, X., He, X., Lin, X., & Zhang, L. (2016). Agent-based en-route diversion: Dynamic behavioral responses and network performance represented by macroscopic fundamental diagrams. *Transportation Research Part C: Emerging Technologies*, 64, 148-163.
- Xiong, C., Jun, J., Nan, X. and L. Zhang. (2016). The effects of personalized real-time information and incentives: behavior research, simulation, and measurements. Project Report. U.S. Department of Energy, Advanced Research Projects Agency – Energy (ARPA-E). DE-AR0000615.
- Ye, X., Pendyala, R. M., & Gottardi, G (2007). An exploration of the relationship between mode choice and complexity of trip chaining patterns. *Transportation Research Part B*, 41(1), 96-113.
- Zhang, B., Chan, W. K. V., & Ukkusuri, S. V. (2015). Adaptive routing and guidance approach for transportation evacuation. In 2015 Winter Simulation Conference (pp. 2008-2019), IEEE.
- Zhang, L. (2006a). Search, information, learning and knowledge in travel-decision making. Ph.D. Dissertation, Department of Civil Engineering, University of Minnesota.
- Zhang, L. (2006b). An agent-based behavioral model of spatial learning and route choice. In the compendium of papers CD of the 85th Transportation Research Board Annual Meetings, Washington, DC.
- Zhang, L. (2006c). Traffic diversion effect of ramp metering at individual and system levels. *Transportation Research Record: Journal of the Transportation Research Board*, 2012: 20-29.
- Zhang, L. (2007). Developing a positive approach to travel demand analysis: Silk theory and behavioral user equilibrium. In Allsop, R. E., and Benjamin, G. H. (eds), *Transportation and Traffic Theory*, Elsevier, pp. 791-812.
- Zhang, L. and Levinson, D. (2004). Agent-Based Approach to Travel Demand Modeling. In *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1898, 2004, pp. 28–36.
- Zhang, L. and Xiong, C. (2012). A positive model of departure time choice and peak spreading dynamics. *Transportation Research Board 91st Annual Meeting Compendium of Papers DVD #12-4514*, January, 2012, Washington D.C.
- Zhang, L., Chang, G. L., Zhu, S., Xiong, C., Du, L., Mollanejad, M., Hopper, N. & Mahapatra, S. (2013). Integrating agent-based travel behavior model with large-scale microscopic traffic simulation for corridor-level and subarea transportation

- operations and planning applications. *ASCE Journal of Urban Planning Development*. 139: 94-103.
- Zhang, L., Levinson, D., and Zhu, S. (2008). Agent-Based Model of Price Competition, Capacity Choice, and Product Differentiation on Congested Networks (August 1, 2008). *Journal of Transport Economics and Policy*, 42(3): 435-461
- Zhang, L., Southworth, F., Xiong, C., and Sonnenberg, A. (2012). Methodological options and data sources for the development of long-distance passenger travel demand models: a comprehensive review. *Transport Reviews*, 32(4), 399-433.
- Zhou, X., & Taylor, J. (2014). DTALite: A queue-based mesoscopic traffic simulator for fast model evaluation and calibration. *Cogent Engineering*, 1(1), 961345.
- Zhu, Z., Xiong, C., Chen, X. and Zhang, L. (2017). Calibrating Supply Parameters of Large-Scale Dynamic Traffic Assignment Models with Simulation-Based Optimization. Presentation at the 2017 TRB Annual Meeting, 17-03350. Washington D.C., Jan 8-12, 2017.