**Network Performance Monitoring and Distributed Simulation   
to Improve Transportation Energy Efficiency**

**TRANSNET-Atlanta**

**Project Overview, Space-time Memory, and Trajectory Simulations**

**June 19, 2018**

**R. Guensler, M.O. Rodgers, M. Hunter, A. Guin,   
K. Abdelghany, J. Laval, H. Liu, F. Gbologah  
Georgia Institute of Technology and Southern Methodist University**

# Introduction and Project Objectives

The TRANSNET-Atlanta team has developed a systems model designed to help minimize surface transportation energy consumption using historic traffic condition data, current system operating conditions, big data analytics, and control architecture interaction with users.   
The conceptual approach entails:

* Monitoring system-wide, real-time, high-fidelity traffic conditions from a variety of sources monitoring at high spatial and temporal resolution (i.e., big data)
* Generating a historic space-time memory for the regional roadway network, retaining the historic monitored traffic conditions as well as near-real-time traffic conditions
* Modeling changes in travel behavior in response to changes in the transportation system (lane closures, pricing, etc.), via a variety of behavioral modeling approaches, and integrating modeled results into the space-time memory
* Using statistical methods within the space-time memory to predict the evolution of traffic conditions (congestion formation and recovery patterns) on corridors throughout the system, in response to temporal changes in travel demand and non-recurrent events revealed through real-time monitoring of changes in traffic conditions
* Developing systems that predict ‘shortest-path’ trajectories through evolving congestion conditions from any origin to destination in the network, based upon pre-defined objective functions such as time, cost, energy use, etc.
* Delivering tailored messages to users to support pre-trip planning (and future in-flight messaging) as well as delivering incentives to influence travel decisions
* Modeling applicable energy consumption for each vehicle trajectory through the system and estimating system-wide energy consumption, travel time, and cost tradeoffs associated with changes in departure times, modes, routes, and ecodriving, given observed and predicted traffic conditions

Instead of focusing on large, centralized simulations, the long-term objective of our project is to develop statistical approaches to do a better job of predicting vehicle trajectories through the system as congestion is evolving. The short-term goal of this project helps resolve the “30-minute” commute dilemma:

A commuter queries an commercial online guidance systems for a route and travel time … they receive a response saying that their commute will take 30 minutes … 10 minutes into the trip the user receives an update indicating that the reminder of the trip will take 30 additional minutes.

This commute dilemma described above results from using streamlines (current speed conditions are assumed to remain static) rather than trajectories (speeds on traffic links that change over time during the commute). The TRANSNET-Atlanta system advises the commuter when they are initially planning the trip that their travel option will take 45 minutes.

Our ultimate goal is to implement agile, real-time simulation capabilities that reside on distributed network computers (e.g., signal control cabinets, OEM vehicle systems, and other user devices) rather than a centralized system. Each user will subscribe to corridor-level STM data feeds, individually predict relevant near-future corridor-level traffic conditions, and chart their own shortest paths through the network. Distributed simulations also react to differences in system predictions reported by other connected users (to adjust their own predictions) and to real-time changes in system conditions caused by incidents or other unanticipated events.

This report outlines the TRANSNET-Atlanta systems model background, systems model implementation, control architecture implementation, and energy and cost methodologies. Our objective is to outline the conceptual and computational techniques used and the suitability and completeness of the data used to construct and calibrate the models for Atlanta and the traveler behavior in question.

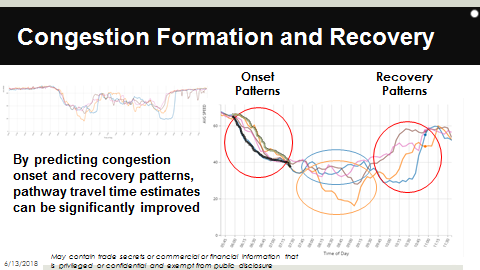
# Project Overview

The TRANSNET-Atlanta project team was composed of researchers from the Georgia Institute of Technology and Southern Methodist University with expertise in traffic simulation, dynamic traffic assignment, travel behavior, and energy modeling. The modeled geographic area was the Atlanta Metropolitan Area, a region with approximately 5.5 million residents and extensive bus and rail transit service. The transportation network model encompasses the area modeled by the Atlanta Regional Commission’s (ARC’s) activity-based travel demand model, covering the entire 20-county metropolitan area.

To achieve the TRANSNET program’s goal of reducing transportation energy consumption, the team was challenged with developing: 1) a system model to predict travel activity and energy use of the transportation system, based on user behavior, and 2) a system designed to deliver incentives to change travel behavior and reduce energy consumption. In our proposal, the team discussed the historical challenges associated with developing reliable large-scale simulation models that directly incorporate user behavior (e.g. Transims, previously applied in Atlanta but ultimately abandoned as impractical). Model developers must account for complex factors that drive travel choices and resulting travel behavior, a dearth of behavioral data from which to develop and calibrate behavioral models, and lengthy model run-times resulting from model complexity and the number of modeled elements (vehicles, signals, etc.) and interactions. Complex regional simulation models also require extensive on-going maintenance due to perpetual changes in roadway configuration, signal timing, etc.

The research team stated in the original proposal, and still believes at the end of this project, that the TRANSNET goal of achieving near-real-time simultaneous delivery of decision inputs to individual participants and network controllers should ultimately be met with alternative simulation approaches that have the potential to be much more “nimble” than traditional approaches. Rather than relying on centralized regional simulation modeling, the team proposed to rely on smaller-scale distributed simulation, coupled with pattern recognition, to assess dynamic responses to observed changes in system performance. Distributed approaches ultimately depend on development of complex communications infrastructures to share information across multiple users; however, these approaches avoid having to simultaneously model the evolving system state and the independent/co-dependent behavior of every participant. Hence, the team proposed to supplement centralized behavioral-based models (designed to assess regional behavior), with high-resolution, real-time monitoring of system operations to support participant decision making as the team assume that second-by-second big data will become ubiquitous through a variety of sources over the relatively near term.

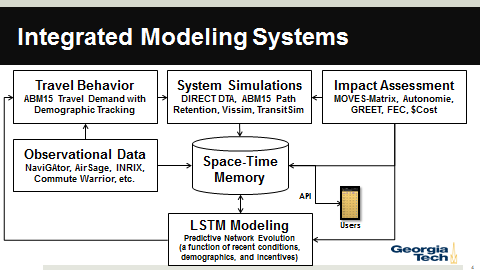
Distributed simulation at the participant level assesses travel time, energy cost, and financial cost participants are likely to experience along potential trajectories, given current system state conditions and the historic evolution of traffic (congestion formation and recovery patterns). The system also allows individual users to respond to a priori observed congestion evolution expected to arise under current conditions and assess how their travel costs will change when regions offer incentives to change departure times, travel routes, or travel modes. Fundamentally, as regions offer larger and larger incentives, the transportation system dynamically responds to these changes in travel activity. As congestion is relieved on one corridor due to incentive implementation, the travel cost for that pathway declines for users on other routes, potentially changing their travel decisions (and leading to a cascade of potential behavioral changes). Hence, the team anticipates that system change resulting from incentive implementation will occur slowly, as users learn over time (and consult navigation and information technologies) and make gradual adjustments to their travel behavior. These kinds of gradual changes make deep learning approaches viable for simulation. The fact that major events, such as lane closures and traffic incidents, cause major disruptions to the system, and occur relatively frequently over time, makes deep learning that incorporate historic events (incorporating long-term memory) viable for simulation.

We believe that our team has taken a unique approach to systems model implementation. Central to our integrated systems modeling approach, and used in all of simulation elements and energy calculations, is a space-time memory (STM) that contains observed historic system performance data at high spatial and temporal resolution. The STM contains a comprehensive archive of system conditions, including speed data for every link in the transportation system for which data are available, from which patterns and relationships between changes in traffic volumes, congestion formation (speed drops), and congestion recovery can be predicted as a function of system conditions. It is not possible to revolve the ‘30-minute commute’ dilemma until the system model is capable of predicting the congestion recovery pathway that the corridor is likely to experience as the commuter is moving through the system.

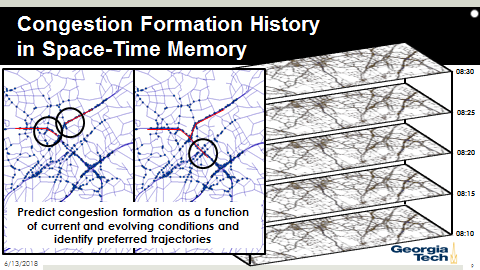
# Integrated Systems Model Components

The TRANSNET-Atlanta systems model is composed of a suite of modeling systems. Regional travel behavior is very complex, especially in regions with large populations, distributed employment centers, and extensive multi-modal transportation infrastructure. Fortunately, the Atlanta Regional Commission operates one of the most modern and complex activity-based travel demand models in the nation. The team uses the ARC’s ABM15 model for the purposes of predicting changes in travel demand associated with the implementation of major regional transportation incentives. Travel demand models employ complex transportation networks, but simplify the prediction of pathways between origin zones and destination zones. The interchange of trips between thousands of zones in the travel demand models requires an iterative approach to route allocation, where the model identifies new alternative paths between zones per iteration based upon generalized supply/demand congestion relationships. When complete, the model has generated millions of pathways between zones, but these pathways do not necessarily reflect the more nuanced interactions between traffic flows, roadway design, and signalized intersection control. To assess how drivers respond to the engineering performance of the system, the tram has integrated Dynamic Traffic Assignment (DTA) and Vissim® microsimulation models. Atlanta also has one of the most comprehensive traffic monitoring systems in the nation, with data feeds from the Georgia NaviGAtor ITS system providing speed and flow data at 20-second, lane-by-lane resolution for the majority of the freeway system. These data feeds can serve as model inputs, and/or be used in a variety of model development and verification efforts.

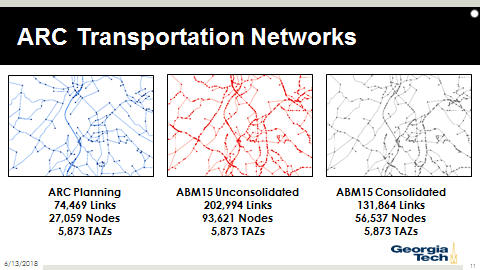
The figure below provides an overview of the systems model architecture and data flows. The core of this system is the space-time memory (STM) which allows the research team to integrate monitored and modeled data, use the data in LSTM statistical modeling approaches, and interact with system users via an API. The chapters that follow in this report focus on each of the modeling subsystem: space-time memory, observational data, travel behavior modeling, simulation modeling, user interaction, deep learning methods, and the variety of impact assessment tools used to quantify changes in energy use.



# Space-time Memory

The space-time memory is the centerpiece of our integrated systems modeling approach and constitutes the most significant accomplishment of the TRANSNET-Atlanta project. The primary purpose of the STM is to store current, historic, and projected system state data, in the formats needed by various systems and model algorithms. The STM is structured to receive system state data from any monitored or modeled source. For example, the TRANSNET-Atlanta project includes data feeds from: 1) the Southern Methodist University DIRECT Dynamic Traffic Assignment (DTA) model, 2) corridor-based speed data provided by AirSage® cellular phone monitoring, and 3) Georgia NaviGAtor ITS machine-vision (video and microwave based detection) monitoring of speed and volume (connections are ready). The team has also integrated data from other demonstration sources, such as INRIX speed data. Modeled data sources can range from regional travel demand models, to dynamic traffic assignment, to VISSIM® simulation models.

Each STM entry records the data source as an enumerated variable. This allows the team to specify which STM data to use in any systems model analysis (e.g., selection of specific data sources or subsets of data within sources). The STM needs to store historic performance data at sufficient spatial and temporal resolution to be useful to deep learning techniques applied to predict future system performance as a function of current observational data.

With respect to spatial resolution, the team is using the underlying network from the Atlanta Regional Commission’s (ARC’s) travel demand model. The ARC planning model currently employs a 75,000-transportation-link (75k-link) and node Planning Network. However, ARC is expanding to a larger network in their forthcoming travel demand model to enhance modeling resolution. The ARC developed a very large network for their next generation model using licensed spatial data from multiple sources and provided the TRANSNET-Atlanta team with this new 203k-link ABM15 Unconsolidated Network at the start of the project. Hence, the TRANSNET-Atlanta team constructed the current STM to match this new high-resolution network. The ARC has since decided that the 203k-link network is too complex (requires multiple days to complete model runs) and is now moving toward the implementation of the ABM15 consolidated network, which uses 132,000 links.

The team spent an extensive amount of time reconciling and verifying the accuracy of the three spatial networks, has submitted an updated ABM15 consolidated network to ARC, and plans to downsize the STM to match the 132k-link network as soon as possible. The research team performed link-to-link matching across all three networks and developed GIS capabilities to transfer data between model link realms as needed so that researchers can link any resolution model output with any other model resolution for use as model inputs or in comparative assessments. Another significant outcome of this TRANSNET project is the conclusion that the development of proper spatial networks is critical to the success of these projects. Regions interested in pursuing advanced modeling systems need to be thinking well in advance about how to structure their link-and-node networks, how to map data across networks, and how to perform detailed quality assurance assessments and continuous improvement evaluations.

With respect to temporal data resolution, the STM data can be implemented for any desired time increment. The STM includes a timestamp variable representing the beginning of the temporal data aggregation period, and a variable representing the aggregation time duration (seconds). Given the structure, the STM can be used at any reasonable predictive time scale depending upon desired use, such as long-term (e.g., seasonal) condition assessments, user pre-trip planning, or dynamic response decision making. The Atlanta STM originally managed data at 15-minute (900-second) resolution, but higher-resolution time-periods were required for proper modeling and comparative analysis of predictions across models. The team increased the STM data resolution to its current five-minute (300-second) resolution so that the research team could link various data sources together (e.g., AirSage® corridor speed data). However, the team also concluded it is not advisable to move to higher-resolution temporal tracking of onroad conditions. When temporal resolutions become too refined, not only do data processing requirements increase exponentially, the data quickly become irrelevant. For example, it is not useful to maintain STM cells at 30-second resolution for arterials, because traffic signal cycles are much longer than 30 seconds. Data in an STM at 30-second resolution simply become unusable for prediction of trajectories through the arterial system. The team has not yet conducted an extensive analysis to quantify the impacts of higher spatial resolutions on modeled predictions of vehicle trajectories through the system, nor has the team performed an optimization analysis. The team structured the temporal resolution at 5-minutes for use in deep learning analyses to help ensure that local data variability would not become a major problem.

The STM can be populated using any monitored or modeled source, provided the spatial resolution of the data matches the network link-and-node structure, and the data elements meet the prescribed STM formats. Whenever available, system state data for each transportation link contain: 1) a standardized set of facility link parameters; 2) link traffic volumes by vehicle source type; 3) average vehicle speed information, to calculate transportation-link-traverse travel time and economic travel cost by mode; and 4) detailed descriptions of onroad operating conditions, in various formats, for calculating energy consumption and emissions. Current state data become historic state data as soon as new data arrive to populate the STM. Additional data fields allow the team to use sample size, data quality (variability and representativeness), and overall data reliability in statistical modeling of system state projections in delivering travel time forecasts to shortest path algorithms and predicting origin-destination travel times by alternative routes, alternative departure times, and alternative modes. Predicted future state conditions (trajectories through the system) are also retained in the STM (using the same formats), specifying the model data source used to predict the future system state. The team believes that retention of predicted data will prove useful in future analytical work.

The regional space-time memory is predominantly composed of observational data. For freeways, the team manages data from the Georgia Department of Transportation (GDOT) Intelligent Transportation System (ITS) known as Georgia NaviGAtor. The streaming data provide traffic counts and monitored freeway speed data (lane-by-lane) at 20-second resolution for the majority of major freeways. The ITS Data Warehouse server receives, cleans, deletes duplicates, aggregates, and loads about 40 million records per day of traffic operations data into the space-time memory for the 203k-link Atlanta metropolitan transportation network. Monitored arterial data are currently available for very few locations via NaviGAtor, although these data are rapidly expanding as the GDOT deploys its new signal control architecture.

We currently populate observational arterial using a feed from AirSage® (vis RESTful API to receive real-time data), which provides per-vehicle travel times by link for the 203k-link network. The data are absorbed into the 8-node MongoDB cluster as they arrive. The AirSage® data are processed to update the STM every five minutes. The real-time updates and the archival STM data are then available to inform predicted traffic states for use in shortest path algorithms used in the Commute Advisor module. The AirSage® data stream produces about 7.5 million records a day, but it is important to remember that the cellular-probe data is still a sparse dataset; only about 1% of the links have data at any given point in time.

The systems model is adept at handling sparsity in observational data because the STM does not rely on any single data source. Real-time data, whenever available, update the STM. However, modeled data can be used to infill STM cells for which data are not available through monitoring systems. Modeling can apply underlying time-of-day/day-of-week historical averages or a fusion of data from multiple sources. For example, shortest path algorithms apply DTA-modeled congestion levels when real-time data are not available. Modeling can also provide STM background data for events that have yet to be observed (which would have been helpful in planning for the I-85 bridge span collapse that resulted in the complete freeway closure experienced in Atlanta).

The STM is capable of integrating data from any public or private data source, once the source data are aggregated or disaggregated to the desired time resolution and properly mapped to the STM network. The data linkages are not complicated and are readily supported by Python™ scripts that employ spatial mapping and matching within GeoPandas®. Hence, the STM can incorporate vehicle speed or other corridor monitoring data from Google, Waze, Here or other private data source (these companies do currently archive such data but are not making the data publicly available). The team demonstrated this capability by integrating two months of INRIX data provided by the ARC. The team expects numerous data sources will become available for link-by-arterial speeds link as vehicle monitoring becomes ubiquitous.

The team has also designed the STM for use with a wide-variety of control architectures (and control architecture interfaces) given the standardized link-based structure and data formats. Hence, as APIs are developed, external apps that provide incentives for changes in travel behavior can receive data from the STM and write observed travel data to the STM. The team structured the STM to be applicable beyond the current scope of travel time and energy use predictions. For example, roadway characteristic information such as number of lanes, lane width, and other design elements are retained in the STM so that they do not have to be queried for other applications. The STM also contains a series of twenty additional user-defined variables that can be populated for other technology-to-market purposes.

The detailed structure of the STM is described in a separate memorandum. The team expects that data STM structures will need to evolve over time, as the team identifies new data sources and explores new uses for the data. The structure of the database has been specifically designed to allow the team to make changes to the data structures over time and post-process all previously collected data on the distributed computing system to add new data elements or change existing data structures.

# TransitSim and RoadwaySim

The TRANSNET-Atlanta team developed a Python™-based regional transit simulator (TransitSim) that can be called by Python™ scripts and interfaced with HTML, Java, and D3 for online interaction and output visualization. transit simulation model employs the content of standard general Transit Feed Specification (GTFS) data feeds to code the transit routes, transit stop locations, and timetables; hence, modelers can readily update the simulator whenever a transit route or schedule changes and a new GTFS file is available. The transit simulator allows users to select an origin and destination (by latitude and longitude, or address) and identify the ‘shortest’ path through the transit network by rail, bus, express bus, and bus/rail combination. The shortest path optimization currently minimizes travel time, but user-selected optimization functions can be implemented, such as lowest financial cost or lowest energy cost.

So that TransitSim could support park-and-ride as an available commute alternative for transit travel, the team also developed a Python™-based roadway simulator (RoadwaySim) that employs the 203k-link roadway network and subroutines in GeoPandas. This allowed the team to integrate full roadway graphs (network maps) into the modeling system. The team used the same 203k-link roadway system used in the STM. Each transit trip node is paired with its closest node in the roadway graph, to support transfers between the roadway network and transit network. Once the roadway map was constructed, the same shortest path algorithms used for transit alone are applied to identify the lowest travel cost paths from origin to destination through the integrated roadway/transit system. Most importantly, RoadwaySim queries the STM for link travel times for future time steps as a vehicle trajectory moves through the network. The research team used the simulator tools to assess the travel time, cost, and energy impacts of 2,400 morning commute trips in a case study presented at the Transportation Research Board Annual Meeting in January 2018 (Li, et al., 2018).

### Transit Graph Construction

To implement TransitSim, the team constructed two separate transit graphs (maps, given the spatial and temporal dependency) using the GTFS data, one for MARTA bus and rail transit and one for GRTA express bus services. In GTFS, the shift/schedule of a transit route is represented by a trip (i.e., a sequence of two or more stops with arrival and departure times). All morning-peak transit trips operated by each transit agency are loaded to the transit graph, i.e. a directed graph, where nodes represent each rail and bus stop, and edges represent transit paths between stops. Because local transit provides transfer opportunities, additional nodes and edges are added to the local transit graph to represent transfer trips. Development of the transit and transfer nodes and edges are described in the following steps and illustrated in the Figure below:

1. Add transit trip nodes - A node is created for each timestamp of a train or bus arriving at a stop (e.g., *A1*). Each node is labeled with the *stop ID* and the timestamp, assuming transit vehicles adhere to their schedules. A single stop will have multiple nodes if multiple arrivals occur at that stop during the morning-peak period.
2. Add transit trip edges - A directed edge is created between each pair of adjacent nodes on a transit trip, e.g. an edge between *A1* (source node) and *A2* to represent traveling from *A1* to *A2.*
3. Add transfer nodes and edges - Each transfer is represented by a transfer node and two directed edges; a walk edge and a wait edge. For each transit trip node, a transfer node is created to represent each of its nearby stops (defined as stops within 0.01 by 0.01 degree of latitude and longitude). For example, *B2’* is a transfer node of *A1*, and the timestamp of *B2’* is set as the timestamp of *A1* plus the walking time to *B2*, calculated by Manhattan distance and an assumed walking speed of 2.0 mph. A walk edge is created from a transit trip node to the transfer node with walking time as weight, e.g. an edge from *A1* to *B2’*. Then a wait edge is created from the transfer node to the next departing transit node with waiting time as weight, e.g. an edge from *B2’* to *B2*.



**Transit Graph (Map) Setup**

### Simulation of Transportation Network Trajectories

TransitSim processes each origin and destination request, identifies the latitude and longitude for the closest address. On the server-side, the combined simulator (TransitSim+RoadwaySim) calculates the ‘shortest-path’ trajectories (currently minimum travel time trajectories) from origin to destination for user-selected mode choices, including drive-only, local transit-only, express-bus-only, local transit park-and-ride, and express park-and-ride. For transit trips, the simulator routines first identify the user’s two closest transit boarding stops to the trip origin location (shortest Euclidean distance), and then identify the two transit alighting stops closest to the trip destination location (shortest Euclidean distance). Transit access and egress assumes walking access. This provides four potential origin-destination transit stop combinations. The transit simulator then first identifies the shortest linked transit pathways for the four origin-destination transit stop pairs. Shortest path routines are based on the weighted single source Dijkstra (1959) algorithm, with travel time as weight. A feasible arrival time constraint also ensures that transit options exclude unreasonable options (e.g. driving for an hour out of town to catch an express bus back into town). A case study application of the combined simulator system is presented later in this report.

Given the complexity of the transit graphs, a typical server request can take about 40 seconds to return eight shortest path options from the server. The team recently developed subsets of the regional roadway network graph to improve the speed of shortest path processing, so that rather than loading the entire roadway network graph, sections are loaded, depending upon the origin and destination selected. This reduced our shortest path processing time by roughly half. The team anticipates that the same process applied to the transit network can reduce the processing time by another 50%.

The most important accomplishment associated with the roadway simulator is that the roadway network simulator pulls travel times for each roadway link in the roadway graph from the STM. Trajectories through the system employ the forecasted travel time across each link given the time the user arrives at that link (i.e., time elapsed to traverse previous links and reach that link). Hence, the travel times through the network employ the projected evolution of congestion contained in the STM. At this time, the simulator predicts transit pathways and trip durations based only on the transit schedules, not on observed operating conditions. That is, transit is never late because transit agencies build in a schedule buffer so that even when buses experience normal traffic, they will arrive at the station in compliance with the posted schedule.

# Conclusions

The TRANSNET-Atlanta project resulted in the creation and implementation of a high-resolution space-time memory (STM) for the entire Atlanta region. This STM, implemented in MongoDB, resolves the entire Atlanta Regional Commission’s (ARC’s) 203,000-link transportation network with a five-minute time resolution. The STM connects directly to Georgia NaviGAtor system, the Georgia Department of Transportation’s (GDOT’s) intelligent transportation system (ITS) that monitors most Atlanta’s freeways and lane-by-lane traffic volumes and speeds at 20-second resolution. Similarly, the STM connects to the AirSage® real-time cell phone monitoring system for speeds on arterials. The STM also contains numerous roadway design parameters and operating condition variables used in energy modeling and for a variety of statistical applications. The STM is the centerpiece of the integrated modeling systems; all models in the system interface with the STM in some form. The research team considers the development and implementation of the STM to be the primary achievement of the project, and the element that will have the largest impact.

The STM serves as the connected data platform and data repository for both monitored and modeled data. Model run predictions from the ARC’s ABM15 regional travel demand model, the regional DIRECT DTA model, and a Vissim® microscopic simulation model for a large subarea of the region integrate with the STM. For example, the team uses the DTA-predicted speeds in the STM to develop trajectories through the system. Hence, a second major contribution of the STM is that the repository can accept monitored and modeled data from sources with different logical and modeling constructs, which will allow “deep learning” methods to use the STM for predictive purposes.

Another major accomplishment of the research team is the development and implementation of the Python™-based TransitSim and RoadwaySim modules that interact with the STM to identify the shortest path (currently optimized for travel time) multi-modal trajectories through the transportation network for any origin-destination pair, departure time, or desired arrival time. The team has developed scripts that allow researchers to send commutes (origin, destination, departure time) through the combined simulators and assess the travel time and energy use differences across modes, as demonstrated with 2,400 commute trips (Li, et al., 22018). The research team also implemented this interface in Commute Warrior for pre-trip planning, so that users could assess differences in the travel time, cost, and energy use associated with: choosing alternative routes, departing earlier or later than planned, walking to transit, using rail park-and-ride, and using the GRTA Xpress bus system. The important consideration here is that the system does not use the current travel times on the network to predict travel time through the system. Instead, the system uses projected travel times for future time steps (modeled) in the STM. As long as projected travel times are accurate, the system helps resolve the ‘30-minute commute’ dilemma. An API is now available that could be used by any stakeholder app to interact with an active STM for the purposes of identifying these shortest-path alternatives and the resulting trajectories through the system. Any stakeholder app that delivers travel incentives to users could use this API to inform the user about how their changes departure times, travel modes, and routes will affect their travel time, distance, cost, and energy use. From a systems perspective, agencies could modify the shortest path algorithms such that the system delivers selective messages to users that focus on reducing individual and corridor energy use. The research team considers this a significant technology-to-market achievement.

The modeling tools developed by the research team are 100% transferable to other regions. The STM is implemented in MongoDB (although the team is still assessing the time penalties associated with loading historic data into the STM and may recommend an alternative system) and almost every single routine developed for data processing is constructed in Python.™ Hence, the modeling tools are also structured for implementation in the cloud computing environment. However, a great deal of work that is required to fully-implement the integrated modeling systems in another region, from properly coding the spatial network, to properly formatting automating the data feeds from sources.

The team believes that this TRANSNET project has fundamentally changed the way that we as researchers are now approaching travel behavior and transportation energy modeling. The extensive experience associated with integrating these modeling tools with each other and with the STM has exposed a number of shortcomings in current modeling approaches. The team anticipates that the journal papers published based upon the experience and modeling results from this project will lead to significant changes in the way that these transportation models are developed and implemented.

# References

ARC, Atlanta Regional Commission. Activity-Based Travel Model Specifications: Coordinated Travel – Regional Activity Based Modeling Platform (CT-RAMP) for the Atlanta Region. Atlanta Regional Commission, 2012.

Argonne National Labs. GREET Model. https://greet.es.anl.gov/. Accessed Jul. 27, 2017.

Bast, H., D. Delling, A. Goldberg, M. Müller-Hannemann, T. Pajor, P. Sanders, D. Wagner, and R. F. Werneck. Route Planning in Transportation Networks. In Algorithm Engineering, Springer International Publishing, 2016, pp. 19-80.

Dijkstra, E. W. A Note on Two Problems in Connexion with Graphs. Numerische Mathematik, Vol. 1, No. 1, 1959, pp. 269–271.

Guensler, R., Liu, H., Xu, Y., Akanser, A., Kim, D., Hunter, M.P. and Rodgers, M.O., 2017. Energy Consumption and Emissions Modeling of Individual Vehicles. Transportation Research Record: Journal of the Transportation Research Board, (2627), pp.93-102.

Guensler, R., H. Liu, X. Xu, Y. Xu, and M. Rodgers (2016). MOVES-Matrix: Setup, Implementation, and Application. Presented at 95th Annual Meeting of the Transportation Research Board, Washington, D.C., 2016.

Ke, J., Zheng, H., Yang, H. and Chen, X.M. (2017). Short-Term Forecasting of Passenger Demand Under On-Demand Ride Services: A Spatio-Temporal Deep Learning Approach. Transportation Research Part C: Emerging Technologies, 85, pp.591-608.

Kim, D., Y. Zhao, M.O. Rodgers, and R. Guensler (2018). Personal Vehicle Ownership and Operating Cost Calculator. <http://costcalculator.ce.gatech.edu/>. National Center for Sustainable Transportation. Georgia Institute of Technology. Atlanta, GA.

Li, H., Y. Wang, X. Xu, H. Liu, A. Guin, M. Rodgers, M. Hunter, J. Laval, and R. Guensler (2018). Assessing the Time, Monetary, and Energy Costs of Alternative Modes (18-05362). 97th Annual Meeting of the Transportation Research Board (presentation only, full paper review, extended abstract in proceedings). Washington, DC. January 2018.

Liu, H., X. Xu, M. Rodgers, Y. Xu, and R. Guensler (2017). "MOVES-Matrix and Distributed Computing for Microscale Line Source Dispersion Analysis." Journal of the Air and Waste Management Association. DOI: 10.1080/10962247.2017.1287788

Lv, Y., Y. Duan, W. Kang, Z. Li, and F.Y. Wang (2015). Traffic Flow Prediction with Big Data: A Deep Learning Approach. IEEE Transactions on Intelligent Transportation Systems, 16(2), pp.865-873.

Ma, X., Z Tao, Y. Wang, H. Yu, and Y. Wang (2015). Long Short-Term Memory Neural Network for Traffic Speed Prediction Using Remote Microwave Sensor Data. Transportation Research Part C: Emerging Technologies, 54, pp.187-197.

Polson, N.G., and V.O. Sokolov (2017). Deep Learning for Short-Term Traffic Flow Prediction. Transportation Research Part C: Emerging Technologies, 79, pp.1-17.

PTV Vissim 9 User Manual. PTV AG, Karlsruhe, Germany. 2016. <http://vision-traffic.ptvgroup.com/en-us/training-support/support/ptv-vissim/>

Ren, H., Y. Song, J. Wang, Y. Hu, and J. Lei (manuscript). A Deep Learning Approach to the Citywide Traffic Accident Risk Prediction. Accessed at: https://arxiv.org/abs/1710.09543.

Simoncini, M., L. Taccari, F. Sambo, L. Bravi, S. Salti, and A. Lori (2018). Vehicle Classification from Low-frequency GPS Data with Recurrent Neural Networks. Transportation Research Part C: Emerging Technologies, 91, pp.176-191

USEPA, U.S. Environmental Protection Agency (2015). eGRID. Retrieved July 1, 2016 from: <https://www.epa.gov/energy/egrid>

USGS, U.S. Geographical Survey. (2016). “Digital Elevation Models” Available at: <http://nationalmap.gov/elevation.html> Accessed on June, 2016.

Wijnands, J.S., J. Thompson, G.D. Aschwanden, and M. Stevenson (2018). Identifying Behavioural Change Among Drivers Using Long Short-Term Memory Recurrent Neural Networks. Transportation Research Part F: Traffic Psychology and Behaviour, 53, pp.34-49.

Wu, Y., H. Tan, L. Qin, B. Ran, and Z. Jiang (2018). A Hybrid Deep Learning Based Traffic Flow Prediction Method and its Understanding. Transportation Research Part C: Emerging Technologies, 90, pp.166-180.

Xu, X., H. Liu, Y. Xu, M. Hunter, and R. Guensler (2016a). “Estimating Project-level Vehicle Emissions using Vissim and MOVES Matrix.” DOI 10.3141/2570-12. Transportation Research Record. Number 2570. pp. 07-117. National Academy of Sciences. Washington, DC. 2016.

Xu, Y., H. Li, H. Liu, M.O. Rodgers, R. Guensler (in 2016b). “Eco-driving for Transit: An Effective Strategy to Conserve Fuel and Emissions. Applied Energy.” APEN8702. DOI: 10.1016/j.apenergy.2016.09.101. October 2016.

Xu, X., Y. Xu, Y. Zhao, H. Liu, H. Cheng, M. O. Rodgers, and R. Guensler (2016c). Fuel and Emissions Calculator (FEC) Version 2.0 (No. NCST-091316). <https://trid.trb.org/view.aspx?id=1424897>

Xu, Y., H. Liu, and R. Guensler (2017). “Understanding the Emission Impacts of HOV to HOT Lane Conversions: Experience from Atlanta, GA.” Journal of the Air and Waste Management Association. UAWM-2016-0183.R1.

Yu, R., Y. Li, C. Shahabi, U. Demiryurek, and Y. Liu (2017), June. Deep Learning: A Generic Approach for Extreme Condition Traffic Forecasting. In: Proceedings of the 2017 SIAM International Conference on Data Mining (pp. 777-785). Society for Industrial and Applied Mathematics.

# Appendix A - Energy Modeling for Transit Trips

## Energy Modeling for Bus Transit Trips

Transit buses are modeled as a specific vehicle source type in the USEPA’s MOVES model; hence, the energy modeling approach for transit bus trips also employs MOVES-Matrix. For heavy-duty vehicles, MOVES calculates scaled tractive power (STP), using the same equation as VSP with the exception that m/M is heavy-duty-vehicle-class-specific (i.e., not set to 1.0 as for light-duty vehicles). As previously described for light-duty vehicles, the energy approach calculates energy use as a function of second-by-second speed, acceleration, and applicable vehicle parameters. Similarly, energy use on any transportation link under any observed or simulated condition is calculated using the Watson plot approach (i.e., coupling the Watson plot with the applicable energy use map for the transit bus source type).

The team is using the Python™ transit simulator described earlier to model all regional transit activity. Modeled users taking transit leave the roadway network simulation at a transit stop, enter the transit simulator, traverse applicable transit routes (bus and/or rail) to their destination transit stop, and re-enter the roadway network. MARTA currently does not monitor vehicle activity at sufficient resolution to generate representative Watson plots for any routes. Each bus transit route is currently assigned a representative Watson plots based upon average route travel speed, based upon route distance and origin-departure to destination-arrival time. Watson plots derived from MOVES are representative of transit operations. The MOVES cycles were all based upon monitored transit operations, exhibiting moderate speeds, low vehicle acceleration rates, midblock stop-start activity for passenger boarding and alighting, and extended idle. As second-by-second data become available routes through Commute Warrior monitoring, the average speed Watson plots currently employed in energy calculations will be replaced with Watson plots derived from route-specific data. The team currently assumes transit buses follow their schedules, bus schedules factor in congestion by time of day, and schedules build in wait times along the route to ensure that buses do not depart any stops earlier than riders expect. Hence, specifically factoring in route-specific congestion impacts may not result in significant changes in modeled energy consumption. The research team anticipates that the high-resolution data collected from transit users on the Commute Warrior app will allow us to assess by fall whether roadway congestion also needs to be built into the transit operations.

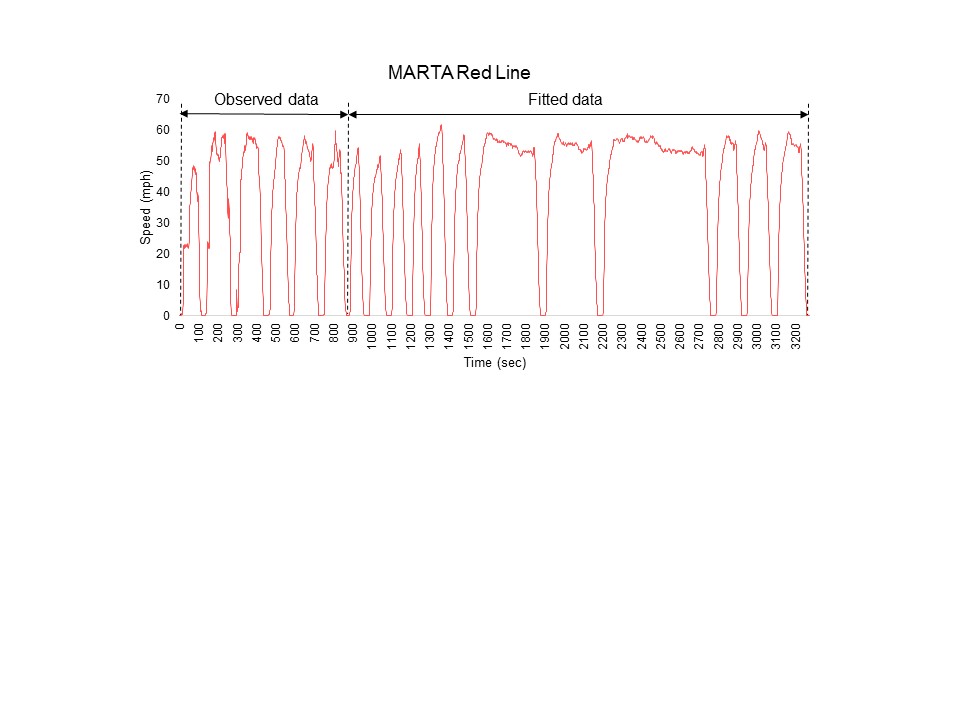
In the Phase I modeling effort, the team is assuming that changes in individual travel behavior associated with targeted incentives are not large enough to significantly shift travel behavior across modes. Hence, the team has not yet factored in the potential impacts of transit demand and capacity. In the Phase I modeling effort, buses are assumed to operate at 40% capacity for the purposes of allocating bus energy use on a per-passenger basis. In the Phase II modeling effort, where large regional incentives designed to affect major changes in travel behavior are simulated (e.g., regional time-of-day parking pricing), significant shifts to transit are expected. This summer, the team will integrate passenger count data from MARTA’s automatic passenger counting systems into the transit simulation framework. Once count data are integrated into the database, and linked to the simulator routes, monitored capacity will be used to allocate bus energy use on a per-passenger basis. In Phase II modeling efforts, the team will integrate bus demand and capacity constraints into the simulator system and address the impacts of predicted route oversubscription. Passenger wait times will need to increase in some scenarios because buses leave full with riders at the station (which drives mode choice away from transit), and increased bus frequency will need to be implemented for other scenarios to ensure adequate capacity (which increases per-passenger energy costs).

## Energy Modeling for MARTA Rail Trips

For rail system emission analysis, the team applied the rail calculator in the Fuel and Emissions Calculator (FEC) developed for the National Center for Sustainable Transportation. The Fuel and Emissions Calculator (FEC) is an operating-mode-based, life-cycle emissions modeling tool developed by the Georgia Institute of Technology researchers that can be downloaded by any interested party (Xu, et al., 2016c). The FEC’s modeling approach estimates energy use and emissions as a function of engine load, which in turn is a function of vehicle service parameters, allowing modelers to account for local on-road operating mode conditions as model inputs. The latest FEC (Version 2.0) includes dedicated algorithms for transit bus, rail, shuttle bus/van, and long-haul combination trucks. To estimate energy consumption and emissions generated by rail system for Atlanta region, the team: 1) collected and processed rail operating mode (speed/acceleration cycle) data to provide input data for rail calculator, and 2) performed the emission analysis using the FEC rail calculator coded with a Python™ script. MARTA’s rail system was analyzed for this study. The energy consumption and emissions were estimated for at the route level and at the individual passenger level.

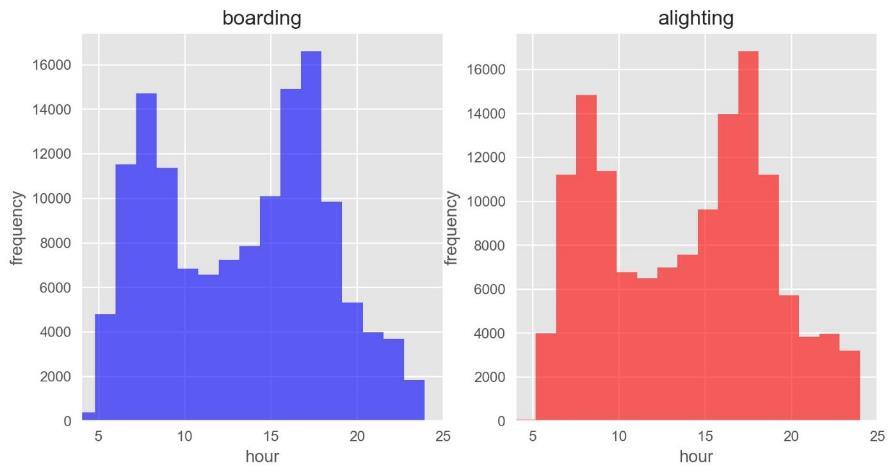
In the first task, the team collected GPS data on the rail system to generate speed/acceleration profiles. Because GPS data are not available in tunnels, the team fitted the space between observed data using above ground speed/acceleration activity generate second-by-second speed data for underground operations. The fitted cycle was generated based on GPS data collected above ground, distance between stations measured with open-source GPS data, and travel times between stops from MARTA’s rail schedule. The rail duty cycle between stations was first divided into four sections: acceleration, cruise, deceleration and idle. Second-by-second data were then generated for each section, based on observed cycle, measured distance, and travel time. Finally, the fitted speed profiles for each section were stitched together to generate the cycle for entire rail routes. Rail cycle data were generated for all four MARTA rail lines in this manner. The fitted cycle for MARTA’s red line is shown in the figure below.

Calculating energy use and emissions on any line uses the generated cycle data for the rail line with the FEC rail calculator. The FEC rail calculator (spreadsheet format) was coded as a Python™ Script to reduce processing time. Next, the team prepared the locally-derived route information and local railcar characteristics as the model inputs. In the analysis module, the energy consumption of all-electric rail operations is estimated by calculating the tractive load, hoteling load, and energy from recovery respectively. Each portion energy consumption element was estimated using the second-by-second speed profile for flat terrain settings (MARTA’s rail grade never exceeds 2%). The total energy consumption is calculated by summing the tractive load and hoteling load, and subtracting energy recovered in regenerative braking. Finally, the energy and emission results are estimated by multiplying energy consumption and emission rates per unit energy consumed. In the FEC, net energy and emission rates for electricity come from the USEPA’s eGRID database (US EPA, 2015).



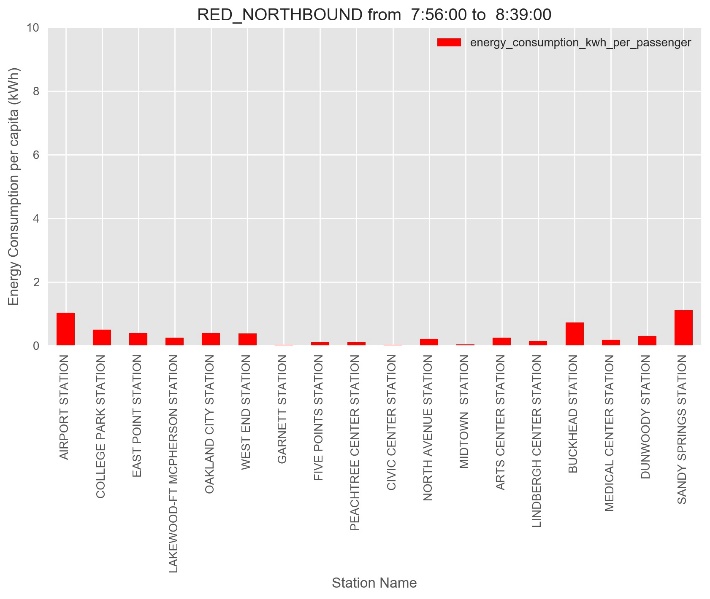
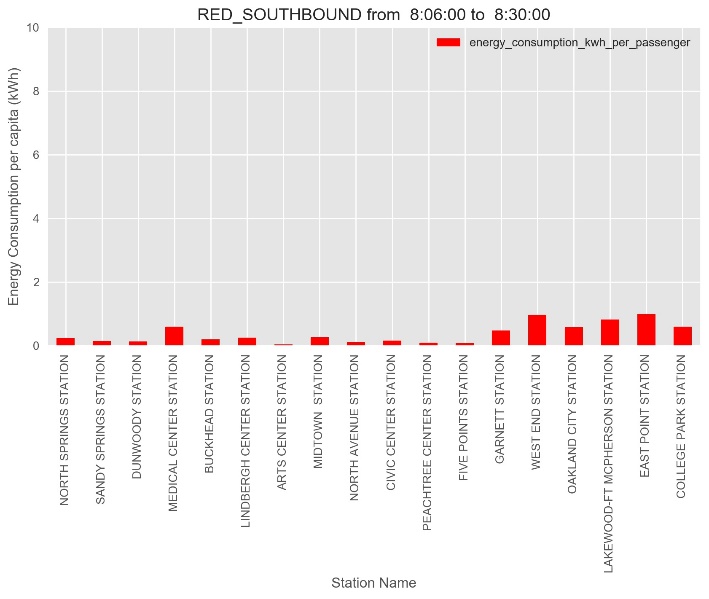
**Figure A-1. Fitted Rail Cycle for MARTA’s Red Line**

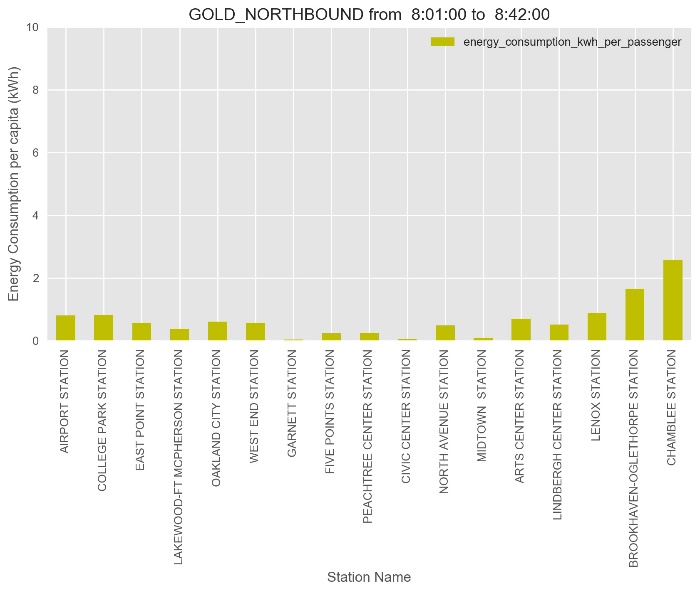
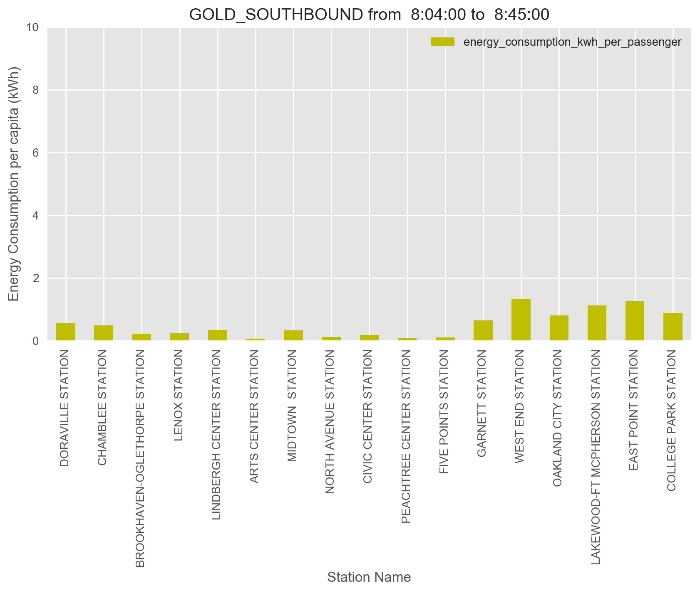
Energy consumption and emissions are calculated at the trip level, and at the passenger level, using real-world rail O-D data for Atlanta. The rail route and trip data were derived from the General Transit Feed Specification (GTFS). The rail trip O-D pairs came from the MARTA smart card data from a typical weekday (but in the future a live data feed could be developed). There were 137,487 valid rail trips on that typical MARTA operation day. The temporal distributions of boardings and alightings are illustrated below:

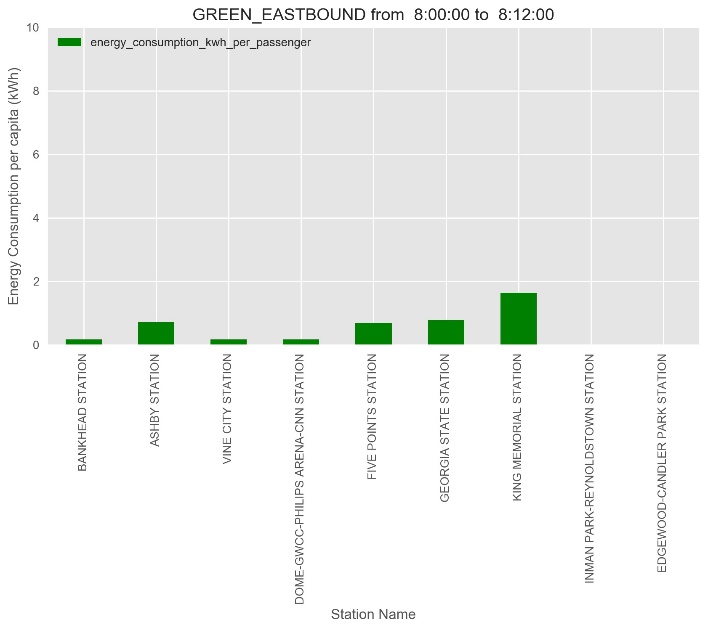
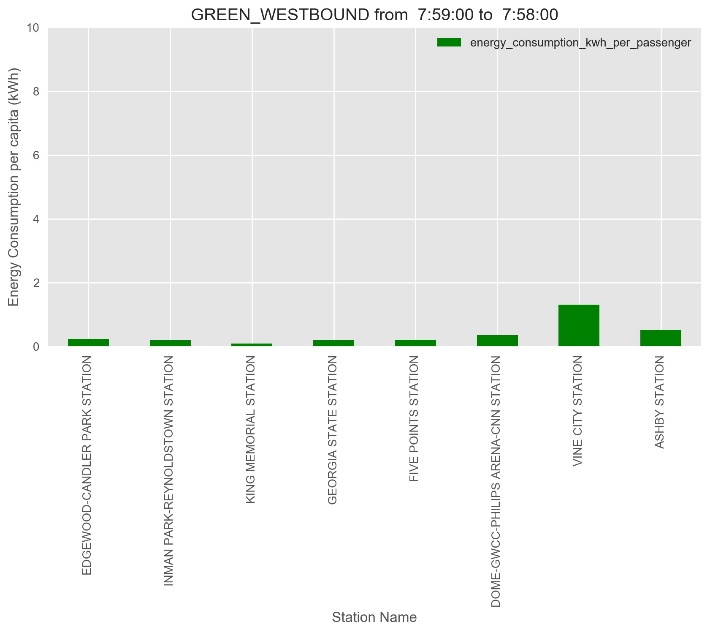


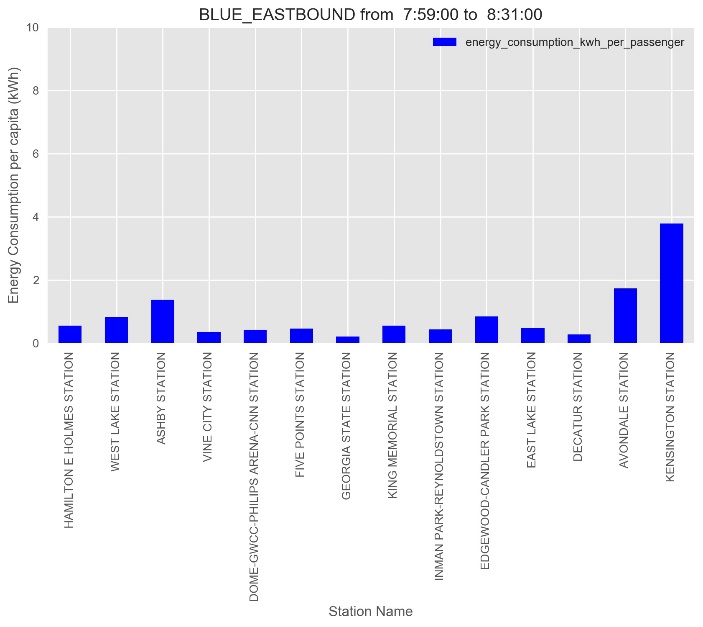
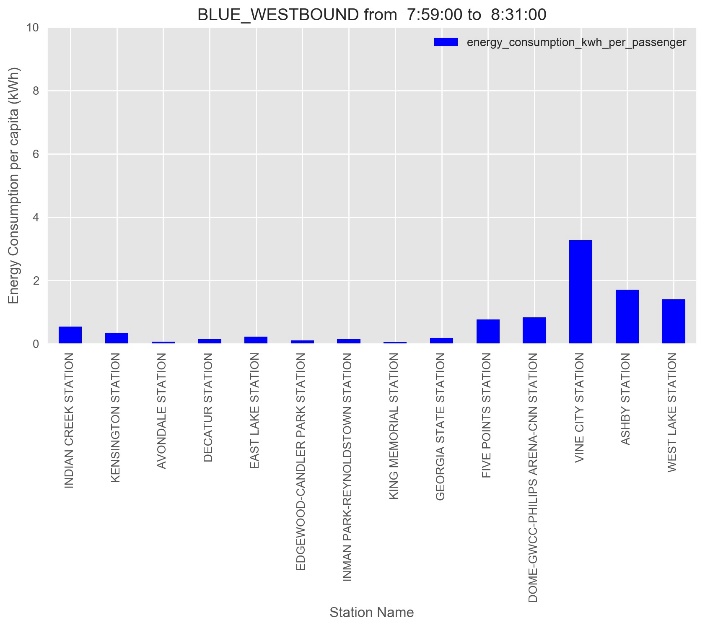
**Figure A-2. MARTA Rail Ridership**

Rail trips are assigned to each rail route based on shortest travel time and distance rules. The total rail energy for each trip is divided by the number of onboard riders to obtain the average energy use per person. Plots of energy consumption per capita during 8:00 a.m. trips from each route are provided below:









**Figure A-3. MARTA Rail Route Energy Consumption from Selected Trips**