SERPM 8.0 Model Update

Model Validation Plan

approved

plan

prepared for

RTTAC-MS

prepared by

Cambridge Systematics, Inc.

February 15, 2017

www.camsys.com

report

SERPM 8.0 Model Update

Model Validation Plan

prepared for

RTTAC-MS

prepared by

Cambridge Systematics, Inc.
2101 West Commercial Boulevard, Suite 3200
Fort Lauderdale, FL 33309

date

February 15, 2017

Table of Contents

List of Tables

List of Figures

Executive Summary

Insert text here.

1.0 Overview

This document provides direction for validation of the Southeast Florida Regional Planning Model (SERPM) 8. The development of this validation plan and the validation procedures and recommendations presented herein are based on guidance provided in the 2011 Travel Model Improvement Program *Travel Model Validation and Reasonableness Checking Manual – Second Edition* (Cambridge Systematics, Inc., 2010), hereafter referred to as the "Validation Manual." The validation standards provided in FDOT validation standards: FSUTMS-Cube Framework Phase II Model Calibration and Validation Standards Final Report (Cambridge Systematics, 2008) are also referenced as a source of guidelines. Information from the Federal Transit Administration (FTA) Section 5309 New Starts guidance is also considered in the development of this validation plan although it is recognized that recent federal legislation is changing these guidelines. Finally, this validation plan builds on experience gained with the Houston-Galveston Area Council (H-GAC), Minneapolis-St. Paul Metropolitan Council (Met Council), and Baltimore Metropolitan Council (BMC) Activity-Based Model (ABM) validation.

1.1 Validation Process

The Validation Manual recommends that a validation plan be developed in conjunction with the model design plan whenever a travel model is estimated or updated. The validation plan should assess the available validation data, determine additional validation data that might reasonably be collected, determine what can be validated, set priorities, and specify any guidelines or standards that might be necessary.

The validation plan should also set the stage for quality model validation documentation. The validation documentation should be an honest assessment of how the model performs. Thus, rather than a blanket statement that the model is valid since it has met some artificial standard such as "the R^2 for assigned traffic volumes versus observed traffic counts exceeds 0.89," the validation documentation will summarize how closely the various model components reproduce observed data and the sensitivities of the various model components. The model validation and model validation documentation are intended to demonstrate levels of confidence that can reasonably be placed in model results as well as providing information to help set priorities for future data collection and model improvement efforts.

Finally, based on experience gained with the aforementioned ABM validations, individual model component calibrations / validations should focus on how well each component reproduces distributions and data that will be passed to subsequent modeling components. This is accomplished through a stepwise calibration / validation process where the ABM is applied and individual model components updated as necessary[1] in a stepwise manner from start to finish. Each model component is adjusted to reasonably match available observed data (typically expanded data from the survey data used for model estimation) in sequence. Data from the model component(s) higher in the modeling chain are used as input to components lower in the chain. In this way, the overall impact of error propagation is reduced. The amount of information produced from the various tests can be overwhelming. Thus, as described in this validation plan, more detailed validation checks will be performed and analyzed as necessary for diagnosis when less detailed checks indicate a possible modeling issue.

[1] Primary updates include adjustments to model constants, followed by updates to coefficients of 0/1 variables, and finally updates to coefficients of continuous variables. In some cases, the addition of a new variable to the model specification may be required.

1.2 Types of Validation Checks

Four primary types of validation checks can be specified:

- **Base year comparisons**. Base year comparisons are, perhaps, the weakest validation checks since they are often performed against the same data that were used for model estimation or since they are results of calibration efforts (e.g. of model constants or model parameters) to match observed base year conditions. The usefulness of base year comparisons can be enhanced by splitting observed data into estimation and validation data sets (provided sufficient data exist) or validating models to observed data using stratifications different from those used for calibration.
- **Temporal comparisons**. Estimating and calibrating a model using data from one year and validating the model using data from a different time period is a strong validation process. It is especially strong if there are substantial socioeconomic, demographic, or transportation supply changes that take place between the two time periods used for the estimation/calibration and validation.
- **Sensitivity testing**. The application of the models and the model set using alternative input data or assumptions is especially important for models, such as the SERPM 8, that are designed to model traveler behavior, not travel patterns. Sensitivity testing of individual model components may include the estimation of the elasticities and cross-elasticities of model coefficients. However, sensitivity testing should also include the application of the entire model set using alternative assumptions regarding the input demographic data, socioeconomic data, or transportation system to determine if the model results are plausible and reasonable.
- **Reasonableness and logic testing**. These tests include the types of checks that might be made under model sensitivity testing. These checks also include the comparison of estimated (or calibrated) model parameters against those estimated in other regions with similar models. Reasonableness and logic checks may also include "components of change" analyses and an evaluation of whether or not the models "tell a coherent story" as recommended by the FTA for New Starts analysis.

1.3 Validation Considerations

1.3.1 Accuracy Requirements and Guidelines

The Validation Manual strongly supports the notion that matching specified standards is neither necessary nor sufficient to prove model validity. It avoids the specification of validation standards for this reason. Nevertheless, it is recognized that past standards may have been set by agencies such as FDOT or others regarding model validation statistics and failure to match those standards will cast doubts on model's validity and usability. Thus, from a practical standpoint, it's important to match, and where possible, exceed the standards. Such standards, the FSUTMS Model Calibration and Validation Standards in particular, will be acknowledged and the necessary model statistics will be calculated. However, the meaning and implication of achieving (or not achieving) the standard will also be discussed.

1.3.2 Aggregation Level

Validation may be performed at two levels of aggregation:

- **Disaggregate validation**. As used in this document, disaggregate validation refers to comparisons performed at the household or individual level. Validation and reasonableness checking measures such as elasticities, prediction success tables, or R^2 are all examples of disaggregate measures providing the base unit for producing the measure is an individual or household. Individual information is required for both the modeled and the observed travel behavior being compared.

- **Aggregate validation**. Aggregate validation refers to comparisons performed after individuals or households have been aggregated over some common variable such as a geographic unit (zone, district, or community) or a socioeconomic unit (e.g. household size, income group, or auto ownership). Aggregate validation allows error for one trip making unit to cancel error for another similar unit in the aggregation scheme. By definition, measures based on traffic counts or transit boardings are aggregate measures since the counts or boardings are aggregations of individuals on roadway links or transit lines. For the SERPM 8 model, two different district structures may be appropriate based on the specific-validation test performed: the 18 Districts and 3 CBD areas or the Public Use Microdata Areas (PUMA) (see Figure 1.1). In other cases, it may be most appropriate to perform comparison for three counties in the SERPM model region (Miami-Dade, Broward, and Palm Beach). Specific district structures for aggregate validation tests are recommended in Chapter 5.

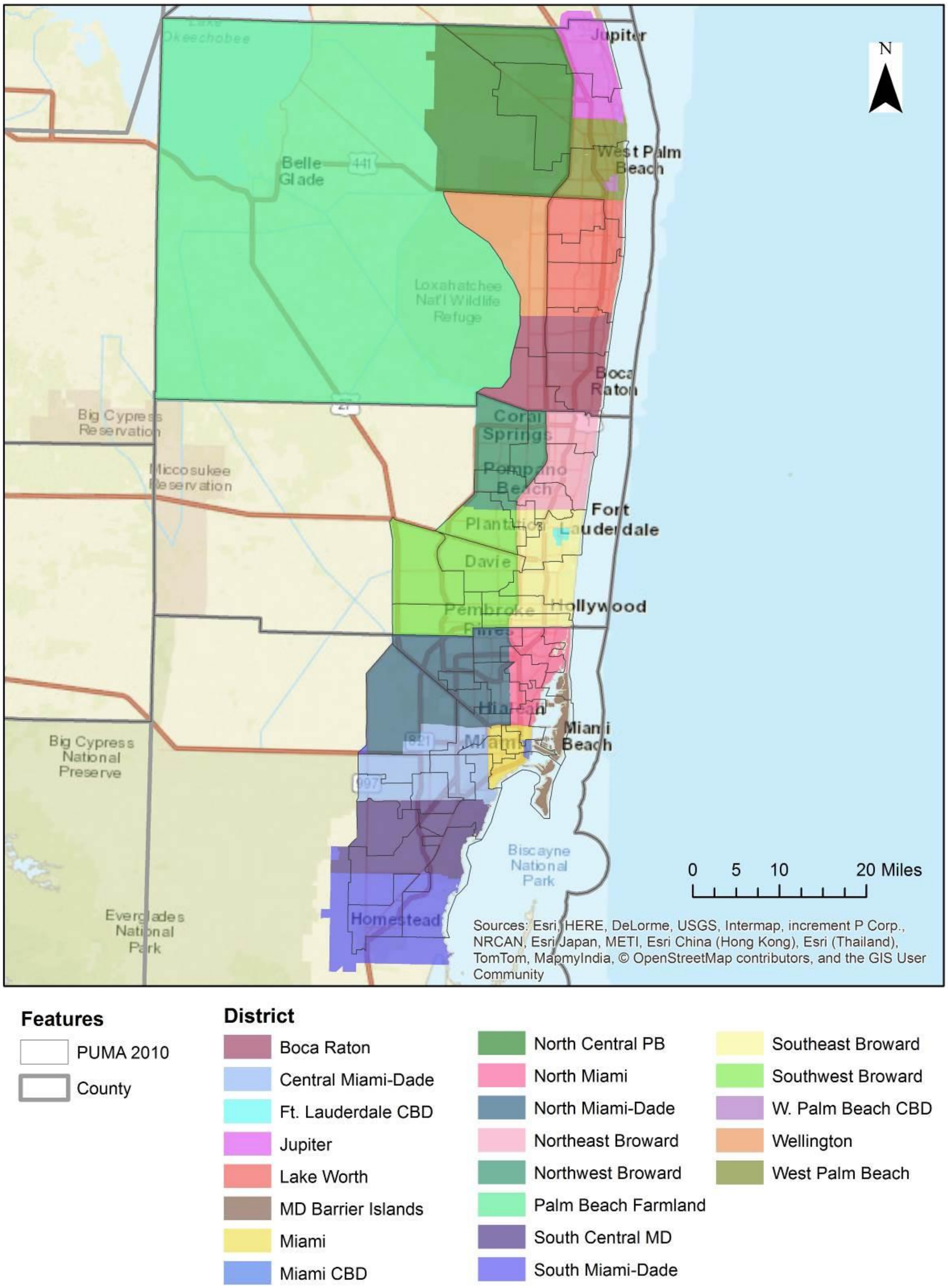
The accuracy of the validation data coupled with the ability to assimilate the information being provided by a validation test at any specific aggregation level must be considered. For example, any model that has a distribution component (e.g. regular workplace location, regular school location, work tour destination choice, etc.) could include validation tests at district level, provided reasonable validation data can be developed and the results can be reasonably interpreted. Two issues must be considered:

- If the “observed” data are developed from a summary of the household survey, the estimate of the “observed” interchange might not be statistically significant. Smith (1979) provides a good summary of the impossibility of building statistically significant trip tables at anything beyond the most rudimentary level. For regular workplace location, it might be possible to construct reasonable home versus workplace matrices from three or five year CTPP data from the American Community Survey (ACS).
- Reviewing raw validation results for a large number of interchanges might be overwhelming (“you can’t see the forest for the trees”). District/CBD to district/CBD distribution summaries would produce values for 441 interchanges and twice that amount if a direct comparison of modeled to observed data was attempted. Thus, innovative approaches might be required such as calculating percent differences and shading the differences in the tables to look for outliers or patterns.

Disaggregate validation tests are typically performed in conjunction with model estimation. Review of the results coupled with stratifying the results by different strata can help guide model estimation. While the tests are typically performed as part of model estimation, it is important to summarize, review and document the results as part of the validation process.

Disaggregate validation tests are expected to reveal, at best, moderate success; low prediction success is the more common result. If high prediction success results are obtained when low or moderate results are expected, a review of the tests should be made. Unexpected results should lead to increased investigation of the model, the validation data, or the tests: either the results are providing information that is important to know, or there is an error somewhere.

Figure 1.1: SERPM Model District and PUMA Map



The disaggregate tests should also be reviewed in light of the aggregate tests. The following matrix shows how various outcomes of disaggregate and aggregate validation results might be interpreted:

<table>
<tr><th colspan="2"></th><th colspan="2">Aggregate Validation Success</th></tr>
<tr><th colspan="2"></th><th>Low</th><th>High</th></tr>
<tr><th rowspan="2">Disaggregate Prediction Success</th><th>Low</th><td>Caution: model improvement might be possible</td><td>A likely result, but model might not be as sensitive as desired</td></tr>
<tr><th>High</th><td>Not an expected result; model sensitivity might be overstated</td><td>Possibly too good to be true–double and triple check the results</td></tr>
</table>

1.3.3 Sources of Error

As documented in the Validation Manual, there are several types of error that can affect models and model validation, including:

- Model specification error;
- Model estimation error;
- Model aggregation error;
- Input data error; and
- Model validation data error.

The model development process seeks to minimize the first four types of error listed above and the model validation process seeks to measure the success of the model development process in minimizing the errors. However, it is often overlooked that model validation data also are subject to error. For example, "observed" average daily traffic counts are, in effect, based on surveys of traffic. Most are estimated from actual traffic counts performed on one or two days over the course of a year and factored to "average daily traffic."

All of the above types of error will affect validation results. At some levels of aggregation, the impacts might be significant. For example, at a disaggregate validation of individual mode choice behavior, the impact might be substantial and an R^2 (or, more appropriately, a rho-square value) of 0.2 might represent a very reasonable model. For other, more aggregate comparisons, such an R^2 value might suggest that the model is not reasonably reproducing the observed data. However, in either case, it must be remembered that the observed data being used for the validation might be a source of some of the error.

When possible, raw model validation data will carefully reviewed to reduce error. For example, instead of using raw 2015 traffic counts, historical counts from 2012 through 2016 will be used to develop targeted counts for model calibration/validation. This approach should notably reduce validation data errors and also substantially enhance confidence in the model for future model applications.

2.0 Model Validation Guidelines and Recent SERPM 7 Validation Efforts

This section provides a brief review of validation procedures suggested in the Validation Manual, Validation Standards, and for FTA Section 5309 New Starts applications, along with past validation efforts for the SERPM model region. This review is intended to provide context for the validation plan outlined in subsequent chapters.

2.1 Validation Manual Guidelines

2.1.1 Validation of Input Data

The Validation Manual recommends that validation start with the inputs to the modeling process: the transportation networks and socioeconomic data. Recommended validation of the transportation networks includes careful review of coded distances, network connectivity, and network characteristics. The Validation Manual also suggests that example paths be built through the highway and transit networks and checked for reasonableness both in terms of path and travel times.

Recommended validation of socioeconomic data includes review of data at micro-analysis zone (MAZ), traffic analysis zone (TAZ), PUMA, district, and county levels. Comparisons to independently collected data such as Census data, ACS data, Public Use Micro-data Sample (PUMS), utility hook-ups, data from the Quarterly Census of Employment and Wages (QCEW), InfoGroup employment data, and school enrollment data are recommended.

Since ABMs are microsimulation-based models that use synthesized populations as input, checking the reasonableness of the synthesized population should be performed as part of the single-pass model calibration / validation. For example, population synthesis results can be compared to 5-year ACS data summaries spanning the calibration / validation year.

2.1.2 Validation of Amount of Travel

For trip-based models, the Validation Manual focuses the validation of trip generation models on measures such as:

- Person trip productions per household or per person;
- Proportion of person trips by trip purpose;
- Correlation of average modeled and observed (from a travel survey) trip rates for different geographies such as districts or area types; and
- Correlation of trips on a household basis (this test may be a disaggregate test by applying the model on observed household data or an aggregate measure by applying the model to households aggregated by the independent variables used for the model).

The above types of checks are also valid for activity-based models. Since activity-based models produce numbers of activities and tours by purpose, stops on each tour, numbers of work-based subtours, and possibly joint travel, each of these items can be summarized by market segment and compared to observed data (albeit, typically the same data that were used for model estimation). In addition, since activity-based models produce individual household and person records analogous to those found in a household survey, model results can be processed to produce traditional trip-based summaries such as person trip productions per capita and the proportions of person trips by trip purpose. These can then be compared to values from trip-based models for the SERPM region or to typical values for similar regions.

2.1.3 Validation of Trip Distribution

For trip distribution-type models, the aggregate checks recommended in the Validation Manual are focused on averages and frequency distributions of travel or travel related information by different length measures. For an ABM, the relevant models with length/distance components include regular workplace location, regular school location, tour primary destination choice, and intermediate stop location choice. Recommended validation measures include:

- Comparison of modeled and observed average lengths/distances by trip purpose and income group (or other socioeconomic group);
- Comparison of modeled and observed length/distance frequency distributions by trip purpose and, possibly, calculation of the "coincidence ratio;"
- Checking the percent of intrazonal location choices; and
- Comparison of modeled and observed location choices by purpose at an aggregate level such as county to county, or districts to districts.

2.1.4 Validation of Mode Choice and Auto Occupancy

The Validation Manual emphasizes the disaggregate validation of mode and auto occupancy choice models via application of the estimated choice models to observed choice data from a survey. If possible, the survey data used for the disaggregate validation should be an independent subset of the survey data used to estimate the models – in other words, the survey data would be randomly divided into two groups with one being used for model estimation and the other being used for disaggregate model validation. Practically, there is rarely sufficient data available to perform such an estimation/validation process. Because of this, the Validation Manual suggests validating using the data used for model estimation, but stratifying the validation by different socioeconomic or impedance values such as:

- Household characteristics such as household size, income level, number of workers, and auto ownership;
- Traveler characteristics such as age, gender, driver license status, and employment status;
- Zonal characteristics such as geographical location, area type, population density, and parking costs; and
- Trip/tour characteristics such as trip/tour distance, time, cost, and purpose.

The Validation Manual also suggests that model sensitivities be checked by reviewing direct- and cross-elasticities of the model coefficients. The elasticities can be compared to those reported elsewhere, are derived from other models, or have been determined empirically.

Similar types of checks can be made for auto ownership models included in ABMs.

2.1.5 Validation of Time of Day of Travel

For time of day choice, the Validation Manual recommends that modeled percentages of tours and intermediate stops by purpose and time period be compared to observed percentages. Such checks can also be performed by market segments. In addition, time of day for tours is represented by an arrival and departure time. The combination of arrival and departure times implies a duration, and comparisons of observed versus modeled tour durations should also be checked.

Disaggregate validation are also recommended for time of day choice models in the Validation Manual, although such tests are typically limited to reapplication of the choice models to the data used for model estimation. Using such an approach, aggregations of modeled and observed trips by time of day for different market/traveler segments such as worker status are recommended. In addition, the Validation Manual recommends sensitivity tests of the time of day model to verify its sensitivity to travel times and costs during specific periods in the day (e.g., morning and evening peak periods).

2.1.6 Validation of Assignment Procedures

Traffic and transit assignments represent the culmination of the modeling process. From the standpoint that the inputs to the assignment processes are based on the previous steps in the process, the assignment validations have often been used to represent a validation of the entire modeling process. Validation has traditionally been focused on the reproduction of traffic volumes and transit line boardings. More recent validations have also focused on the reproductions of reasonable speeds on roadway facilities.

The Validation Manual suggests that vehicle-miles of travel (VMT) be summarized for the region, per household, and per capita. The region-wide summaries should be by facility type with comparisons to VMT summaries obtained from regional traffic count programs.

The Validation Manual suggests checking the modeled-to-observed traffic volumes on a more disaggregate basis after the regional VMT and per household or per capita VMT estimates have been deemed acceptable. The following measures are suggested:

- Modeled versus observed volumes by screenline;
- Modeled versus observed volumes for all links with counts;
- Coefficients of determination (R^2) by link type (e.g. Functional class or volume group); and
- Root mean squared errors (RMSE) or percent RMSE (%RMSE) by link type.

The Validation Manual also suggests that speeds be reviewed for reasonableness. The Validation Manual suggests summarizing link speeds by facility type and area type, showing the minimum, maximum, and average speeds for each category. It also suggests comparing the assigned speeds with speeds used for distribution and mode choice and comparing estimated to observed speeds by highway segments, if the observed data are available.

As described in Section 4.1.3, roadway speed data are now available for the National Highway System from the FHWA's National Performance Management Roadway Data Set (NPMRDS). The final assigned travel speeds depend on input free-flow speeds[2] in addition to the volumes of traffic to be assigned, the link capacities, and the volume-delay functions. The NPMRDS data will be processed to estimate link free-flow speeds as well as congested speeds.

For transit assignments, the Validation Manual states that the primary validation checks are modeled versus observed boardings for the region, by mode and, possibly, sub-mode, and by trip length. Optional, additional checks include modeled versus observed:

- Boardings per trip (transfer rates);
- Screenline volumes;
- Boardings by route or group of routes; and
- District-to-district transit trips.

The aforementioned transit validation checks are appropriate for general model validation. However, if use of the model for FTA New Starts applications is anticipated, validation guidelines recommended in FTA New Starts workshops and the FTA website should be considered. Those guidelines are summarized in the following section.

2.2 FTA New Starts Guidelines

The FTA has an active interest in encouraging the development of travel demand forecasting models that provide useful information on potential transit markets. Part of this interest arises from the fact that travel demand forecasting models provide key inputs to the FTA's evaluation of new fixed guideway transit projects funded under the Section 5309 New Starts program. FTA is also interested in having non-New Starts transit planning being based on credible forecasts of future transit ridership demand. While the MAP-21 legislation is prompting FTA to develop new guidance for forecasts, FTA guidance from past years is still relevant to model validation.

FTA's guidance on forecasting models has been communicated to the profession through a series of workshops held between 2004 and 2009. FTA's guidance related to model development and validation falls into four basic themes:

- Overall expectations;
- Supporting data;
- Model structure and parameters; and
- Requirements for model validation.

Each aspect of FTA guidance is described below. Each section includes one or more source references describing the workshop year and session number where the guidance was presented. For instance (2009, #1) refers to the slide presentation for Session #1 occurring in the 2009 workshop as documented on the FTA website (Federal transit Administration, 2006/2007/2009).

[2] Some regions use the results of a previous forecast for the region, for the same or approximately the same forecast year, as the starting speeds for the traffic assignment process in order to reduce the number of iterations of both traffic assignment equilibration and speed feedback for the full modeling process.

2.2.1 Overall Expectations

FTA is interested in reliable information for New Starts projects that provides a useful, big-picture understanding of potential New Starts projects (2009, #1). FTA desires insights rather than just numbers, meaning that the forecasts must provide information on:

- Nature of the transportation problems for specific travel markets;
- Ability of the alternatives to improve transit service;
- Ridership response for specific travel markets; and
- Benefits accruing to those markets (2009, #2).

Since 2009, FTA has required that travel forecasting models be tested to confirm that the models properly describe actual ridership patterns. To be meaningful, these tests must include:

- Identification of key transit travel markets in current data;
- Focused testing of a model's grasp of key markets;
- Detection and correction of actual sources of error; and
- Tests over both time and transit system changes (2009, #2).

2.2.2 Supporting Data

FTA requires data that describe the current role of transit serving the region's mobility needs (2009, #4). These data have been required since May 2009 as a condition for entry into Preliminary Engineering/Project Development. The data support model development and Before-and-After Studies. The data must be suitable for the purpose of identifying the characteristics of key transit ridership markets including geographic location, trip purpose and access mode. Typically, these data are collected using an on-board rider survey although intercept or telephone surveys may be used if circumstances warrant.

FTA is interested developing data that represent the full universe of transit riders. That means that survey data collection must account for non-response biases that may occur among different trip purposes, socioeconomic classes, times of day, and length of time spent on the surveyed vehicle. The traditional (and minimally acceptable) approach for survey expansion is to develop a separate factor for each combination of route, time period, and direction. FTA prefers survey expansion strategies that use a unique factor for each train/bus route at each station/stop with aggregation only where needed to avoid very large factors. Ideally, expansion accounts for both boarding and alighting locations. Data collected should include:

- Trip origin and destination locations
- Trip origin and destination activity/purpose
- Transit access mode
- Park-and-ride location
- Full set of transit lines used along with boarding and alighting locations from origin to destination
- Traveler characteristics such as driver's license status, age, work/student status, and gender
- Household characteristics such as household size, number of workers, vehicle availability, and income

Finally, FTA expects that the data have been subjected to quality control checks to confirm the reasonableness of the collected data.

2.2.3 Model Structure and Parameters

FTA has issued guidelines regarding the expected structure of travel forecasting models (particularly mode choice) and model coefficients and the relationships among the coefficients (2006, #15). While the FTA will allow coefficients outside the specified ranges, they require "compelling" evidence supporting the out-of-range coefficients. FTA concerns related to model structure and parameters include:

- **Unusual model coefficients**. FTA guidance suggests that unusual model coefficients implying unusually high or low sensitivities should be avoided. The coefficient of in-vehicle travel time should be between -0.02 and -0.03 and the out-of-vehicle travel time coefficient(s) should be between 2.0 and 3.0 times the coefficient of in-vehicle travel time.
- **Non-logit decision rules**. Non-logit decision rules include model parameters that are used when a variable exceeds (or is less than) a specified level, such as minimum transit travel times or auto access times that must be less than transit in-vehicle travel time. These rules should be avoided wherever possible because they can cause sudden changes in behavior when trip characteristics cross an arbitrary threshold.
- **Bizarre (over-specified) alternative-specific constants**. Bizarre, unexplainable alternative-specific constants should be avoided. An example of bizarre alternative-specific constants is a series of constants stratified by income group that are not monotonically changing in magnitude as the income group changes. Over-specified constants might be characterized by constants that vary by mode for, say, each type of area type-to-area type interchange.
- **Alternative-specific constants for fixed guideway transit modes**. This issue relates to the magnitude of the rail mode alternative-specific constant in relation to the local bus constant. The FTA has indicated that the implied travel time savings for the rail alternative-specific constant should be less than 20 minutes and, preferably, less than 12 minutes.
- **Inconsistencies between path and mode choice models**. Conformance between parameters in transit path selection and mode choice utility expressions for transit choices is a primary concern. This issue has been a concern in the development of transit impedance variables for the model estimation. However, once the estimation is complete, the conformance of the relationships must be verified.
- **Accuracy of bus running times**. The level-of service matrices must match the actual baseline (and build) conditions. This implies that bus speeds should be related to auto speeds when buses run in mixed flow and, in exclusive lanes or on exclusive guideway, bus and fixed guideway travel times must reflect reasonable operating characteristics and schedules.

2.2.4 Requirements for Model Validation

Perhaps the most important element of FTA guidance on model development relates to good practices in Model Testing (2009, #7). FTA guidance states that conventional modeling practice is inadequate for developing models suitable for New Starts forecasting. Under the conventional approach, models are estimated (or asserted) to match aggregate measures of trip rates, trip length distributions and mode shares. Models are then adjusted with K-factors or mode-specific constants and results are checked against traffic volumes or transit line boardings. As needed, adjustment factors are used to improve the correspondence between model results and observed traffic or ridership. FTA expectations for model testing are considerably more involved than the model validation approach typically employed. Model testing involves the following four key activities:

- **Data Matching.** All aspects of the model (input data, intermediate computations, and output results must match the observed supply and demand characteristics of the regional transportation system. That means that during calibration, the transit components should be assessed to confirm the model's understanding of the transit network (line-haul and access components), appropriateness of the shortest paths, and the ability of the mode choice model to represent the specific markets that are attracted to transit. Since these components are dependent on other model elements such as highway path building, trip generation, and trip distribution, these other components also need to match observed conditions.
- **Model Assessment.** The resulting model must be assessed in terms of its ability to be explainable. Parameters such as trip rates, distribution patterns, mode choice coefficients, and mode-specific constants must have specific behavioral explanations. The relative values of these parameters must be sensible across socio-economic classes, travel modes, and other segmentations.
- **Model Testing.** Models must be subjected to meaningful testing that includes more than simple parametric sensitivity testing or replication of calibration year data. Ideally, models should be tested against a different year that spans a major change in the transit system. If such a change is not available, then testing should include an assessment of the reasonableness of a horizon-year forecast with a major transit alternative.
- **Documentation.** To be usable, a model must have complete model development documentation and a user's guide. Beyond these requirements, FTA expects a formal assessment of model plausibility, results of forecast testing, and an assessment of what markets or situations the model can, and cannot, effectively forecast.

The FTA emphasis in its guidance is that the focus of model testing should be on model performance supporting the understanding of markets for transit rather than matching statistical validation standards. That means that model testing should focus on:

- **Transit Travel Patterns.** FTA is most interested in evidence that the model understands the circumstances under which transit can successfully attract a market. FTA tests include a comparison of observed and modeled district-to-district flows by mode and market segment, production and attraction totals, and transfer rates.
- **Aggregate Transit Paths.** FTA guidance suggests that an expanded (observed) transit trip table be assigned to the underlying transit networks to confirm that both the survey data base and the transit networks properly represent observed conditions. Assigned and observed line loads, on/off distributions, time of day patterns, modes of access/egress, park-ride usage, and walk distance distributions should be tested.
- **Disaggregate Transit Paths.** FTA encourages development of prediction success tables that check the consistency between the path builder and observed behavior of individual responses from the transit rider survey. This includes the sequence of mode (e.g., bus-rail-bus), the number of transfers, and park-ride locations.

2.3 Past SERPM 7 Validation Efforts

The existing SERPM 7 model was validated to year 2010 using many of the checks recommended in the Validation Manual and detailed above. The SERPM 7 model was based on the Florida Add-On 2009 NHTS, transit on-board survey, and 2010 traffic counts. The results of the model validation are documented in the Model Development Report[3]. This review of the SERPM 7 model validation effort provides context for the validation of the SERPM 8. Since the SERPM 7 was calibrated and validated to reproduce observed

[3] Southeast Florida Regional Planning Model – SERPM 7.0 – Coordinated Travel – Regional Activity Based Modeling Platform (CT – RAMP) – Model Development Report – Final Report, February 2015.

conditions in 2010, it can be expected that the SERPM 8 should produce similar rates and validation statistics.

2.3.1 Input Data

The SERPM 7 zonal input data include household and population socioeconomic data, employment data, school enrollment data, parking supply data, hotel and motel room data, airport enplanements data and vehicle volumes at external stations. The Model Development Report summarized person and household socio-economic input data, employment and student enrollment input data by counties and for the entire model region.

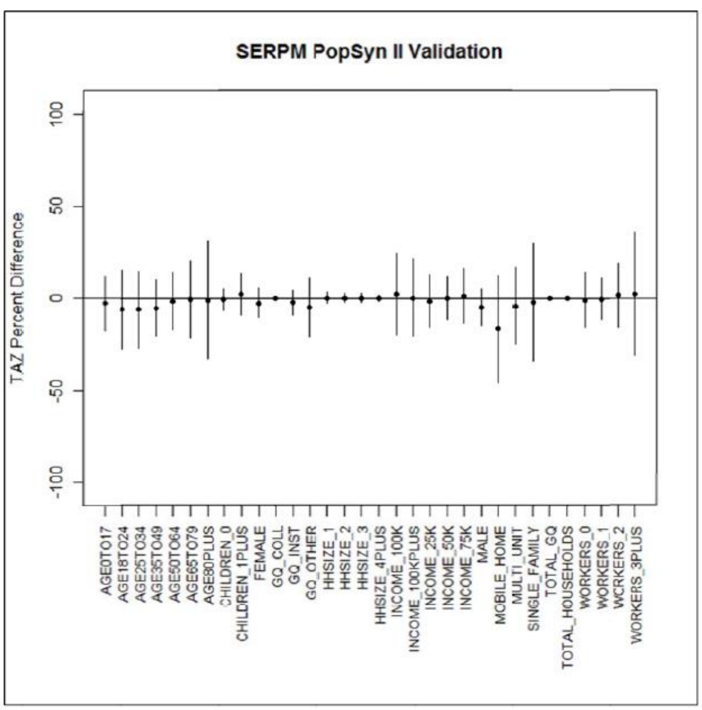
The highway network includes all streets for 8 facility types from freeway to collector. The report conducts a comparison of model estimated free-flow speeds (average posted speed, free-flow speed, and average speed) as well as RMSE to observed INRIX speed data by facility type and area type. A comparison of bus travel times obtained from 2010 timetables to estimated travel times for each county in transit network is also provided in the report.

2.3.2 Core ABM Demand Sub-Models and Procedures Validation

Population Synthesis

In SERPM 7, a synthetic population is created using the PopSynII model by disaggregating zonal households and person records to MAZs. The synthetic population attempts to match zonal level marginal distributions of households and person records by various characteristics. The report presents the households/persons contribution towards the controls by household/person characteristics, as shown in Figure 2.1.

Figure 2.1 Base Year Synthetic Population Validation



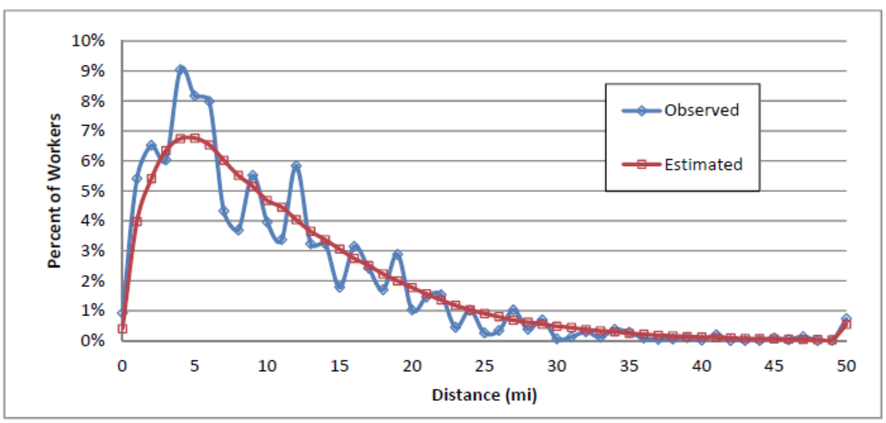
Long term level

The Usual Workplace and School Location Choice at long term level includes 4 submodels for work, grade school, high school, and college. A multinomial logit destination choice model assigns a regular workplace TAZ and MAZ for every workers in the synthetic population. The out-of-home usual workplace location choice model is calibrated targeting at matching NHTS origin-destination distance frequency distribution and the 2010 CTPP commuting flow patters. The average home to work distance by worker class is shown in Figure 2.2. The validation report also provides the comparison of estimated and observed commuting flows at county level, average out-of-home work location distance, the comparison of estimated and NHTS trip length frequency distribution (Figure 2.3), and the estimated and CTPP worker flow comparison. Similar validations are done for school location choice model.

Figure 2.2 Average Out-of-Home Work Location Distance

Worker Class	Observed	Estimated
Full-Time Workers	10.6 miles	10.7 miles
Part-Time Workers	7.5 miles	7.6 miles
Income less than $30k	7.1	8.3
Income $30k - $60k	9.3	9.2
Income $60k - $100k	10.5	10.4
Income more than $100k	12.5	11.7

Figure 2.3 Work Location Choice Distance Frequency Calibration



Mobility Level

There are three sub-models in the mobility level: the free parking eligibility model, household car ownership model, and transponder ownership for use of toll lanes model. The car ownership model is a nested logit choice model which predicts the number of autos available in a household, five alternatives of the model are presented in Figure 2.4. The car ownership model estimation results are validated to the ACS 2006-2010 observed data by household size, household income, and household workers. The auto ownership by household size validation results are shown in Figure 2.5. Total cars by residence TAZ is also validated to registered personal use vehicles data obtained from Florida Department of Motor Vehicles as presented in Figure 2.6.

Figure 2.4 Auto Ownership Nesting Structure

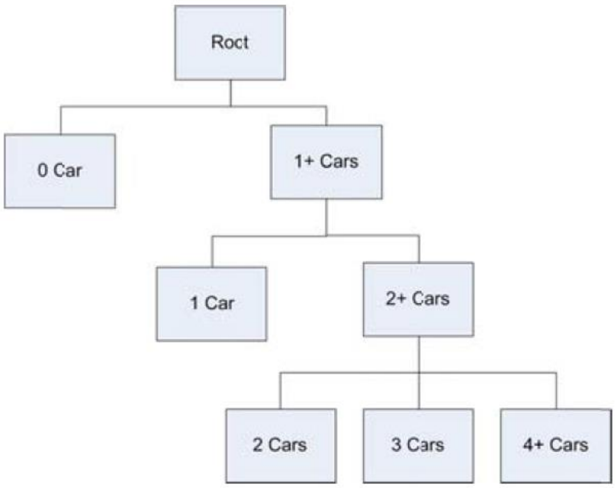
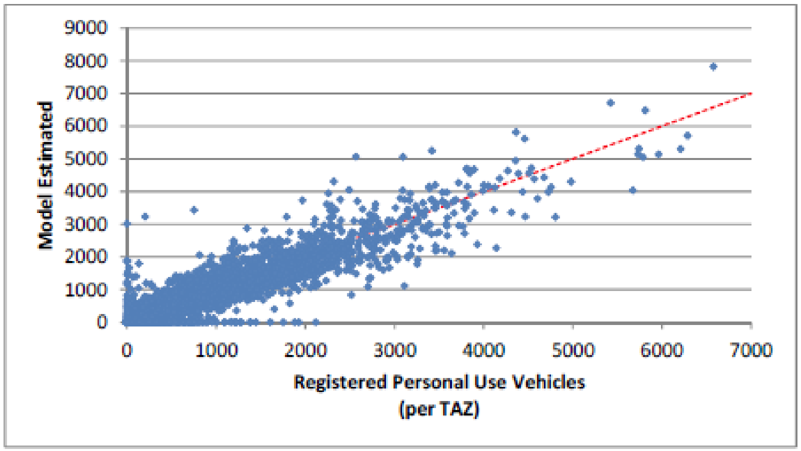


Figure 2.5 Auto Ownership Model Validation to Household Size

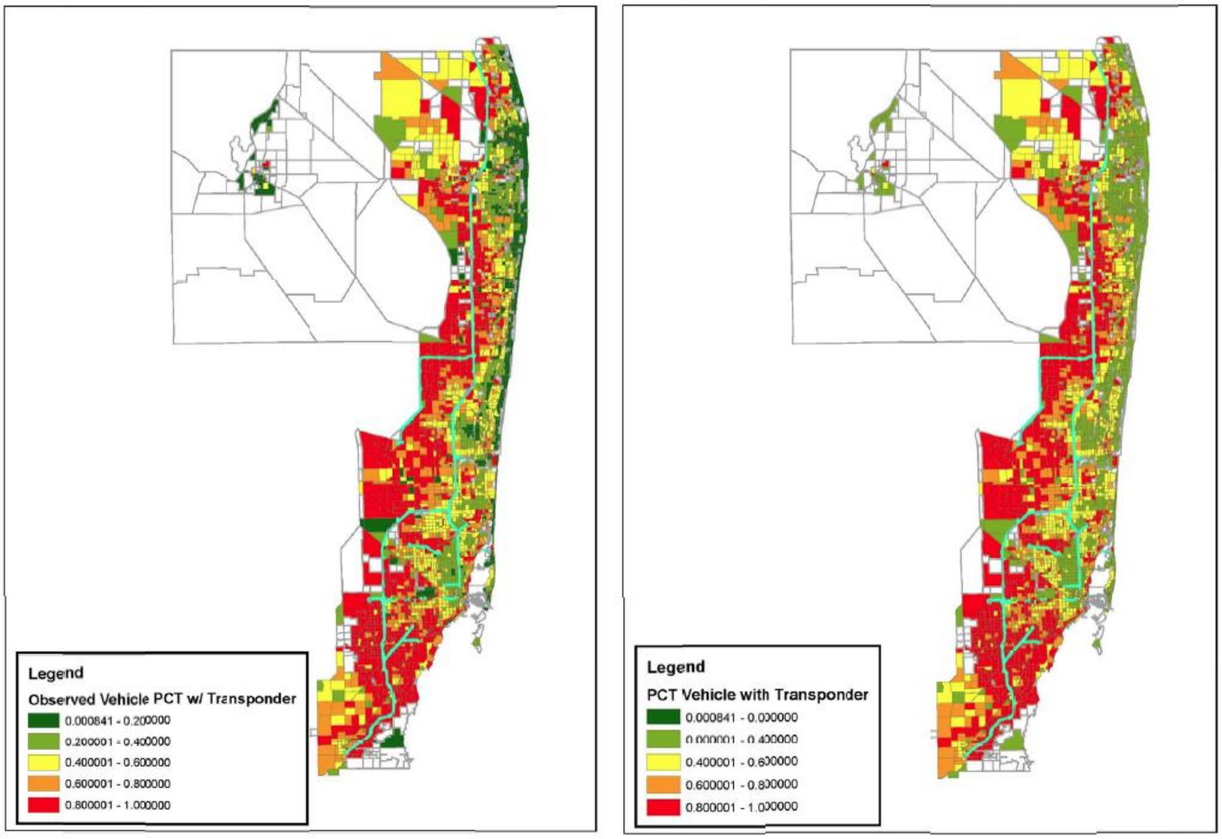
	Auto Ownership	Household Size 1	2	3	4	5+
Observed ACS 2006-2010	0	18%	6%	4%	3%	4%
	1	73%	37%	26%	17%	17%
	2	7%	52%	46%	51%	39%
	3	1%	5%	21%	20%	24%
	4+	0%	1%	3%	9%	16%
	Total	100%	100%	100%	100%	100%
Estimated	0	15%	8%	7%	4%	4%
	1	83%	29%	27%	18%	14%
	2	2%	56%	46%	48%	44%
	3	0%	7%	18%	20%	23%
	4+	0%	1%	3%	10%	15%
	Total	100%	100%	100%	100%	100%

Figure 2.6 Validation of Auto Ownership Model to 2010 Registered Vehicles



The multinomial logit Toll Transponder Ownership model estimates toll transponder (SUNPASS) ownership of a household. The observed and estimated share of SUNPASS equipped vehicles are mapped in Figure 2.7.

Figure 2.7 Transponder Ownership Validation



Daily pattern/schedule level

The daily pattern/schedule level has five main sub-models: the Coordinated Daily Activity Pattern (DAP) model, Individual mandatory activities/tours model, Joint travel tours model, Individual non-mandatory activities/tours model, and Individual at-work sub-tours model.

The Coordinated Daily Activity Pattern (DAP) model is a multinomial logit model which models personal DAPs and the generation of individual tours. The DAPs include three patterns: mandatary patter (M), non-mandatory pattern (NM), at-home pattern (H). The DAP choice structure is presented in Figure 2.8. The validation report provided the comparison results of estimated and observed DAP share by person type (Figure 2.9).

The individual mandatory time of day choice model is validated by comparing estimated and observed tour departure and arrival time distribution by work, university, and school. The joint tour sub-model are validated to observed share of joint tour by purpose.

Figure 2.8 DAP Type Choice Structure

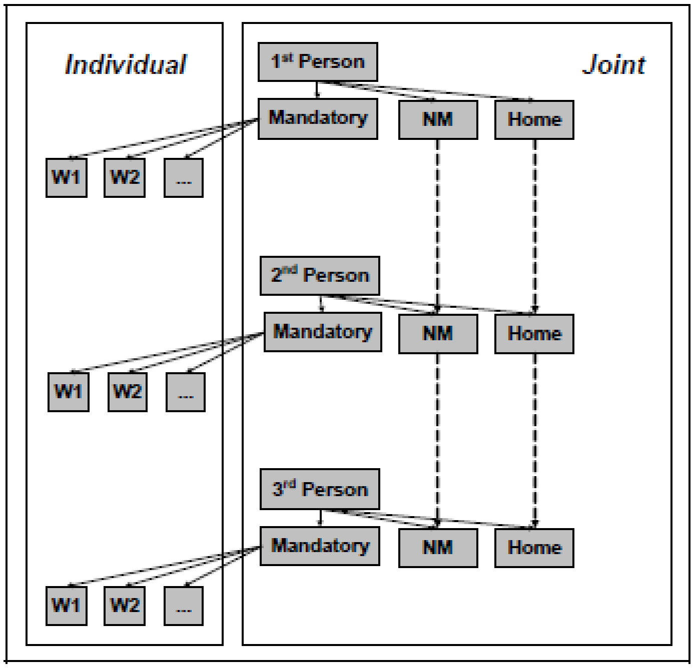
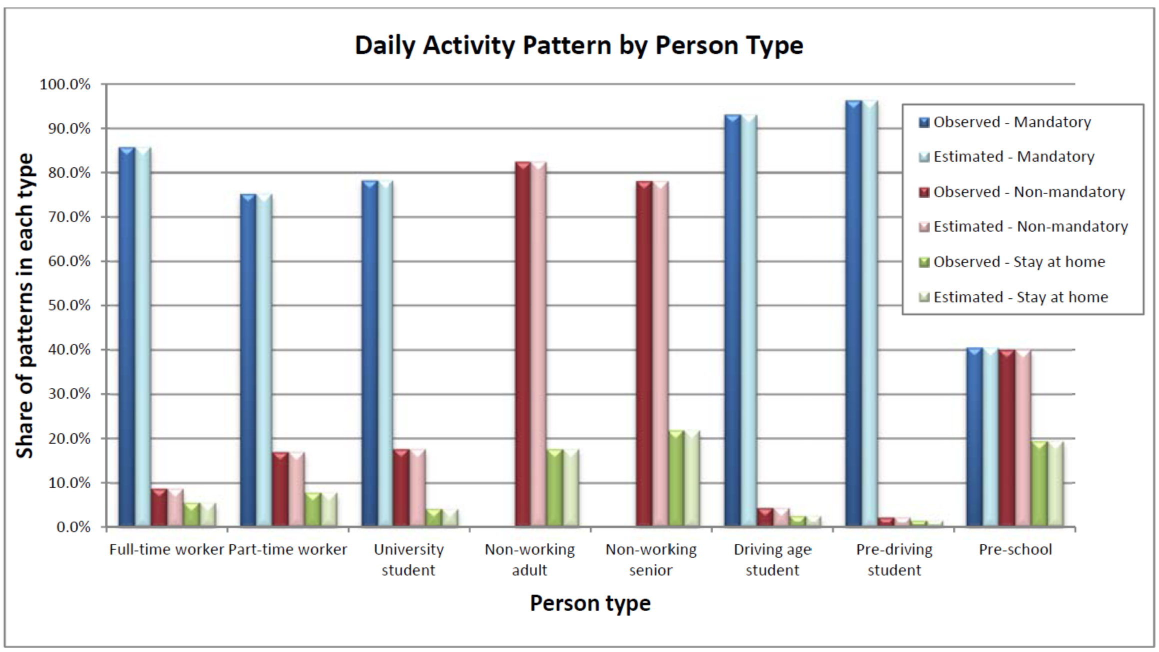


Figure 2.9 Daily Activity Pattern by Person Type



Tour Level

Tour mode choice model, intermediate stop frequency model, intermediate stop location choice model are the three sub-models at the tour level. The Tour Mode Choice Model predict the main tour mode used to travel from origin to primary destination, tour modes included in the model are presented in Figure 2.10. Estimated tours by tour mode and purpose are validated to observed data, Figure 2.11 provides validation results for work tours.

Figure 2.10 Tour Mode Choice Model Structure

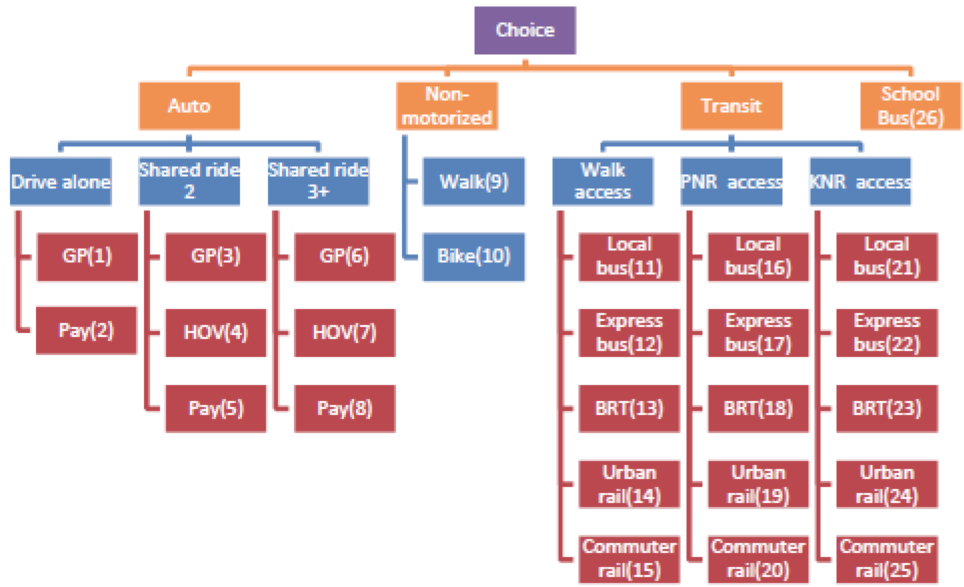


Figure 2.11 Work Tours by Tour Mode and Purpose

Tour Mode	Observed Tours				Estimated Tours			
	No Veh	Veh<Adult	Veh>=Adult	Total	No Veh	Veh<Adult	Veh>=Adult	Total
Drive-Alone	-	278,860	1,067,131	1,345,991	0	279,710	1,069,590	1,349,300
Shared 2	16,003	171,407	183,676	371,086	15,389	170,494	182,393	368,276
Shared 3+	9,200	65,020	83,057	157,278	8,430	64,938	83,867	157,234
Walk	13,267	18,502	3,389	35,158	13,120	18,429	3,160	34,709
Bike	5,772	463	3,201	9,436	6,250	640	3,050	9,940
Walk-Transit	30,219	29,686	13,125	73,030	31,099	29,929	12,490	73,517
PNR-Transit	-	3,349	12,323	15,672	0	3,470	11,410	14,879
KNR-Transit	908	3,022	2,508	6,438	1,050	2,820	2,380	6,250
School Bus	-	-	-	-	0	0	0	0
Total	75,370	570,310	1,368,410	2,014,090	75,337	570,429	1,368,339	2,014,105

The Intermediate Stop Frequency model estimates the number of intermediate stops between origin and primary destination, assigns trip purpose to each stop. Intermediate stop frequencies are validated to NHTS data by tour purposes. The intermediate stop location model then determines the location of intermediate stops. The validation provided trip length distribution of out-of-direction of calibration results and NHTS observed data.

Trip Level

The mode used for each trip along the tour is modeled by the Trip Mode Choice Model. Trip mode switch distribution is validated to NHTS and transit board surveys, below provides the differences between observed and estimated transit trips by trip mode and tour mode.

Figure 2.12 Observed and Estimated Transit Trips by Mode

	Transit Trips by Trip Mode				
	Local Bus	Express Bus	BRT	Urban Rail	Commuter Rail
Work Tours					
Observed Trips	124,547	7,425	2,843	35,482	8,473
Estimated Trips	124,125	7,809	4,230	37,047	8,424
Difference	*(422)*	*384*	*1,387*	*1,565*	*(49)*
University Tours					
Observed Trips	20,342	485	442	6,749	1,276
Estimated Trips	22,151	687	554	5,172	1,182
Difference	*1,809*	*202*	*112*	*(1,577)*	*(94)*
School Tours					
Observed Trips	24,313	368	400	4,388	1,157
Estimated Trips	25,491	369	489	3,787	920
Difference	*1,178*	*1*	*89*	*(601)*	*(120)*
Maintenance Tours					
Observed Trips	95,386	1,040	4,072	6,584	1,106
Estimated Trips	89,820	861	2,606	7,168	1,999
Difference	*(5,566)*	*(179)*	*(1,466)*	*584*	*893*
Discretionary Tours					
Observed Trips	10,598	116	452	732	123
Estimated Trips	10,662	96	261	642	202
Difference	*64*	*(20)*	*(191)*	*(90)*	*79*
Work-Based Tours					
Observed Trips	8,781	178	196	1,603	36
Estimated Trips	10,600	48	186	615	28
Difference	*1,819*	*(130)*	*(10)*	*(988)*	*(8)*
All Tours					
Observed Trips	283,968	9,611	8,406	55,537	12,170
Estimated Trips	282,849	9,870	8,326	54,431	12,755
Difference	*(1,119)*	*259*	*(80)*	*(1,106)*	*585*

2.3.3 Base Year

Highway Assignment

The highway assignment model is validated to traffic count database and screenlines, e.g. facility type validation for the southeast Florida region as shown in Figure 2.13.

Figure 2.13 Facility Type Validation, SE Florida Region

Facility Type	Model RMSE	Estimated Volume	Count Volume	Estimated To Count Ratio	Number of Observations
Freeways	20.2%	17,981,475	18,032,571	1.00	235
Uninterrupted Roadways	46.5%	2,235,035	1,942,941	1.15	213
High Speed Arterials	33.8%	53,192,393	52,173,198	1.02	3,673
Low Speed Collectors	57.3%	9,699,909	10,591,829	0.92	1,695
Ramps	64.2%	7,546,736	6,930,200	1.09	686
HOV Lanes	25.4%	1,748,683	1,721,433	1.02	82
Toll Roads	22.5%	5,475,976	5,732,702	0.96	178
HOT Lanes	24.2%	539,538	488,660	0.90	22
All Groups	38.6%	97,880,207	97,148,324	1.01	6,767

Transit Boardings Validation

The transit assignment is done for eight transit sub-modes that include a combination of model and access. These are: walk to bus, express bus, rail and commuter rail, drive to bus, express bus, rail and commuter rail. The transit assignment validations are done by mode and operator. Metrorail boardings by station, access distance distribution, and access/egress mode share to stations are validated to on-board survey results. Tri-Rail boardings are validated to on-board surveys for average daily boardings by station and drive access distance distribution.

3.0 Validation Process

The validation process of an ABM is much more elaborate, comprehensive and data intensive when compared to that of a four-step trip-based model system. CS proposes to conduct the SERPM 8 validation task in various steps as described below.

3.1 Validation Data Compilation

The first step of this process is data assembly and assessment. This will include reviewing and summarizing all kinds of travel survey data (household travel, transit, etc.), count data (by facility type, vehicle class, transit boardings, etc.), speed data, and any other publicly available data (census, ACS, etc.). This review is essential to ensure there is sufficient data to validate the SERPM system. It also provides us with a logical way of compiling data for the various model components. At this step, validation targets will also be set for various model components. This is described in more detail under Chapter 4.0 of this report.

3.2 Single-Pass Calibration of Model Components on 2015 Base Year

The estimated models will be applied in a single pass using congested speed network data used for the model estimation process. In other words, single-pass application will not involve any feedback loops during the model calibration process. Model constants will be adjusted, if necessary, to better replicate expanded data from the household and transit on-board survey data. The single-pass process involves the following steps:

- **Apply the estimated models** as is using skims and socioeconomic data used for model estimation – This will be used to set a benchmark for all the subsequent model calibration work. The applied model results will be compared to the validation targets for every model component to see how well the model is performing.

- **Calibration of individual core ABM components** – Based on the above step, every component will be calibrated by adjusting constants, revising coefficient values, and re-estimating certain models to include new variables, if necessary. This is described in more detail in Chapter 5.0 of this report.

- **Examine error propagations** – As all the ABM components are linked to one another and applied in sequence, each subsequent model component is affected by models upstream. Doing a single-pass, stepwise calibration therefore helps understand the magnitude and direction of error propagation through the model system.

The CT-RAMP model application process will add some challenges and an opportunity to the single-pass, stepwise calibration process. Specifically, CT-RAMP applies the ABM using a “depth-first” approach rather than a “breadth-first” approach. The depth-first approach applies the entire modeling sequence for each household and its members before proceeding to the next household in the synthetic population. In a breadth-first approach, each model component is applied for the entire population prior to proceeding to the subsequent model component. The challenge of the depth-first approach will be computer processing time since uncalibrated model components lower in the modeling sequence must be applied each time an upper level model component is calibrated.

The opportunity provided by the depth-first approach is the ability to monitor the effects of the calibration process on error propagation. Several key statistics will be selected for each model component (e.g. average tour length) and monitored regarding whether they diverge from or converge to the observed statistics as upper level model components are calibrated. The divergence or convergence of the key statistics for lower level model components will be considered in the calibration of each upper level model component only if it appears that substantial corrections will be necessary as the stepwise calibration proceeds to the lower level components.

This task will be done for a base year of 2015 with all 2015 inputs – synthetic population, zonal socioeconomic data, networks and skims.

3.3 Full Feed-Back Validation to 2015 Base Year

After all the individual ABM components are calibrated, a full model run with feedback loops will be run to evaluate the impact of feedback loops on the calibrated components. Feedback loops enable the model system to use congested characteristics on travel behavior, that is, the input data (times, etc.) are as realistic as possible. This will affect the number, destination, mode, timing and routing of tours and trips. The SERPM region includes large cities and some of the state's busiest highway travel corridors. Having adequate feedback loops in the model system is essential to obtain accurate estimates of air quality, transit, HOV and studying peak spreading in the region.

There are several effects of feedback loops on the ABM components:

- **Tour generation** – The feedback of congested skims affects accessibility measures that influences a person's tour generating behavior. This could either result in fewer or more trips depending upon the OD pair. It could also affect trip chaining and tour formation.
- **Daily activity patterns** – The accessibility to various activities and opportunities is affected by congested travel times, which will in turn affect the daily activity patterns of travelers.
- **Destination/location choices** – The congested skims influence the trip lengths directly, which affects the distribution patterns of tours and trips.
- **Mode choices** – Mode choice is directly affected by the feedback of congested times.
- **Time-of-day** – Congestion causes people to alter time of travel, which induces peak spreading. So these models also are affected by the feedback loops.
- **Highway and transit assignments** – These are affected by the equilibration process of the iterations within each loop as well as between subsequent loops.

Once the number of feedback loops is determined, the key results from each ABM component will be compared against the observed (survey) data as well as the results of the single-pass validation. While the results of these comparisons may suggest that additional calibration of ABM model components is warranted, it is more likely that adjustments of the assignment process or volume-delay functions used for the determination of congested highway speeds are necessary since the single-pass calibration / validation process used congested highway skims (based on observed speeds) as input. The full-feedback validation will focus on how well modeled assignment results match observed traffic counts and transit boardings.

3.4 Backcast Using SERPM 8

As part of the sensitivity testing, a backcast to 2010 using the validated SERPM 8 will be made as a key check on its performance. The backcast year to 2010 will simplify the test since all input and validation data should be readily available for that year as a result of the SERPM 7 model calibration and validation. Several substantial transportation projects (e.g. the Port of Miami tunnel and the I-595 reconstruction / 595Express lanes) were opened between 2010 and 2015. In addition, the US was just emerging from the Great Recession in 2010; the Florida economy in 2015 was substantially better. The total population of Broward, Miami-Dade, and Palm Beach counties grew by about 7 percent between 2010 and 2015 and employment grew by about 14 percent over the same time period.

3.5 Sensitivity Testing

As part of the validation process, a series of sensitivity tests will be performed using the ABM. This will include executing a number of different scenarios with varying input parameters for several of its components. These tests will be focused on the most important factors that will have a direct impact on the projection of future travel behavior in the region. The tests are meant to examine how the ABM responds to certain plausible scenarios and the sensitivity of the model to various level-of-service parameters and other variables. A subset of the following tests will be undertaken:

- Socioeconomic and demographic factors:
  - Alternate growth rates of population, employment;
  - Alternate growth rates of different market segments such as aging of population, presence of more females in the workforce, increase in low income households, etc.
- Auto Mode Parameters:
  - Adjustments to fuel costs.
- Impact of new highway projects:
  - New managed lanes, or pricing scenarios;
  - Widening of highways.
- Impact of new transit projects:
  - Extension of rail lines;
  - Addition of new transit modes like LRT.

3.5.1 Elasticity Tests

Elasticity is a convenient, quantitative measure of travel demand response to price and service changes which influence demand[4]. For elasticity measures to be applicable, the transportation system change must be based on a relative measure. In other words, it must involve a quantifiable percentage increase or decrease in the system parameter involved. For example, while elasticity measures can be used to describe the response to a change in the overall amount of transit service, they cannot be used to describe the response to a new transit system. Transportation elasticities are informally adopted from the economist's measure "price elasticity." The price elasticity of demand is loosely defined as the percentage change in quantity of service demand in response to a 1 percent change in price. For instance, a price elasticity of -0.3

[4] TCRP Report 95, *Transit Pricing and Fares: Traveler Response to Transportation System Changes*, 2004.

indicates that for a 1 percent increase (or decrease) in the price of a service, there is 0.3 percent decrease (or increase) in the demand for that service.

A model is said to be inelastic if the absolute value of the calculated elasticity is less than 1.0 and elastic if 1.0 or greater. Roughly, an elasticity of 1.0 implies that a 1 percent increase in "price" of a "product" will produce a 1 percent increase in demand for the product. A negative 1.0 elasticity simply implies that a 1 percent increase in "price" will produce a 1 percent decrease in demand. For the SERPM 8, both direct and cross-elasticities will be estimated for certain key explanatory variables. Direct elasticities are based on the change in demand for a mode based on a change in cost (price) for that mode, while cross-elasticities are based on the change in demand for an alternate mode based on the change in cost for a given mode.

4.0 Validation Data Assessment

This section identifies the data available for validation of the activity-based model. A variety of data are needed to perform the validation tests described in Section 5.0 but those tests are limited by the available data. While some tests identified in Section 5.0 require data that are not currently available, identifying and assessing the existing data help focus Section 5.0 toward meaningful tests that can be readily performed.

4.1 Existing Data Sources

This section describes in detail the existing data and data sources available for use in validation of the ABM.

4.1.1 Socioeconomic Data

Typically, input socioeconomic data validation sources are the same as those used to develop the data. Few regions have multiple sources of the same socioeconomic data for a particular year. The main socioeconomic data available are the Census, American Community Survey (ACS), ACS Public Use Micro-data Sample (PUMS), Census Transportation Planning Package (CTPP), Quarterly Census of Employment and Wages (QCEW), and Longitudinal Employment Household Dynamics (LEHD).

Bureau of Economic and Business Research (BEBR) Population Estimates and Projections

BEBR produces Florida's official state and local population estimates and projections. These estimates and projections are used for distributing state revenue-sharing dollars to cities and counties in Florida and for budgeting, planning and policy analysis by state and local government agencies, businesses, researchers, the media, and members of the general public. The population projections by age, sex, race, and Hispanic origin for Florida and by county are now available for 2020 to 2045 with estimates for 2015.

2010 Census Data

The decennial U.S. Census provides information on all persons and households in the country and can be viewed at census block level geographic resolution (similar to the transportation analysis zone (TAZ) level). The 2010 Census focused on population and housing questions. Questions that were asked in the 2000 Census regarding income, auto ownership, and employment are now part of the ACS and were not asked in the 2010 Census, and are therefore no longer available on SF1. The 2010 SF1 data will be used to examine univariate distributions of households and persons across particular variables (e.g., households by household size and persons by sex or age).

5-Year American Community Survey (ACS) Data

The ACS is conducted continuously by the Census Bureau and provides a great deal of information that can be used for validation. Because the ACS is conducted continuously, the Census Bureau can make data available every year rather than every 10 years like the decennial census (though for a smaller samples of the population).

Rather than surveying about one in every six households once every 10 years, as had been done with the long form, the ACS samples about one in every 40 addresses every year, or 250,000 addresses every month. For areas with large populations (65,000 or more), survey estimates are based on 12 months of ACS data. The Census Bureau produces estimates for all areas, down to the census tract and block group levels, based on five years of ACS data. One and five year estimates based on survey data including the data from 2009 are currently available.

The ACS provides data on housing and population not available from the 2010 Census. Information includes:

- Population characteristics
  - Age
  - Sex
  - Relationship to head of household
  - Income
  - Employment information including labor force status, industry and occupation
  - Journey to work information
- Household characteristics
  - Vehicles available
  - Income
  - Tenure
  - Housing value
  - Rent

5-Year Census Transportation Planning Products (CTPP) Data

The CTPP is based on data from the ACS. In addition to providing information on the place of work of residents and the journey to work noted above for the ACS, the CTPP also provides zonal level information in the form of cross-classifications across variables. For instance, one could examine households by size, number of vehicles, and income rather than univariate distributions provided in the data sources described above. The 2006-2010 CTPP is the latest currently available data, the next 2012-2016 CTPP data products will likely be released in late 2018, or early 2019. The 2009-2013 county to county commuting flow data (by means of transportation) is also available in Census Bureau Commuting (Journey to Work) website.

ACS Public Use Microdata Sample (PUMS)

PUMS files from the ACS show the full range of population and housing unit responses collected on individual ACS questionnaires. For example, they show how respondents answered questions on occupation, place of work, and so forth. The records contain information from the completed ACS questionnaires for most questions for the selected subsample of housing units and group quarters persons including questions on age, sex, tenure, income, education, language spoken at home, journey to work, occupation, condominium status, shelter costs, vehicles available, and other subjects. Many multi-variate customized tabulations can be summarized from the PUMS files.

Longitudinal Employment Household Dynamics (LEHD)

The LEHD provides even more detailed data on workers and employers than the CTPP. Employer characteristics are reported for each employer with geocoded addresses. Worker/individual characteristics include wage records, personal characteristics (e.g., gender, age, ethnicity), and location of residence.

Compared with the worker flow and journey to work data from the CTPP, the LEHD data enumerates the entire workforce population who are covered by unemployment insurance (rather than a small sample), includes a more comprehensive geographic coverage including counties with low population, and includes second jobs for workers. LEHD and Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES) do not count self-employed workers, unpaid workers, or uniformed military personnel.

Quarterly Census of Employment and Wages (QCEW)

The QCEW program provides quarterly and annual establishment and employment data by county classified using the North American Industry Classification System (NAICS). The QCEW employment data covers 98 percent of U.S. jobs and includes total wages for each employment category and geographic region.

InfoGroup

The Infogroup Business USA dataset provides data on businesses in the United States, at national and subnational geographic units. Data are collected from multiple sources, including direct calls to businesses. Statistics are presented on company counts, employee counts, and sales by industry (NAICS), company, and company location.

The Florida Model Task Force created a focus group to review the pressing issues regarding the Florida region InfoGroup data the Department purchased in 2010. The issues identified lies in: natural disasters destroyed businesses, government employment, hospital coverage, big box employment, and higher education coverage.

4.1.2 Travel Survey Data

Several travel surveys are available for validation of the SERPM 8 including the 2015 Southeast Florida Household Travel Survey and transit on-board surveys.

Household travel survey

The ongoing 2015 Southeast Florida Regional Travel Survey (RTS) will collect detailed travel data from a sample of 5,000 households across Miami-Dade, Broward and Palm Beach counties. The data collected from the HTS will be used to develop activity-based travel demand models that forecast travel behavior based on population socio-demographics, land-use characteristics, and transportation services and infrastructure. The 2015 Southeast Florida Household Travel Survey results are scheduled to be released around May 2017.

Transit on-board survey

Different transit on-board survey data sources and years of data collection will be reconciled in order to develop 2015-year calibration and validation transit ridership targets for SERPM 8. The MDT Metrobus and the Metromover survey data, including the recently acquired survey data – 2013 MDT Central Garage, 2014 and 2015 MDT South Garage and the 2015 Metromover surveys, will be processed for use in developing targets. The MDT Metrobus and the Metromover survey data will be converted into Production-Attraction format (as needed for modeling purposes) and defined by access mode, trip purpose, and time periods consistent with SERPM 8. The data for other systems and agencies have already been processed under various previous studies and will be directly utilized.

For Palm Tran and MDT for which APC data is available at sufficient details, the 2015 ridership data from various sources and agencies will be collected to develop validation targets for SERPM 8. For Tri-Rail, Metromover and Metrorail, station-level boarding targets will be developed. National Transit Database (NTD) and American Public Transportation Association (APTA) ridership data will also be reviewed for various agencies and any discrepancies between various data sources in order to determine the best possible course of action. The transit surveys and 2015 expansion challenges are listed in Table 4.1.

Table 4.1 Transit Surveys

Agency/System	On-Board Survey Year	Notes on 2015 Expansion
Tri-Rail	2013	O/D information not sufficient; MIA station closed when the survey was conducted. Supplement data using the 2008 survey.
Palm Tran	2015	System-wide (or district-level) information should be generated only; may need to supplement information with the 2009 on-board survey
BCT	2010	2015 expanded data available through previous studies
MDT Metrobus	2012-2015	Low sample rate; does not cover all routes in the system; only system-wide targets likely
MDT Metromover	2015	Only high-level targets by purpose and access modes
Metrorail	2009	No survey available for modeling purposes after the new line serving MIA station became operational
I-95/I-595 Express	2012	Major service re-organization on routes operated by BCT

Surveys will be expanded to 2015 ridership-levels, time of day period definitions consistent with SERPM 7 will be utilized. The expanded data will be used to develop validation targets, including:

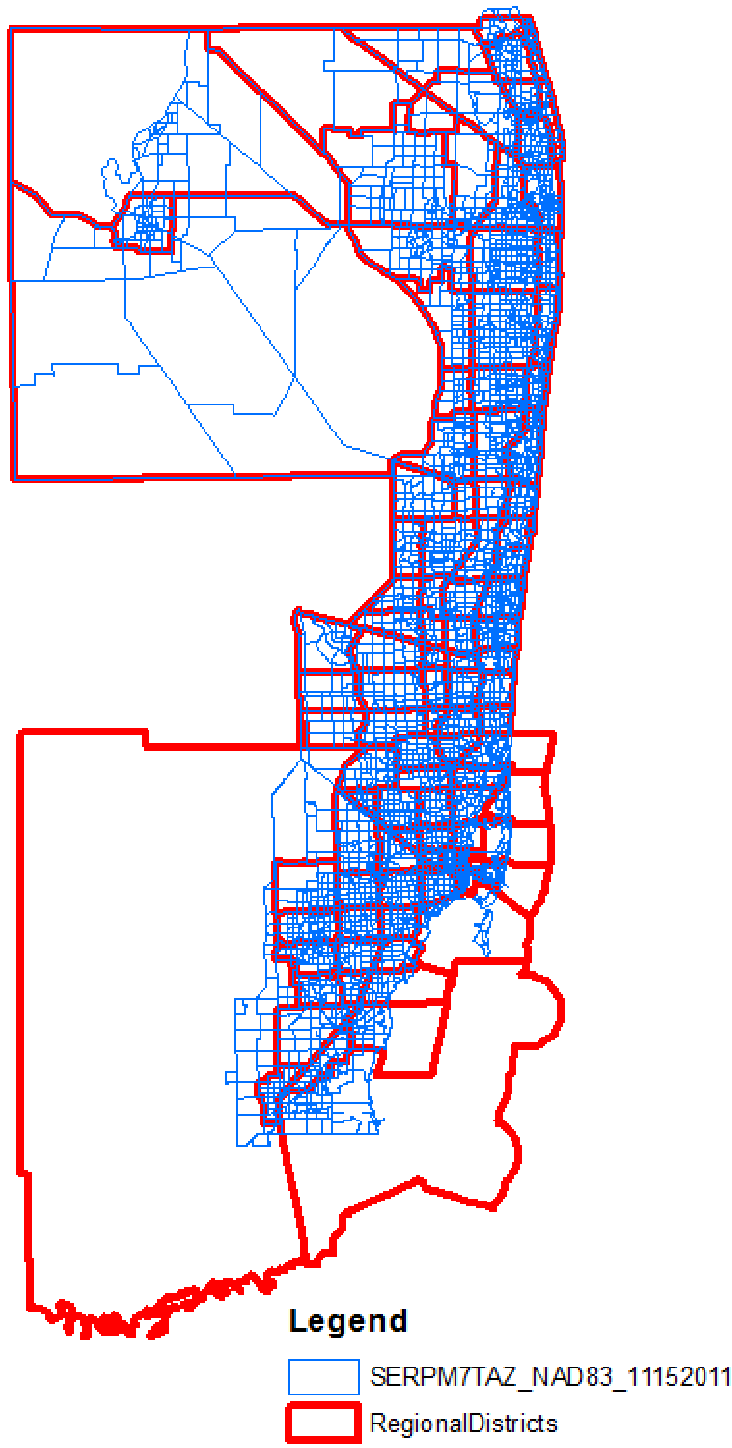
- Boardings by route, by station, by stop-group,
- District-to-district flows on various systems,
- Average trip lengths on Tri-Rail, Metrorail and Metromover,
- Mode-to-mode transfers,
- Development of unexpanded survey trip tables by agency for survey trip table assignment to validate the path-building and mode choice parameters.

Streetlight Data

Streetlight data will be acquired as part of the household survey project. Each MPO developed a zone system with a total of 174 zones for trucks and 100 zones for passengers across the region. There will be two data acquisitions for this region. The first will cover 2015 and the second will refresh for the same zones in 2016. The dataset will include trip-ends for heavy and medium trucks, and passenger vehicles and will come with a Streetlight index that may be used to scale the data. Figure 4.1 shows the Streetlight data person travel districts as they overlay the model TAZs. The Streetlight data districts do not extend to the north or west of the model region, implying that they will not capture travel with an external origin or destination.

These data will be used to update the aggregate travel models including the special generator and external travel models. Although the Streetlight data does not provide the detailed information available in a household survey, and its algorithms for imputing information such as trip purposes are not publicly documented, it does provide a much larger sample than the household survey data. The Streetlight trip data can be used to validate county, area type, and district scale trip patterns. It is particularly useful in providing a second opinion in cases where travel model results are not consistent with expanded survey results.

Figure 4.1 Streetlight Personal Travel Districts



4.1.3 Transportation System Utilization

Transportation system utilization data includes can include traffic counts, transit boarding counts, speed data, among others. This section describes the relevant data available for model validation.

Traffic Counts

A new set of 2015 traffic counts in the region will be assembled for model validation purposes. During this process, the 2015 (or latest) traffic counts using historical data will be reviewed at the particular station and the values of upstream and downstream count stations, anomalies will be reviewed by MPOs and adjusted for model calibration and validation.

Transit Boarding Counts

Various transit on-board survey and ridership survey provide transit ridership data from 2010 to 2015. These data will be used to evaluate transit assignment model outputs.

Travel Time / Speed Data

The available travel time and speed data will be compiled and evaluated for model travel time and speed calibration and validation. Roadway traffic performance data are available by link on the National Highway System from the FHWA's NPMRDS. These data can be linked in a geospatial environment through the ***Traffic Message Channel*** (TMC) location code. The TMC is an international standard that was developed to deliver traffic and travel information to drivers, traffic reporters, and traffic management centers. Typically, TMC roadway segments begin and end at major interchanges or at intersections with other major roads. TMC location codes are generally assigned only to higher functional class roads, where traffic congestion or incidents may impact traveler decision-making.

Although TMC location codes are an industry standard, they are not readily accessible by the general public. In the U.S., these location codes have been integrated into proprietary commercial road database products developed by firms such as INRIX, HERE (formerly NAVTEQ), and TomTom. The TMC location codes are linked directly to road segments in each proprietary database. Consequently, the precise coordinates of the geographic begin and end points for each TMC location code, as well as the geographic alignment and overall length of the roadway section represented by the TMC can vary depending on which database is used.

In creating the NPMRDS, FHWA contracted with HERE to use a subset of their proprietary NAVSTREETS® geospatial road network database to serve as the base map for displaying travel time information. The NPMRDS base map contains those road segments in NAVSTREETS® that are part of the National Highway System (NHS). This road network database is distributed as a geospatial shape file that can be downloaded by authorized state and metropolitan agencies from the NPMRDS web site. However, the NPMRDS road shapefile does not include the TMC location codes as explicit attributes of the road segments; instead, a separate crosswalk table is provided that matches each road segment identifier to one or more TMC location codes. This crosswalk table must first be joined to the road shape file in order to associate each road segment to a TMC location code.

Once the NPMRDS data have been conflated to the SERPM roadway network, methods have been developed to estimate mean link travel times for each link in the network. Specifically, the travel times from the NPMRDS are reported for five-minute intervals (epochs) and will need to aggregated to estimate congested time of day travel times to be used in SERPM 8. The aggregation must use a weighting process to properly account for variation in traffic volumes within the time periods and also over the days of data selected to estimate the congested times (e.g. data from Tuesday through Thursday for selected weeks in April-May and September-October).

Note that the NPMRDS data may be used for the estimation of congested network travel time skims used for model estimation and the single-pass, stepwise calibration process. In addition, free-flow speeds for the SERPM network can be developed from the NPMRDS and can be used to validate (or update) free-flow speeds used for the full-feedback model validation. Finally, the estimated NPMRDS congested speeds will be used in the validation of the final congested speeds produced by SERPM 8.

4.2 Model Parameters

A key piece of information that will be used in model validation is model parameters estimated in other regions. As discussed in more detail in Section 4.0, comparison of model parameters to other regions provides a relatively simple reasonableness check. The most likely comparison source for model parameters are the models recently developed for the San Diego Region and the Sacramento Region, which have similar structures to the SERPM model. The Denver and Atlanta models also have somewhat similar structures and may also be used for parameter comparisons. Additionally, model parameters from selected trip-based models may be used for certain model components. The most likely component to use to compare to trip-based methods is mode choice.

4.3 Backcast Year Survey and Model Inputs and Counts

The backcast year has not been decided yet but it could year 2010. It depends on the extent of data available to be fed into the ABM process. Several data items are necessary for the purpose of backcasting, including the following:

Survey Data

- Regional household activity-travel survey
- Transit on-board transit survey
- Other surveys

Count Data

- Highway traffic count data
- Highway speed data (note that the NPMRDS data are not available prior to July of 2013 so more traditional sources of speed data such as estimated speeds from floating car speed studies will need to be used)
- Transit ridership data.

Supply Data

- socioeconomic data, networks, inputs, scripts, outputs
- Population synthesizer data.

Other Data

- Census data or ACS data closer to the backcast year.

5.0 Planned Validation Tests

This section of the model validation plan discusses specific validation tests that are planned. These include tests of the socioeconomic and network input data used, tests for ABM component model, system-level validation tests for highways and transit, and temporal validation and sensitivity tests.

Because the validation effort represents a major undertaking, each proposed test was assigned a priority describing its importance. Priorities are based on the following considerations:

- **Level 1 priority tests** include those tests that can be produced with relatively little effort and provide good measures of the validity of the models. This priority level includes aggregate validation measures like VMT and VHT summaries, tours per person and per household, and average tour durations and distances.
- **Level 2 priority tests** include more detailed and non-standard validation tests. These tests can be considered to be more directly focused on the tour-based models being developed for the ABM. Since the tests are more detailed and/or non-traditional, they may be more difficult to produce or interpret, or it may be more difficult to acquire data or information from other regions for comparison. In addition, the Level 2 tests typically produce "information overload" and, for that reason, will be used for diagnosis of the model results if a Level 1 priority test suggests an issue with the model calibration / validation.
- **Level 3 priority tests** include those tests that are likely to be costly or difficult to produce or interpret, including those tests that do not have readily available validation data. These tests are listed primarily to suggest other diagnosis measures that might be considered if model calibration / validation results suggest a problem with a model component.

The priorities were assigned to each validation test with the goal of producing the most return for the investment in the validation testing.

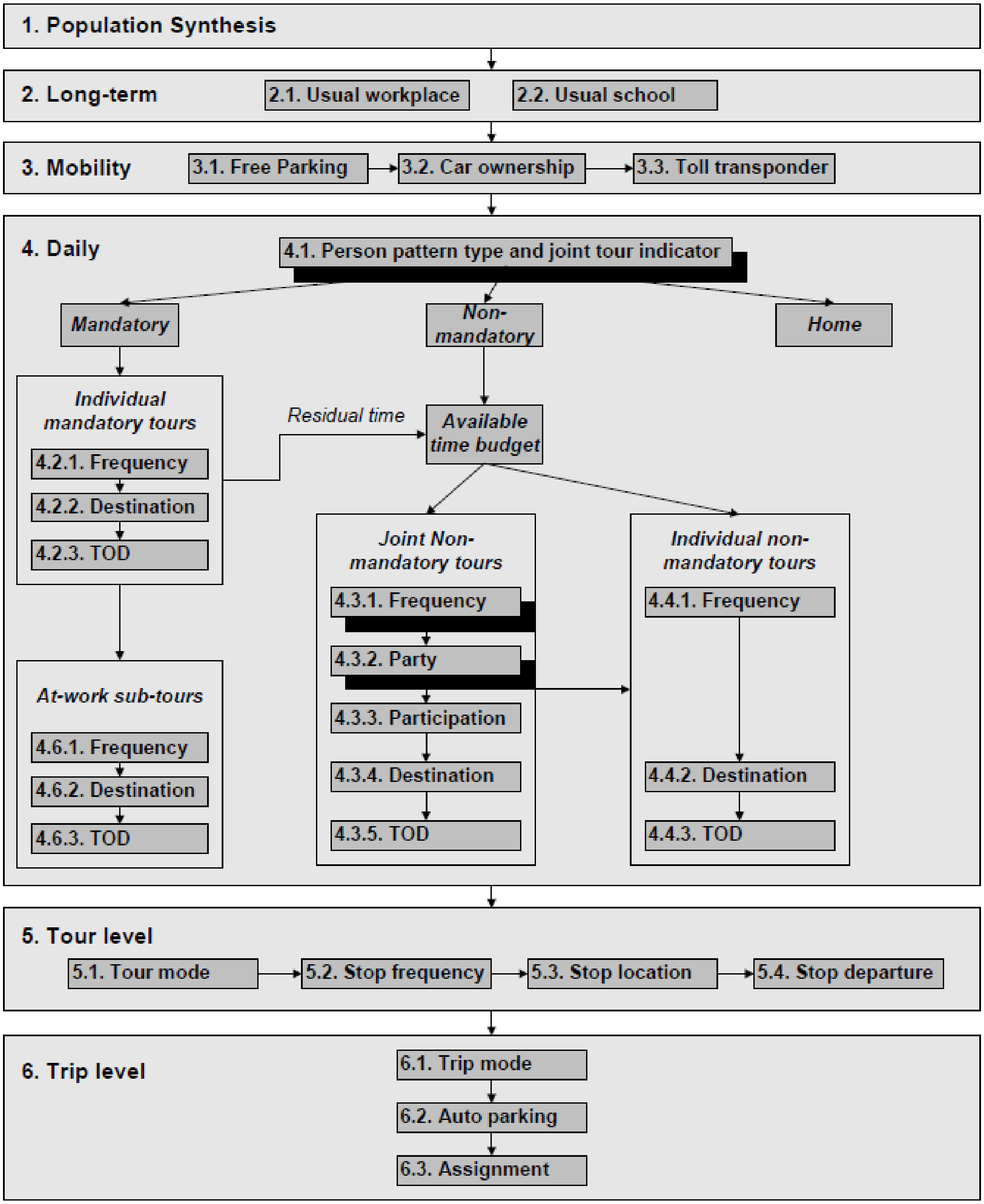
As suggested by the priority levels, the tests may also be classified into basic tests (the Level 1 priority tests), and debugging tests (the Level 2 and, possibly, Level 3 priority tests). The "basic" tests will be conducted for all ABM components and it will provide enough insights into the performance of each component. The "debugging" tests are designed to provide more information at disaggregate levels and are intended to diagnose any errors or examine any anomalies found in the "basic" test results.

Section 5.1 provides background on the structure and flow of the new ABM. The validation tests for inputs to the ABM are discussed in Section 5.2. Section 5.3 discusses validation tests specifically for the population synthesis procedure and each of the estimated model components. System level validation tests discussed in Section 5.4 will rely on the final outputs of the model. These include highway and transit trip tables, highway volumes, and transit boardings by time of day.

5.1 Model Design Background

Figure 5.1 details the structure and flow of the SERPM 7.0 ABM. The SERPM 8.0 ABM will have a similar structure as the existing with changes only to the model component parameters and alternatives. The model design plan was prepared and delivered in a separate document (Cambridge Systematics, Inc., 2017).

Figure 5.1 Model Process Flow



5.2 Input Data Validation

Two key types of data are required for any travel model process: socioeconomic data and transportation network data. These data represent the basic building blocks used with travel models to forecast regional travel. Not only do these inputs affect forecasts, but they are also used in model estimation, calibration, and validation, which can have important consequences for each step. Thus, ensuring the reasonableness of socioeconomic and network data is critical to the overall modeling process.

5.2.1 Socioeconomic Data

The zonal-level data available as inputs into the model include:

- Number of households
- Population
- Median Income
- Employment by Type (Retail, Office, Industrial, Government, Medical)
- K-12 Educational employment and enrollment
- Area Type

The CS team will have little opportunity to assess the correctness of the input socioeconomic information and instead will rely on the accuracy of inputs provided by FDOT District 4. Since the household-based data will form control totals for PopSynII, and the employment-based and educational enrollment data will be instrumental in several model components, the data should be checked if FDOT District 4 has not previously validated the data.

Table 5.1 presents a list of validation checks by level of aggregation that should be considered. Validation checks at the zone level should be based on the compatibility of the zone system with 2010 census block definitions. No specific criteria have been specified for the validation tests; rather, the reasonableness of the data should be gauged by potential impacts on model results.

Table 5.1 Socioeconomic Data Validation Tests

AGGREGATION LEVEL	DATA ITEM	VALIDATION TEST	PRIORITY

Area Type by District (~79 segments)	• Number of households
	• Population in households
	• Median household income
	• Employment by Type
	  • Retail,
	  • Office,
	  • Industrial,
	  • Government,
	  • Medical
	  • K-12 Educational
		• Compare to 2015 ACS and BEBR for households and population – data should match at district level
		• Estimate median household income from zone-level data and compare to available ACS 5-year estimates – results should be "close"
		• Compare employment by type to InfoGroup or LEHD data – results should be "close"
			Level 1 (Basic Test)

Zone	• Color coded GIS plots of percent differences between Zone data, and 2015 ACS data, and LEHD for:
	  • Number of households
	  • Population in households
	  • Median household income
	  • Employment by type
	• Color coded plots of Zone data for:
	  • K-12 educational enrollment
	  • Area type
		• Review trends and look for outliers in term of large percentage differences	Level 2 (Debugging Test)

5.2.2 Transportation Network Data

Transportation network data are the other key input to travel demand forecasting models. The key interfaces between the networks and the ABM consist of:

- Level-of-Service (Skim) Matrices prepared by network path-builders and used as an input to the travel demand forecasting models; and
- Trip Table Matrices by mode created by the ABM and passed back to the path-builder to determine vehicular traffic by highway facility and transit ridership by station and/or route.

To be useful as both a front-end and a back-end to the ABM procedures, the transportation network data and the path-building procedures that use these data must be verified. Verification will include the following steps.

- **Highway network verification**. Highway networks will be spot-checked to confirm accuracy.
  - On a more systematic level, estimated highway travel times for a series of highway journeys will be compared to observed travel times to confirm approximate agreement. This test will help to confirm the appropriateness of the entire network processing procedures including assignment of free flow speeds and capacities, and the how volume-delay functions relate traffic to reduced (congested) travel speeds.
  - Toll and HOV facilities and skims will also be reviewed for accuracy (e.g. toll charges) and reasonableness. As with the validation of general purpose highway skims noted above, estimated travel times for a series of journeys using toll or HOV facilities will be compared to observed travel times to confirm approximate agreement.

For model estimation, reasonable speeds will need to be used to build highway skims. It was decided to build highway skims for the model estimation at the outset and develop the speed processing procedures in parallel with the model estimation process, with checks performed of the highway skims.

Validation of the highway skims will rely on observed travel time data. As noted in Section 4, use of NPMRDS data to develop congested highway travel times. While these data should provide reasonable link specific travel times, we will also look to the reported travel time data from the household survey to verify the network skims. Of course, these data have the well documented reporting issues of rounding starting and ending times to the nearest five minutes on a clock face. In addition, depending on how activities were defined, reported departure and arrival times may include “terminal times”. Nevertheless, understanding these issues will help us interpret the results.

The NPMRDS data will also be used to estimate free-flow speeds on links. Since posted speeds are used for the starting point of the SERPM assignment process, scatterplots of the NPMRDS free-flow speeds versus the posted speeds will be produced. The scatterplots will be stratified by functional class and area type (or county). If NPMRDS free-flow speeds are systematically above or below posted speeds for different functional class-area type strata, an adjustment to the estimation of free-flow speeds for the assignment process might be warranted.

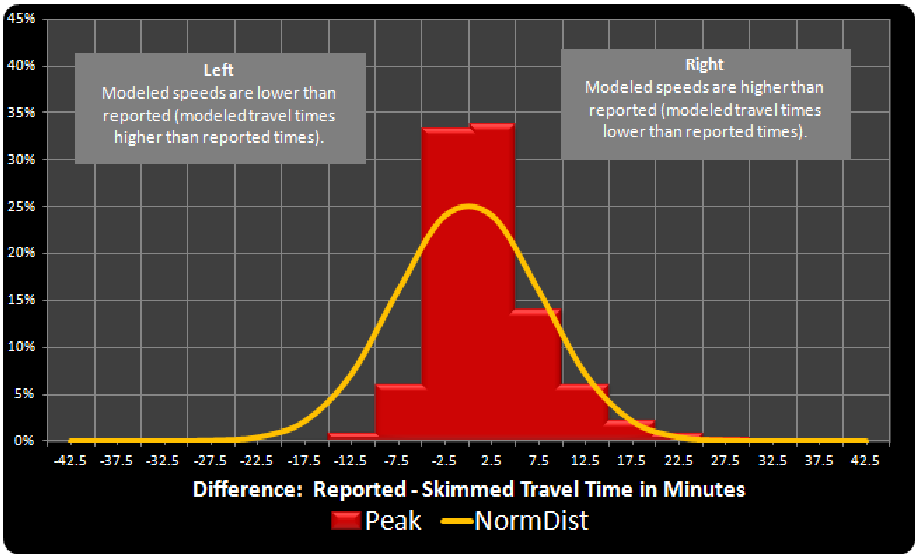
The suggested validation process will involve posting the modeled interchange specific travel times on the household survey trip data for auto drivers and producing scatter plots and TLFDs of the modeled versus reported travel times. The following household survey trip data are required for each auto driver trip:

- Origin zone
- Destination zone
- Time of day (AM peak, midday, PM peak, night)
- Expansion factor
- Reported travel time

The skim data, which will be posted on the survey records, will be based on the estimated congested speeds from the NPMRDS data. In addition to the modeled travel time for the appropriate time of day, the skimmed distance will also be posted. This will allow us to calculate “reported” interchange travel speeds to filter out outliers (e.g. implied speeds greater than, say, 80 miles per hour). The differences between reported and modeled travel times will be calculated for each observed auto driver trip. Since the survey data probably include terminal times, modeled origin and destination terminal times will be added to the skimmed origin to destination travel times prior to calculating the differences[5]. TLFDs for the modeled and reported travel time differences will be produced for each time of day as shown in Figure 5.2. An expectation of the results is that the distribution will be somewhat normally distributed around a mean of zero. The median difference in travel times should also be reported.

[5] If the terminal times are based on origin and destination area types, this step can be performed in the comparison procedures. This way, it may be possible to adjust the modeled terminal times to achieve a better fit between the modeled and reported travel times.

Figure 5.2 Example TLFDs of Modeled and Reported (Survey)



- **Transit network verification.** Transit networks will be examined using the following tests:
  - Review of rules for developing walk-access, park-ride, and kiss-ride access links.
  - Review of rules for estimating mixed-flow (bus) running times as a function of congested highway times.
  - Review of resulting estimates of bus running time on a time point-to-time point basis by time of day.
  - Review coded fixed guideway running time estimates by time of day against timetables.
  - Review rules and coded headways by time-of-day for routes against timetable values.
  - Aggregate path-building test through an assignment of a transit trip table estimated from the on-board survey to network to confirm that line loads and station on-offs by mode of access and time of day reasonably match observed values.
  - Disaggregate path-building test of selected individual paths from the on-board survey to confirm that reported and path-builder paths are similar. If possible, this will be automated so that prediction success tables of modeled versus reported boardings can be compared for the numerous selected paths. The experience in Denver was that about 60 percent of the modeled paths had the same number of boardings per linked trip as the reported paths.
  - Examine resulting estimated travel times (by time of day) contained in the level-of-service matrices to confirm that travel times estimates are properly scaled (e.g., minutes) and are located in the proper file and table.

As with the socioeconomic data tests, there will not be any specific criteria by which these network tests will be measured. Reasonableness of the data will be gauged by potential impacts on model results.

5.3 Component Validation Tests

This section details the specific tests planned for validation of the individual components of the ABM. The model components and sections in which they are discussed are listed in Table 5.2. Validation tests have generally been grouped as follows:

- Comparison of model parameters to other regions
- Disaggregate validation tests
- Aggregate validation tests

Table 5.2 Components of the SERPM 8.0 System

Model Name	Decision-Making Unit	What is Predicted	Validation Test
Population Synthesis	N/A	Household size and composition, household income, number of workers in the household, presence of children, person age, occupation, gender, employment status, student status	Table 5.3 Synthetic Population Generator Validation Tests
Residential Location Choice	Household	Residence location zone (MAZ)	Table 5.3 Synthetic Population Generator Validation Tests
Work from home	Worker	Workplace location within or outside of the home	Table 5.4 Usual Workplace Location Validation Tests
Out of home workplace location choice	Worker	Workplace location zone (MAZ)	Table 5.4 Usual Workplace Location Validation Tests
Preschool location choice	Persons age 0-4	Preschool location zone (MAZ)	Table 5.5 Usual School Location Validation Tests
Grade school location choice	Persons age 5-13	School location zone (MAZ)	Table 5.5 Usual School Location Validation Tests
High school location choice	Persons age 14-17	School location zone (MAZ)	Table 5.5 Usual School Location Validation Tests
University location choice	College Student	School location zone (MAZ)	Table 5.5 Usual School Location Validation Tests
Employer Parking Provision and Reimbursement Model	Workers whose workplace is in parking-priced area	Weather workers has free on-site parking, parking reimbursement, no free/subsidized parking	Table 5.6 Employer Parking Provision and Reimbursement Validation Tests
Auto Availability	Household	Number of vehicles available	Table 5.7 Auto Availability Validation Tests
Auto technology	Household	Connected / Automated technology level	N/A
Willingness to Ridesource	Household	Whether household members will use ridesourcing	N/A

Coordinated Daily Activity Pattern (DAP)	Household	Personal DAPs and the generation of individual tours by purpose.	Table 5.8 Daily Activity Pattern Model Validation Tests
Individual Mandatory Tour Frequency	Person	Number and purpose of mandatory tours for each person.	Table 5.9 Individual Mandatory Tours – Tour Frequency Model Validation Tests
Individual Mandatory Tour Time of Day Choice	Tour	Tour departure and arrival half-hour periods	Table 5.10 Individual Mandatory Tour Time of Day Choice Model Validation Tests
Joint Tour Frequency	Household	Joint tour frequency (0,1,2) by purpose of the joint tours	Table 5.11 Joint Non-Mandatory Tours – Tour Frequency Model Validation Tests
Joint Tour Composition	Joint Tour	Person type (adults only, children only, adults & children) in the tour	Table 5.12 Joint Non-Mandatory Tours – Tour Party Composition Model Validation Tests
Joint Tour Participation	Persons	Whether each person corresponds to each joint tour.	Table 5.13 Joint Non-Mandatory Tours – Tour Participation Model Validation Tests
Joint Tour Primary Destination Choice	Joint Tour	Tour primary destination zone (MAZ)	Table 5.16 Non-Mandatory Tour Destination Choice Validation Tests
Joint Tour Time of Day Choice	Joint Tour	Tour departure and arrival half-hour periods	Table 5.17 Joint Non-Mandatory Tour Time of Day Choice Model Validation Tests
Individual Non-Mandatory Tour Frequency	Person	0-3 tours of each type of Non-mandatory activity	Table 5.15 Individual Non-Mandatory Tours – Tour Frequency Model Validation Tests

Individual Non-Mandatory Tour Primary Destination Choice	Tour	Tour primary destination zone (MAZ)	Table 5.16 Non-Mandatory Tour Destination Choice Validation Tests
Individual Non-Mandatory Tour Time of Day Choice	Tour	Tour departure and arrival half-hour periods	Table 5.17 Joint Non-Mandatory Tour Time of Day Choice Model Validation Tests
At-Work Sub-Tour Frequency	Work Tour	Number and purpose of tours by 9 alternatives of work-based sub-tours	Table 5.18 At Work Sub-tours – Tour Frequency Model Validation Tests
At-Work Sub-Tour Primary Destination Choice	Tour	Tour primary destination zone (MAZ)	Table 5.19 Non-Mandatory Tour Destination Choice Validation Tests
At-Work Sub-Tour Time of Day Choice	Tour	Tour departure and arrival half-hour periods	Table 5.20 Work Sub-tour Time of Day Choice Model Validation Tests
Tour Mode Choice Model	Tour	Main tour mode	Table 5.21 Tour Mode Choice Model Validation Tests
Intermediate Stop Frequency Model	Person	Number of intermediate stops (0-3) on the way to/from the primary destination	Table 5.22 Intermediate Stop Generation Model Validation Tests
Intermediate Stop Purpose	Stop	Stop purpose	Table 5.22 Intermediate Stop Generation Model Validation Tests
Intermediate Stop Location Choice Model	Tour	Intermediate stop location (MAZ)	Table 5.23 Intermediate Stop Location Model Validation Tests

Intermediate Stop Departure Time Model	Tour	Stop departure time period	Table 5.24 Stop Time of Day Choice Model Validation Tests
Trip Mode Choice Model	Trip	Mode for each trip along the tour	Table 5.25 Trip Mode Choice Validation Tests
Parking Location Choice	Trip	Parking-priced zone (MAZ) of the terminal end of tour	Table 5.26 Parking Choice Validation Tests

The following sub-sections list the each type of model components and detail the types of validation tests that will be performed for each component. Some of the sub-sections represent multiple model components, as shown in Table 5.2, because validation tests will be largely the same (e.g., work, school, and other tour mode choice models will follow the same general procedures for validation, as will each of the tour time of day choice models).

Disaggregate validation tests will generally be performed by applying the estimated model parameters to the estimation data and comparing modeled results to reported behavior. The disaggregate validation procedure recommended for the ABM development is to use the entire estimation dataset for both model estimation and model validation. Stratification of disaggregate validation results by different socioeconomic, geographic, mode, or other travel data can also be performed. In many of the model components, the stratifications used for the validation results can mirror the stratifications used for aggregate validation.

Input data for aggregate validation tests will generally be the results from previous steps in the ABM process. The exception to this rule for the aggregate validation of the individual model components will be that the travel impedance matrices used for the modeling process will be the same as those used for the model estimation. This approach will obviate the need to run multiple iterations of the entire ABM for the validation of the individual model components. However, the final process validation as represented by the traffic and transit assignments will be based on running the ABM from “top to bottom” including feedback iterations.

Note that the disaggregate validations will focus specifically on the performance of the individual model components. The aggregate validations of the individual model components will help in the assessment of the impact of cumulative error in the model validation.

5.3.1 Synthetic Population Generator

The SERPM 8.0 model will still use the PopSynII synthetic population generator. Table 5.3 summarizes the reasonableness tests for household data that will be produced for base year 2015. The primary reasonableness tests will be socioeconomic distributions stratified by various geographic strata. Table 5.3 also shows the priority level for the tests.

Table 5.3 Synthetic Population Generator Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOME	PRIORITY
Regional	• Two dimensional cross-tabulations: • Households by household income and household size • Households by household income and number of employed household members (full- or part-time) • Persons by age and sex • Persons by age and employment • Persons by age by school grade	• Confirm that control variables from 2015 ACS data have been maintained (within ±3 percent) • Review trends of non-control variables for reasonableness over time	Level 1 (Basic Test)
County	• Same as for regional	• Review trends over time for reasonableness	Level 1 (Basic Test)

5.3.2 Usual Workplace Location

Usual workplace location models have been estimated for other regions. Thus, it will be possible to compare estimated model parameters to those for other regions. In addition to the household survey data, comparisons to CTPP and LEHD data can be performed. Table 5.4 lists the validation tests that will be performed.

Table 5.4 Usual Workplace Location Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected TAZ against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)

Aggregate	• Modeled versus observed (from household survey) home-to-regular workplace impedance histograms with impedances based on:
	  • Congested auto travel time
	  • Straight-line travel distance
	• Modeled versus observed (from household survey) average impedances (same as above) stratified by[1]:
	  • Household income level
	  • Number of workers in the household
	  • Area type of household
	  • Full- or part-time worker (if those population synthesis data are available)
	  • Sex of worker
		• Similar to home-based work trip length frequency distributions
		• Modeled to observed averages should be ±5 percent
		• Modeled to observed impedance histogram coincidence ratios (reasonableness test only)
			Level 1 (Basic Test)

	• Modeled versus observed home-to-workplace flows for:
	  • District-to-district
	• Modeled versus observed work-at-home for:
	  • District
		• Look for anomalies
		• Comparisons may be made to observed (expanded) household survey data, 5-year ACS data, and 2010 LEHD data
			Level 1 (Basic Test)

	• Modeled versus observed work-at-home by[1]:
	  • Household income level
	  • Number of workers in the household
	  • Full- or part-time worker
	  • Age of worker (under 18 and 18 or older)
	  • Sex of worker
	  • Industry of worker
		• Comparisons may be made to observed (expanded) household survey data and/or 5-year ACS data	Level 2 (Debugging Test)

[1] The stratification variables listed are likely to be explanatory variables in the model. Summaries by other socioeconomic variables not used as explanatory variable (e.g. ethnicity) can be performed, but should be considered only as Level 3 priority validation measures.

5.3.3 Usual School Location

Like usual workplace location models, usual school location, including grade school location, high school location and university location models have been estimated for other regions. Thus, it will be possible to compare estimated model parameters to those for other regions. Table 5.5 lists the validation tests that will be performed.

Table 5.5 Usual School Location Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected TAZ against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate	• Modeled versus observed (from household survey) home-to-regular school impedance histograms with impedances based on: • Congested auto travel time • Straight-line travel distance	• Similar to home-based school trip length frequency distributions • Modeled to observed averages should be ±5 percent • Modeled to observed impedance histogram coincidence ratios (reasonableness test only)	Level 21 (Basic Test)
	• Modeled versus observed (from household survey) average impedances (same as above) stratified by[1]: • Household income level • Grade level of student	• Modeled to observed averages should be ±5 percent	Level 1 (Basic Test)
	• Modeled versus observed school-at-home by[1]: • Geographic location (district) • Household income level • Grade level of student	• Comparisons may be made to observed (expanded) household survey data (if survey contains sufficient observations to include this as a regular school location choice)	Level 1 (Basic Test)
Aggregate	• Modeled versus observed (from household survey) home-to-school flows for: • District-to-district for elementary and high school grade levels • If applicable, school district of residence-to-school district of school for elementary and high school grade levels • Major colleges and universities[2]	• Look for anomalies • If address locations or zip codes of students can be obtained from school districts and/or colleges and universities, the data could be used for validation	Level 2 (Debugging Test)

1 The stratification variables listed are likely to be explanatory variables in the model. Summaries by other socioeconomic variables not used as explanatory variable (e.g. ethnicity) can be performed, but should be considered only as Level 3 priority validation measures.

2 Major community colleges and other university campuses will be specified based on consultation with FDOT District 4 staff.

5.3.4 Employer Parking Provision and Reimbursement (Free Parking Eligibility)

The Employer Parking Provision and Reimbursement model forecasts whether a worker whose workplace is in parking-priced areas receives free on-site parking, parking reimbursement, no free/subsidized parking. These options are, of course, applicable to workers with a regular workplace outside the home. Thus, the results of this model can be summarized both at the home location to get information by, say, socioeconomic group, and the regular workplace location. Table 5.6 lists the validation tests that will be performed.

Table 5.6 Employer Parking Provision and Reimbursement Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Aggregate	• Number of modeled workers with regular workplaces in parking-priced areas with: • Free on-site parking • Parking reimbursement • No free/subsidized parking • Proportions of workers with free, subsidized, and no free/subsidized parking by: • Income group • Worker status (full-time, part-time)	• Modeled to observed totals should be ±5 percent	Level 1 (Basic Test)
District of Residence	• Proportions of workers with free, subsidized, and no free/subsidized parking by: • Income group • Worker status (full-time, part-time)	• Relatively accurate match	Level 2 (Debugging Test)
District of Parking-Priced Area	• Proportions of workers with free, subsidized, and no free/subsidized parking by: • Area type • Average posted parking cost range (ranges to be determined)	• Look for anomalies	Level 2 (Debugging Test)

5.3.5 Auto Availability

The auto availability model predicts the number of motorized vehicles owned, leased, or otherwise belonging to fleet of vehicles possessed by the household. It will consider household characteristics as forecast by the population synthesizer along with regular workplace and school accessibility information.

Extensions to the validation tests performed for the population synthesizer will be the primary validation tests. However, validation of the vehicle availability model should also consider the impacts of accessibility on vehicle availability. A number of trip based models have attempted to incorporate transit accessibility into vehicle availability models under the hypothesis that increased transit accessibility reduces the need for additional autos. The same type of argument can be made for accessibility via non-motorized modes. Thus, two special tests comparing vehicle availability per person of driving age (age 16 or older) are planned to test the veracity and reasonableness of the impact of transit or non-motorized accessibility on vehicle availability. Table 5.7 lists the proposed validation tests.

Table 5.7 Auto Availability Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of modeled vehicle availability against observed estimation data (percent correct). • Prediction success cross-tabulation of predicted autos owned by reported autos owned.	• Percent prediction success likely to be very low. • Percentages of households with predicted autos within ±1 auto of reported autos owned should be relatively high.	Level 3 (Debugging Test)

Aggregate	• Regional and county-specific three dimensional cross-tabulations of households by vehicle availability (0, 1, 2, 3+ autos) by:
	  • Household income and household size
	  • Household income and number of employed household members
	  • Household income and number of members age 16 or older
		• Regional and county-specific cross-tabulations should be compared to
		  • Expanded household survey data
		  • 5-year ACS results
			Level 1 (Basic Test)

Aggregate	• Mean and standard deviations of number of autos per person of driving age (16 or older) stratified by:
	  • Transit accessibility[1] and household income
	  • Walk accessibility[2] and household income.
	  • Residential density and household income
		• Modeled to observed means of autos per driving age person for four transit accessibility levels, three walk accessibility levels, or residential density (ranges to be determined) controlling for household income
		• Compare differences between modeled means for each transit accessibility level, walk accessibility level or residential density for reasonableness
			Level 2 (Debugging Test)

[1] Transit accessibility will be defined as no access, high access, drive-only access, and other where:

- No access implies that none of the selected regular workplaces and school locations for household members is accessible by transit.
- High access implies that all of the selected regular workplaces and school locations for household members are accessible by walk access to transit
- Drive only access implies that all of the selected regular workplaces and school locations for household members are accessible by transit but only if drive access to transit is used.
- Other implies that only some of the selected regular workplaces and school locations for household members are accessible by transit, or that some locations are accessible by walk access to transit and others are accessible by only drive access to transit.

[2] Walk accessibility will be defined as no access, high access, and other where:

- No access implies that none of the selected regular workplaces and school locations for household members is within one mile of walking distance.
- High access implies that all of the selected regular workplaces and school locations for household members are accessible within one mile of walking distance.
- Other implies that only some of the selected regular workplaces and school locations for household members are accessible within one mile of walking distance.

5.3.6 Coordinated Daily Activity Pattern Model

The coordinated daily activity pattern model forecasts the type of daily activity pattern for up to 5 members of the household simultaneously. The alternatives are mandatory travel, non-mandatory travel, or stay at home. Tables of the number of people (or the percentage of people) having each of the alternatives can be created. Such summaries can be stratified by meaningful socioeconomic, demographic, and geographic strata. Recommended validation measures are shown in Table 5.8.

Table 5.8 Daily Activity Pattern Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions (e.g., Houston, Twin Cities)	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of modeled daily activity pattern choices against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate (for Region)	• Numbers or percentages of residents making mandatory, non-mandatory, or no travel by type[1]: • By household size and income group • By household size and vehicle availability • By gender and age group • By employment status • By student status	• Compare modeled to observed numbers or percentages from the expanded household survey • Review for reasonable patterns • Compare percentages of those making no travel ("immobiles") to information summarized by Madre et al. (2007)	Level 1 (Basic Test)
Aggregate (by District)	• Same as for region	• Same as for region	Level 2 (Debugging Test)

[1] The stratification variables listed are likely to be explanatory variables in the model. Summaries by other socioeconomic variables not used as explanatory variable (e.g. ethnicity) could be performed, but should be considered only as Level 3 priority validation measures.

5.3.7 Individual Mandatory Tours

Tour Frequency

Based on the DAP chosen for each person, individual mandatory tours, such as work, school and university tours are generated at person level. The model is designed to predict the exact number and purpose of mandatory tours for each person who chose the mandatory DAP type at the previous decision-making stage: 1 work tour, 2 work tours, 1 school tour, 2 school tours, or 1 work tour with 1 school. Since the DAP type model at the household level determines which household members engage in mandatory tours, all persons subjected to the individual mandatory tour model implement at least one mandatory tour. Validation tests for the joint travel model are listed in Table 5.10.

Table 5.9 Individual Mandatory Tours – Tour Frequency Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of modeled individual mandatory tour choices against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)

Aggregate (for Region)	• Numbers of persons with mandatory tours making 1 work tour, 2 work tours, 1 school tour, 2 school tours, or 1 work tour with 1 school tour by:
	  • Person type by sex
	  • Household auto availability
		• Compare modeled to expanded observed numbers or percentages from the household survey
		• Review for reasonable patterns
			Level 1 (Basic Test)

Aggregate (by District)	• Same as for region	• Same as for region	Level 2 (Debugging Test)

Tour Time of Day

The tour time of day choice models simultaneously forecast the arrival and departure time periods for the tour primary activity locations. For most of the day, the time interval size for time period alternatives is 30 minutes, resulting in 861 different alternatives. Since it is not reasonable to review 861 alternatives, the aggregate validation will combine time periods and use other procedures to simplify the analyses.

Tour time of day choice provides two important items that can be validated: time of day of travel and activity duration. Time of day of travel can be determined from both the arrival and departure times at the primary activity location. Activity duration can be determined by the time difference between the arrival and departure times. Table 5.11 summarizes the validation tests anticipated for the tour time of day choice model validation.

Table 5.10 Individual Mandatory Tour Time of Day Choice Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected TAZ against observed estimation data	• Prediction success likely to be low.	Level 3 (Debugging Test)
Aggregate (for Region)	• Histograms of departure times to the primary tour destination by purpose • Repeat above for arrival times from the primary tour destination • Repeat above for tour duration	• Compare modeled to expanded observed numbers of tours • Review for reasonable patterns	Level 1 (Basic Test)
	• Primary tour durations by type[1]: • Gender and person type	• Compare modeled to expanded observed numbers of tours • Review for reasonable patterns	Level 2 (Debugging Test)

5.3.8 Joint Non-Mandatory Tours

Tour Frequency

The joint travel frequency model will forecast the number of joint non-mandatory tours made by individuals from the same household. There are 5 non-mandatory tour purposes and the options of making no non-mandatory joint tours, one joint tour, or two joint tours during the day resulting in the following 21 options:

- No joint non-mandatory tours
- 1 shop tour
- 1 maintenance tour
- 1 eating out tour
- 1 visiting tour
- 1 discretionary tour
- 2 shop tours
- 2 maintenance tours
- 2 eating out tours
- 2 visiting tours
- 2 discretionary tours
- 1 shop tour + 1 maintenance tour
- 1 shop tour + 1 eating out tour
- 1 shop tour + 1 visiting tour
- 1 shop tour + 1 discretionary tour
- 1 maintenance tour + 1 eating out tour
- 1 maintenance tour + 1 visiting tour
- 1 maintenance tour + 1 discretionary tour
- 1 eating out tour + 1 visiting tour
- 1 eating out tour + 1 discretionary tour
- 1 visiting tour + 1 discretionary tour

Fully-joint travel means that all important aspects of tour-making are shared by two or more household members, including origin, destination, mode, time of departure from origin and destination (as well as arrival times) and purpose. Validation tests for the joint travel model are listed in Table 5.12.

Table 5.11 Joint Non-Mandatory Tours – Tour Frequency Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of choices against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate (for Region)	• Numbers or percentages of households making no joint travel tours, 1 joint tour (5 purposes), or 2 joint tours (15 purpose combinations) by: • Household size • Income group • Vehicle availability • Number of workers	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 1 (Basic Test)
	• Average number of joint travel tours per household: • By household size and income group • By household size and vehicle availability • By household size and number of workers	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 2 (Debugging Test)

Tour Party Composition

Joint non-mandatory tour party composition is modeled for each tour, and determines the person types that participate in the tour. The following options are possible:

- only adults,
- only children, and
- adults and children.

Validation tests for the tour party composition model are listed in Table 5.13.

Table 5.12 Joint Non-Mandatory Tours – Tour Party Composition Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of choices against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate (for Region)	• Numbers or percentages of households making 1 joint tour (5 purposes), or 2 joint tours (15 purpose combinations) for each of the 3 tour party composition types	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 1 (Basic Test)

Tour Participation

Joint tour participation is modeled for each person and each joint tour. If the person does not correspond to the composition of the tour determined in the joint tour composition model, they are ineligible to participate in the tour. Similarly, persons whose daily activity pattern type is home are excluded from participating. The model relies on a heuristic process to assure that the appropriate persons participate in the tour as per the composition model. Validation tests for the tour participation model are listed in Table 5.14.

Table 5.13 Joint Non-Mandatory Tours – Tour Participation Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of choices against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate (for Region)	• Average numbers of adults and children participating in joint tours by the 3 tour party composition types for each of the 5 primary joint tour purposes (see Table 5.15 for a summary table template)	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 1 (Basic Test)
	• Same as above stratified by household size	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 2 (Debugging Test)

Table 5.14 Joint Non-Mandatory Tours – Tour Participation Validation Template

Tour Purpose	Average Number of Participants by Tour Composition Type		
	Only Adults	Only Children	Adults & Children
Shop			
Maintenance			
Eat Meal			
Visit			
Discretionary			

Tour Destination

The joint tour primary destination choice model determines the location of the tour primary destination. The destination is chosen for the tour and assigned to all tour participants. Since the primary destination choice is performed at the tour level, SERPM can perform joint and individual non-mandatory tour destination choice in the same model step. Validation tests for both joint and individual non-mandatory tour destination choice are discussed in Section 5.3.11**Error! Reference source not found.**.

Tour Time of Day

After joint tours have been generated and assigned a primary location, the tour departure time from home and arrival time back at home is chosen simultaneously. The model uses the same structure as the mandatory tour time of day model. However, a unique condition applies when applying the time-of-day choice model to joint tours. That is, the tour departure and arrival period combinations are restricted to only those available to all participants on the tour, after scheduling mandatory activities. Once the tour departure/arrival time combination is chosen, it is applied to all participants on the tour.

As with joint non-mandatory tour destination choice, joint non-mandatory tour time of day choice is performed simultaneously with individual non-mandatory tour time of day choice. Validation tests for the joint and individual non-mandatory tour time of day choice are discussed in Section 5.3.11.

5.3.9 Individual Non-Mandatory Tours

Tour Frequency

The individual non-mandatory tour frequency model generates all non-mandatory tours for individuals traveling alone. There are six kinds of non-mandatory activities:

- shop,
- escort,
- other maintenance,
- eat out,
- visit, and
- other discretionary.

Between 0-3 tours may be generated for each purpose. No more than five non-mandatory tours per person are allowed, and certain infrequent combinations are excluded from the choice set. Ultimately, only 160 different choices are modeled. Validation tests for the joint travel model are listed in Table 5.16.

Table 5.15 Individual Non-Mandatory Tours – Tour Frequency Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of choices against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate (for Region)	• Numbers or percentages of individuals by person type, making 0, 1, 2, or 3 tours, for each of the 6 purposes	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 1 (Basic Test)
	• Average number of individual tours by type per household: • By household size and income group • By household size and vehicle availability • By household size and number of workers	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 2 (Debugging Test)

Tour Destination

The individual non-mandatory tour destination choice model determines the location of the tour primary destination. Since the primary destination choice is performed at the tour level, SERPM can perform joint and individual non-mandatory tour destination choice in the same model step. Validation tests for both joint and individual non-mandatory tour destination choice are discussed in Section 5.3.11**Error! Reference source not found.**.

Tour Time of Day

After individual non-mandatory tours have been generated and assigned a primary location, the tour departure time from home and arrival time back at home is chosen simultaneously. The model uses the same structure as the mandatory tour time of day model. However, a unique condition applies when applying the time-of-day choice model to individual tours. That is, the tour departure and arrival period combinations are restricted to only those available for the individual after scheduling mandatory and joint tour activities. Validation tests for the joint and individual non-mandatory tour time of day choice are discussed in Section 5.3.11.

5.3.10 Joint and Individual Non-Mandatory Tour Destination Choice

Validation of both joint and individual non-mandatory tour destination choice will be performed at one time. While information regarding whether the tour was a joint tour (and all individuals on the tour have the same destination information) is available, checks at that level of detail will be for debugging purposes only. Validation tests for the tour participation model are listed in Table 5.17.

Table 5.16 Non-Mandatory Tour Destination Choice Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected ZONE against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate	• Modeled versus observed (from household survey) tour destination location impedance histograms by tour purpose with impedances based on: • Congested auto travel time • Straight-line travel distance	• Modeled to observed averages should be ±5 percent • Modeled to observed impedance histogram coincidence ratios (reasonableness test only)	Level 1 (Basic Test)
	• Modeled versus observed (from household survey) average impedances by tour purpose stratified by[1]: • 4 tour party types (individual, joint–only adults, joint–only children, joint–mixed) • Household income level	• Modeled to observed averages should be ±5 percent	Level 2 (Debugging Test)

[1] The stratification variables listed are likely to be explanatory variables in the model. Summaries by other socioeconomic variables not used as explanatory variable (e.g. ethnicity) can be performed, but should be considered only as Level 3 priority validation measures.

5.3.11 Joint and Individual Non-Mandatory Tour Time of Day Choice

As with validation the joint and individual non-mandatory tour destination choice models, validation of the joint and individual non-mandatory tour time of day choice models will be performed at one time. The tour time of day choice models simultaneously forecast the arrival and departure time periods for the tour primary activity locations. For most of the day, the time interval size for time period alternatives is 30 minutes, resulting in 861 different alternatives. Since it is not reasonable to review 861 alternatives, the aggregate validation will combine time periods and use other procedures to simplify the analyses.

Tour time of day choice provides two important items that can be validated: time of day of travel (departure and arrival times) and activity duration. Time of day of travel can be determined from both the arrival and departure times at the primary activity location. Activity duration can be determined by the time difference between the arrival and departure times. Table 5.18 summarizes the validation tests anticipated for the tour time of day choice model validation.

Table 5.17 Joint Non-Mandatory Tour Time of Day Choice Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected TAZ against observed estimation data	• Prediction success likely to be low.	Level 3 (Debugging Test)
Aggregate (for Region)	• Histograms of departure times to the primary tour destination by purpose • Repeat above for arrival times from the primary tour destination • Repeat above for tour duration	• Compare modeled to expanded observed numbers of tours • Review for reasonable patterns	Level 1 (Basic Test)
	• Primary tour durations by type: • Gender • Person type	• Compare modeled to expanded observed numbers of tours • Review for reasonable patterns	Level 2 (Debugging Test)

5.3.12 At Work Sub-tours

Sub-tour Frequency

Work-based sub-tours are relevant only for those persons with at least one work tour. The at-work tour frequency model predicts the number and purpose of tours that start at work. These underlying activities are mostly individual (e.g., business-related and eating-out purposes), but may also be for household or person maintenance tasks. In addition to the null alternative (no work sub-tour), there are eight actual tour alternatives in the model:

- 1 eat out tour,
- 1 business tour,
- 1 other tour,
- 2 business tours,
- 2 other tours,
- 1 eat out & 1 business tour,
- 1+ eat out and 1+ other, and
- 1+ business and 1+ other

Validation tests for the at work sub-tour frequency model are listed in Table 5.19.

Table 5.18 At Work Sub-tours – Tour Frequency Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of modeled individual mandatory tour choices against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate (for Region)	• Numbers of sub-tours and sub-tour rates (per work tour) by purpose by: • Area type of work location • Worker type (full-time, part-time, student, other) • Sex of person making parent work tour • Household income of person making parent work tour	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 1 (Basic Test)
	• Proportions of workers (making work tour) with 0 sub-tours and with 1+ sub-tours by: • Area type of work location • Worker type (full-time, part-time, student, other) • Sex of person making parent work tour • Household income of person making parent work tour	• Compare modeled to expanded observed numbers or percentages from the household survey • Review for reasonable patterns	Level 1 (Basic Test)

Sub-tour Destination Choice

The work sub-tour primary destination choice model determines the location of the primary sub-tour destination. Validation tests for the at work sub-tour frequency model are listed in Table 5.20.

Table 5.19 Non-Mandatory Tour Destination Choice Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected ZONE against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate (for Region)	• Modeled versus observed (from household survey) sub-tour destination location impedance histograms by sub-tour purpose with impedances based on: • Congested auto travel time • Straight-line travel distance	• Modeled to observed averages should be ±5 percent • Modeled to observed impedance histogram coincidence ratios (reasonableness test only)	Level 1 (Basic Test)
	• Modeled versus observed (from household survey) average impedances by sub-tour purpose stratified by[1]: • Area type of work location • Worker type (full-time, part-time, student, other) • Sex of person making parent work tour • Household income of person making parent work tour	• Modeled to observed averages should be ±5 percent	Level 2 (Debugging Test)

[1] The stratification variables listed are likely to be explanatory variables in the model. Summaries by other socioeconomic variables not used as explanatory variable (e.g. ethnicity) can be performed, but should be considered only as Level 3 priority validation measures.

Sub-tour Time of Day

The sub-tour time of day choice models simultaneously forecasts the arrival and departure time periods for the sub-tour primary activity locations. For most of the day, the time interval size for time period alternatives is 30 minutes, resulting in 861 different alternatives. Since it is not reasonable to review 861 alternatives, the aggregate validation will combine time periods and use other procedures to simplify the analyses.

Sub-tour time of day choice provides two important items that can be validated: time of day of travel and activity duration. Time of day of travel can be determined from both the arrival and departure times at the primary activity location. Activity duration can be determined by the time difference between the arrival and departure times.

After individual sub-tours have been generated and assigned a primary location, the sub-tour departure time from work location and arrival time back at the work location is chosen simultaneously. The model uses the same structure as the mandatory and non-mandatory tour time of day models. However, the tour departure and arrival period combinations are restricted to only those available based on the time window of the parent work tour. Table 5.21 summarizes the validation tests anticipated for the sub-tour time of day choice model validation.

Table 5.20 Work Sub-tour Time of Day Choice Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected TAZ against observed estimation data	• Prediction success likely to be low.	Level 3 (Debugging Test)
Aggregate (for Region)	• Histograms of departure times to the primary sub-tour destination • Repeat above for arrival times from the primary sub-tour destination • Repeat above for tour duration	• Compare modeled to expanded observed numbers of tours • Review for reasonable patterns	Level 1 (Basic Test)

	• Primary tour durations by:
	  • Area type of work location
	  • Gender
	  • Worker type
		• Compare modeled to expanded observed numbers of tours
		• Review for reasonable patterns
			Level 2 (Debugging Test)

5.3.13 Tour Mode Choice Model

The draft SERPM 8 Model Design Plan discusses several options for modifying the SERPM tour mode model. The net effect of whatever option is chosen should be a net reduction of the numbers of submodes considered under each main mode. At the same time, ridesourcing (e.g. using a smart phone app to order a shared ride provider, using a shared vehicle service) will be added as a submode under auto. It is anticipated that the following nine submodes will be present in the SERPM 8 model regardless of final design decisions:

- Auto drive alone
- Auto shared ride 2
- Auto shared ride 3+
- Auto ridesourcing
- Transit walk access
- Transit park and ride access
- Transit kiss and ride access
- Walk
- Bike

Note that the K-12 school tour purpose will include a tenth mode, school bus.

Separate mode choice models will be developed for the following aggregate tour purposes:

- Work tours
- University tours
- K-12 school tours
- Maintenance tours
- Discretionary tours
- Work sub-tours

Table 5.21 summarizes the expected tour mode choice model validation tests. The validation tests will be performed for each of the six tour purposes listed above. Some details may change based on final decisions regarding the SERPM 8 tour mode choice models.

Table 5.21 Tour Mode Choice Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from: • other regions • SERPM 7 • FTA Guidelines	• Primarily comparison for reasonableness. Coefficients outside of FTA guidelines will need explanation (for use in New Starts model justification).	Level 1 (Basic Test)
	• Conformance with FTA guidance related to non-logit decision rules, bizarre constants, alternative-specific constants, and path-builder/mode choice inconsistencies	• Results outside of FTA guidelines will need explanation.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected mode against observed estimation data	• Prediction success likely to be low.	Level 3 (Debugging Test)
Aggregate	• Numbers of purpose specific tours by mode[1] and mode shares: • For the region • By district (at both tour origin and tour destination) • By area type (at both tour origin and tour destination) • Purpose specific tour mode shares by auto distance, transit in-vehicle time, walk time, or bicycle time • Purpose specific tour transit mode shares by walk time to transit • County-to-county purpose specific tour mode shares	• Compare modeled to expanded observed numbers of tours and tour mode shares • Review for reasonable patterns • Compare to 5-year CTPP home-to-regular workplace by means of transportation	Level 1 (Basic Test)
	• Purpose specific tour mode shares[1]: • Household size • Income group • Autos available per worker equal 0.0, greater than 0.0 but less than 1.0, equal to 1.0, and greater than 1.0 • Autos available per household member of driving age equal 0.0, greater than 0.0 but less than 1.0, equal to 1.0, and greater than 1.0 • Gender by age group	• Compare modeled to expanded observed numbers of tours and tour mode shares • Review for reasonable patterns • Compare patterns to those for other regions (e.g. Houston, Twin Cities) • Review regarding success of the model in "telling a coherent story"	Level 1 (Basic Test)

[1] Summaries by the following nine modes will be produced: Auto drive alone, Auto shared ride 2, Auto shared ride 3+, Auto ridesourcing, Transit walk access, Transit park and ride access, Transit kiss and ride access, walk, bike. School bus is a tenth submode available for K-12 school tours.

5.3.14 Intermediate Stop Generation

The intermediate stop frequency model first simulates the presence and number of stops on each half-tour. A distribution of stop purposes is then used to simulate the stop purpose by tour type, half-tour, time period, and person type. This approach does not consider inter-dependencies between the stop purposes (e.g. all stops may have the same purpose). However, the stop purpose is not used by the intermediate stop location model or the subsequent determination of the stop arrival and departure times or trip mode choice. We will consider replacing the stop purposes with a generic term to represent "intermediate stop," i.e. remove the intermediate stop purpose Monte Carlo simulation entirely. This will avoid confusion around the data that the model is actually producing.

Aggregate validation tests will focus on the numbers of intermediate stops for the various tour types and the variation in intermediate stop generation rates by various socioeconomic, demographic, and geographic market segments. The numbers of intermediate stops might be expected to vary by tour purpose and half-tour "direction." The recommended aggregate validation tests are listed in Table 5.22.

Table 5.22 Intermediate Stop Generation Model Validation Tests

<table>
<tr><th>AGGREGATION LEVEL</th><th>VALIDATION MEASURES</th><th>EXPECTED OUTCOMES</th><th>PRIORITY</th></tr>
<tr><td>Regional</td><td>• Comparison of model coefficients to those from other regions</td><td>• No expectations; comparison only.</td><td>Level 1 (Basic Test)</td></tr>
<tr><td>Disaggregate</td><td>• Prediction success of modeled numbers of intermediate stops by tour against observed estimation data</td><td>• Prediction success likely to be low</td><td>Level 3 (Debugging Test)</td></tr>
<tr><td rowspan="2">Aggregate (for Region)</td><td>• Number of stops by half-tour (16 options: 0 out/0 in, 0 out/1 in, 0 out/2 in, 0 out/3 in, 1 out/0 in…3 out/3 in) by primary tour purpose
• Mean number of intermediate stops per half-tour by primary tour purpose:
  • By time of day (time of day based on departure time to the primary tour destination or arrival time from the primary tour destination for each appropriate half-tour, aggregated to AM peak, mid-day, PM peak, evening, and night)</td><td>• Compare means of modeled numbers of stops to unexpanded means from surveyed numbers to determine whether they lay within the confidence intervals
• Review for reasonable patterns</td><td>Level 1 (Basic Test)</td></tr>
<tr><td>• Mean number of intermediate stops regardless of half-tour by primary tour purpose for each:
  • Gender / age group combination
  • Person type</td><td>• Compare modeled to expanded observed average numbers of intermediate stops
• Review for reasonable patterns</td><td>Level 2 (Debugging Test)</td></tr>
</table>

5.3.15 Intermediate Stop Location Choice

The stop location choice model predicts the location of stops other than the primary destination. The intermediate stop location model is, in essence, a destination choice model. Thus, validation tests similar to those used for the tour destination choice can be employed. Impedances between stops will be calculated for all half-tours with one or more stops. In other words, if a half-tour has no stops, the half-tour and its associated impedances will not be included in the summaries. The recommended validation tests are listed in Table 5.23.

Table 5.23 Intermediate Stop Location Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected ZONE against observed estimation data	• Prediction success likely to be very low	Level 3 (Debugging Test)
Aggregate	• Modeled versus observed (from household survey) stop impedance (e.g. trip distance) histograms by tour purpose with impedances based on: • Congested auto travel time • Straight-line travel distance • Percent of stops by out-of-direction distance for each tour purpose	• Modeled to observed averages should be ±5 percent • Modeled to observed impedance histogram coincidence ratios (reasonableness test only)	Level 1 (Basic Test)
	• Modeled versus observed (from household survey) average impedances by tour purpose stratified by: • Person type • Gender • Household income level	• Modeled to observed averages should be ±5 percent	Level 2 (Debugging Test)

5.3.16 Stop Time of Day

Time of day choice is predicted in 30-minute intervals (48 periods across the day). Unlike the tour time of day choice models, the stop time of day choice model predicts only a single time period for each stop activity. This time period corresponds to the activity start time if the stop occurs on an outbound half-tour or the activity end time if the stop occurs on a return half-tour.

At the point in which stop time of day choice is applied in the model chain, the start and end times of a tour's primary activity have already been simulated. Thus, the difference between stop start time and primary activity start time (outbound half-tour) or stop end time and primary activity end time (return half-tour) represents the duration of that stop (inclusive of the travel time between activity locations). Of course, an outbound half-tour stop must occur earlier than the primary activity start time and a return half-tour stop must occur later than the primary activity end time.

In cases where multiple stops exist on a single half-tour, the stops are modeled in sequence starting with the one closest in temporal proximity to the primary activity (the stop location choice model sequences stops the same way). For outbound half-tour stops, this means stops are modeled in reverse chronological order, and for return half-tour stops, this means stops are modeled in chronological order. In such cases, the second stop's timing would be bounded by the start/end time of the first stop's simulated time period (rather than that of the primary activity), and a third stop's timing would be bounded by the start/end time of the second stop's simulated time period.

Since the stop times implicitly include travel time, the validation tests will focus on the implied durations of the stops. The recommended validation tests are listed in Table 5.25.

Table 5.24 Stop Time of Day Choice Model Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from other regions	• No expectations; comparison only.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected TAZ against observed estimation data	• Prediction success likely to be low.	Level 3 (Debugging Test)
Aggregate (for Region)	• Histograms of stop durations by primary tour purpose by: • Half-tour direction (outbound or inbound) • Number of stops on half-tour • Tour mode	• Compare modeled to expanded observed data • Review for reasonable patterns	Level 1 (Basic Test)
	• Primary stop durations by tour purpose by: • Gender • Person type	• Compare modeled to expanded observed data • Review for reasonable patterns	Level 2 (Debugging Test)

5.3.17 Trip Mode Choice

Trip mode choice is similar to the mode choice performed for many trip-based models. However, trip mode choice in the ABM has the benefit of being conditional on tour mode choice. Trip mode choice validation tests shown in Table 5.25 will be similar to the tour mode choice validation tests described previously.

Table 5.25 Trip Mode Choice Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from: • other regions • SERPM 7 • FTA Guidelines	• Primarily comparison for reasonableness. Coefficients outside of FTA guidelines will need explanation (for use in New Starts model justification).	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected mode against observed estimation data	• Prediction success likely to be low.	Level 3 (Debugging Test)
Aggregate	• Numbers of purpose specific trips by mode[1] and mode shares by tour purpose: • For the region • By tour mode • Tour purpose specific trip mode shares by auto distance, transit in-vehicle time, walk time, or bicycle time	• Compare modeled to expanded observed numbers of tours and tour mode shares • Review for reasonable patterns • Compare to 5-year CTPP home-to-regular workplace by means of transportation	Level 1 (Basic Test)

5.3.18 Parking

The parking location choice model is segmented by work and non-work tours and applies only to the primary tour destination where the terminal end of the tour is a destination in a parking-priced MAZ. Modeled parking priced areas include downtown Miami, Fort Lauderdale and West Palm Beach, as well as the Jackson Memorial Hospital area in Miami. The parking model forecasts the parking location through the trade-off of walking distance between each candidate parking location and the trip destination is measured using the MAZ to MAZ distances measured over an all-streets network and the parking cost of the MAZ. The parking costs are obtained from the input MAZ data.

Due to lack of observed disaggregate parking choice data in SE Florida, the SANDAG parking lot choice model was transferred 'as is' to SERPM 7.0. It will be assessed and re-estimated for SERPM 8.0. Reasonable validation of the parking model will depend on the amount of information that can be gleaned from the 2015 travel survey. If insufficient 2015 survey data are available, the model validation tests will be assessed for reasonableness. The recommended validation tests are listed in Table 5.27.

Table 5.26 Parking Choice Validation Tests

AGGREGATION LEVEL	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Regional	• Comparison of model coefficients to those from: • other regions • SERPM 7	• Primarily comparison for reasonableness.	Level 1 (Basic Test)
Disaggregate	• Prediction success of selected mode against observed estimation data	• Prediction success likely to be low.	Level 3 (Debugging Test)
Aggregate	• Walk distance histograms for all stops in parking priced areas by tour purpose (work and non-work) by parking price area • For work tours where the stop is as the work location, average walk distances by employer parking provision (free, subsidized, and no free/subsidized parking) • Average walk distances by tour purpose by: • Income group • Gender • Person type	• Compare modeled to expanded observed data from 2015 travel survey • Review for reasonable patterns	Level 1 (Basic Test)

5.4 System Level Validation

Once the single-pass, stepwise calibration / validation is complete, the SERPM 8 model will be run in its normal application mode with multiple iterations of speed feedback and system level results will be checked. The full model run will include application of the visitor, external-external, internal-external, truck, and airport models. System level validation tests are focused on the results of highway and transit assignment procedures and represent some of the most traditional validation measures:

- For highways, validation tests focus on regional and subregional VMT, vehicular volumes and VMT by different facility types, screenline crossings, etc.

- For transit, validation tests are focused more on boardings and alightings by submodes, stations, and specific routes as well as transfer rates (boardings per linked trip).

The remainder of this section describes the planned system level validation tests for highways and transit in greater detail.

5.4.1 Highways

The system level highway validation is an overall validation of the travel modeling process with a very specific focus on the reproduction of observed traffic volumes in a region. Experience in other areas has shown that even a poorly specified model can be made to reproduce observed traffic volumes for a base year. The validations of the individual modeling procedures described above are intended to help ensure that the regional travel model is, in fact, reasonable. However, if the individual model components are all deemed to be reasonable but the modeled traffic volumes do not reasonably reproduce observed traffic volumes for a base year, the model is not valid.

The system level highway validation will focus on several classes of measures:

- Vehicle-miles of travel (VMT);
- Individual link traffic volumes;
- Intra-regional traffic flows as defined by screenlines; and
- Congested roadway speeds.

Table 5.28 summarizes the system level highway validation measures for the auto modes, which will focus much more specifically on numerical guidelines than the validation of the individual model components. A primary reason for this is the availability of independently collected data such as traffic counts for the system level validation.

Table 5.27 System Level Highway Validation Tests

VALIDATION FOCUS	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY

Vehicle-miles of travel	• Comparison of modeled VMT to VMT estimated from traffic counts by:
	  • region
	  • functional class (with particular attention to toll roads, HOT, and HOV facilities)
	  • area type
	  • county
	• Comparison of total modeled VMT per resident to other regions
		• Modeled regional VMT should be within the following percentages of estimated VMT:
		  • ±1 percent for the region
		  • Percentages shown in Table 5.29 by functional class, area type, and county on links with counts
		• Modeled VMT per person per day should be around 25-35[1]
			Level 1 (Basic Test)

Individual link volumes	• Scatterplot of modeled versus observed daily traffic volumes by link
	• Root mean squared error (RMSE) and percent RMSE by:
	  • region
	  • functional class (with particular attention to toll roads, HOT, and HOV facilities)
	  • volume group
	• Coefficient of variation by volume group
	• Anomalous links
	  • 100 links with the greatest absolute differences between modeled and observed volumes
	  • Links with 0 volumes
	  • Links with very high v/c ratios
		• Visual inspection for large errors in modeled link volumes or for general trends in errors
		• RMSE and %RMSE by functional class should be within targets shown in Table 5.30; %RMSE by volume group should be within the targets shown in Figure 5.3
		• Coefficient of variation by volume group should be within the targets shown in Figure 5.4
			Level 1 (Basic Test)

Screenlines	• Percent deviation by screenline	• Percent deviation should be within the targets shown in Figure 5.5	Level 1 (Basic test)
Congested Speeds	• Comparison of modeled and observed average congested speeds by functional class (with particular attention to toll roads, HOT, and HOV facilities) and area type for each time of day • Scatterplots of modeled to observed congested speeds by time of day for links with speed data	• Average assigned congested speeds should be within ±5 miles per hour of average observed congested speeds • 90 percent of the link specific assigned congested speeds should be within ±5 miles per hour of average observed congested speeds	Level 1 (Basic Test)

[1] The 2001 NHTS documents the range as 17-24 VMT per person per day. This range was referenced in the Model Validation and Reasonableness Checking Manual – Second Edition. More recent data in many areas show higher values. Since values can vary by region depending on a variety of locally specific factors, an estimate of the current value for the SERPM region should be obtained from local sources (and perhaps the existing trip based model) for use in comparison.

Table 5.29 presents the targets for differences between assigned and observed VMT by functional class and area type, which is the state of practice guidance. SERPM's link speeds are based on posted speeds so the VMT differences by functional class will also be stratified by posted speed. The overall targets by functional class in Table 5.29 will also be used for each posted speed category. Failure to satisfy the overall targets by posted speed will prompt an investigation of assignment procedures. Table 5.30 shows the targets for RMSE and RMSE% by functional class. Figure 5.3, 5.4 and 5.5 present the assignment results by RMSE% by volume group, percent difference by volume group, and screenline scatterplots from various studies.

Table 5.28 Traffic Assignment VMT Targets by Functional Class

Stratification	VMT Target[1,2]
Functional Class	
Freeways	±7%
Expressways	±7%
Principal Arterials	±10%
Minor Arterials	±10%
Collectors	±15%
All Links	±1%
Area Type	
CBD	±10%
Fringe	±10%
Urban	±10%
Suburban	±10%
Rural	±10%
Total	±1%
Total	±1%

Source:

1 Adapted from Giaimo, Gregory, Travel Demand Forecasting Manual 1 – Traffic Assignment Procedures, Ohio Department of Transportation, Division of Planning, Office of Technical Services, August 2001.

2 VMT by functional class will also be stratified by posted speed of the link.

Table 5.29 RMSE and %RMSE Targets by Functional Class[1]

Functional Class	Target Values	
	RMSE[2]	%RMSE[3]
Freeways	<12,500	<20%
Expressways	<7,500	<30%
Principal Arterials	<3,750	<30%
Minor Arterials	<3,000	<40%
Collectors	<2,250	<70%
All Links	n/a	<40%

1 $\text{RMSE} = \sqrt{\frac{\sum(\text{Count}-\text{Assigned})^2}{(\text{Number of Observations}-1)}}$

$\%\text{RMSE} = \frac{100\times\text{RMSE}}{\text{Avg. Count}}$

2 Based on one-half lane of capacity and assumption of 8 percent peak hour factors for interstates, freeways, system ramps, expressways, and external connectors; based on 10 percent peak hour factor for other functional classes

3 Rules of thumb from model validation efforts for several regions.

Figure 5.3 Traffic Assignment %RMSE Targets by Volume Group

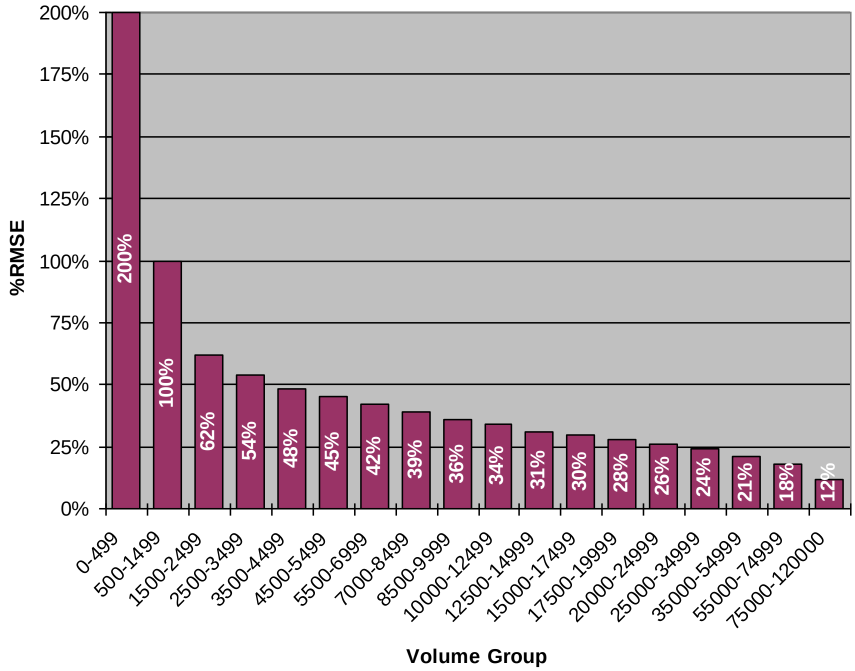
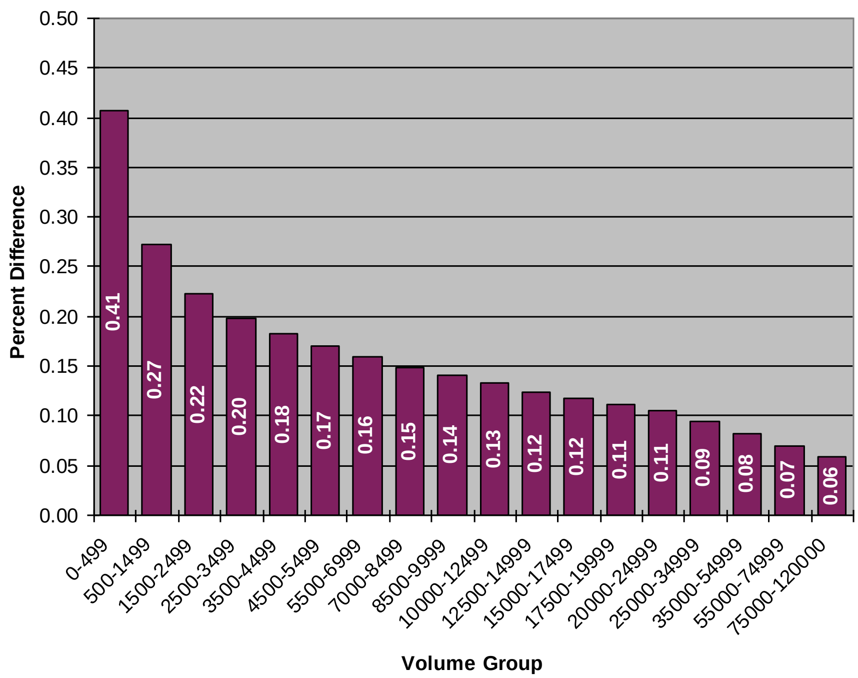
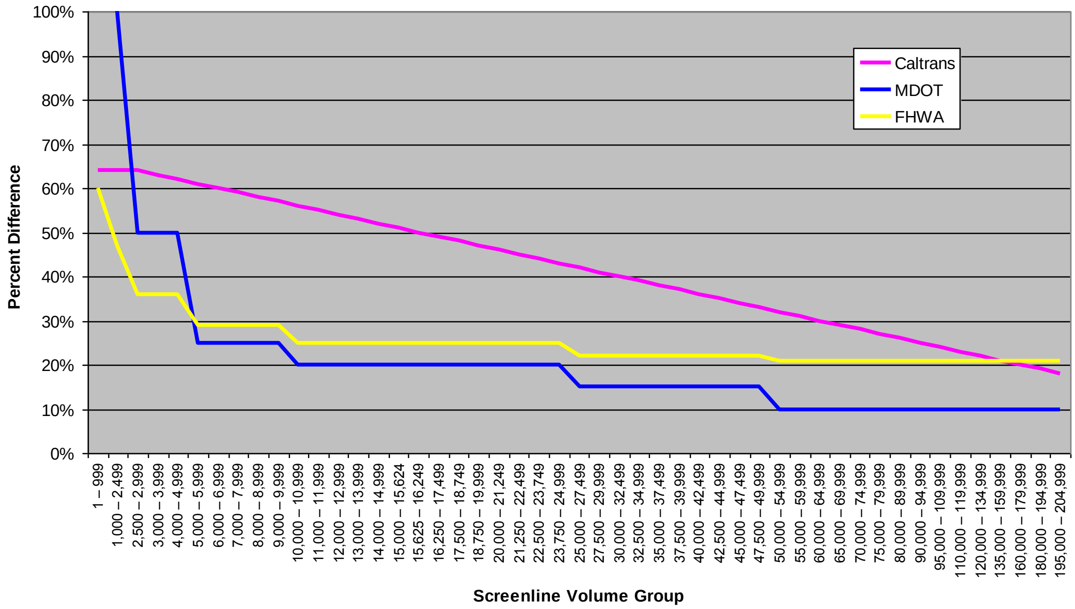


Figure 5.4 Traffic Assignment Percent Difference by Volume Group[1]



1 $$Percent\ \text{Difference} = \frac{|\text{Assigned} - \text{Count}|}{\text{Count}}$$

Figure 5.5 Screenline Crossing Percent Difference by Volume Group[1,2]



1 $\textit{Percent Difference} = \frac{|\text{Assigned} - \text{Count}|}{\text{Count}}$

Adapted from *Las Vegas Travel Demand Model Guidelines for Estimation, Calibration, & Validation*, prepared for Regional Transportation Commission of Southern Nevada, prepared by Fehr & Peers Transportation Consultants, March 25, 2005, page 28.

2 The MDOT standard may have recently changed to 10 percent for all cutlines and 5 percent for all screenlines.

5.4.2 Transit

The system level transit validation is an overall validation of the travel modeling process with a very specific focus on the reproduction of observed transit boardings and transit volumes in the region.

The system level transit validation will focus on several types of measures, including:

- Person-miles of travel (PMT) on transit;
- Transit boardings by route, station, and access mode;
- Boardings per linked trip;
- Transit screenlines; and
- Park-and-Ride lot utilization.

Table 5.31 summarizes the aggregate validation measures that are proposed for the system level transit validation.

Table 5.30 System Level Transit Validation

VALIDATION FOCUS	VALIDATION MEASURES	EXPECTED OUTCOMES	PRIORITY
Boardings by mode by time of day	• Comparison of modeled to observed boardings by: • region by time of day • mode by time of day • line and station by time of day and access mode	• Modeled boardings for the region should be within ±5 percent of the estimated boardings • Modeled boardings by mode by time of day should be within ±10 percent of the estimated boardings • Modeled boardings by line by time of day should be within ±20 percent of the estimated boardings	Level 2 (Basic test)
Park-and-Ride lot utilization	• Modeled daily drive access vehicle trips to park-and Ride lots to observed parking at lots	• Modeled drive access trips to lots should be in the range of 1.0 to 2.0 to account for turnover of parking spaces • Modeled boardings per observed parked vehicle should be in the range of 4-10 for each park-and-Ride lot	Level 2 (Basic Test)
Person-miles of travel	• Comparison of modeled PMT to PMT estimated from boarding and alighting counts (or PMT estimated from transit on-board survey data) by: • mode • access mode	• Modeled PMT for the region should be within ±5 percent of the estimated PMT • Modeled PMT by mode by time of day should be within ±15 percent of the estimated PMT • Modeled PMT by access mode should be within ±15 percent of the estimated PMT	Level 3 (Debugging Test – Depends on time required to process detailed boarding and alighting count data)
Transit screenlines	• Modeled daily screenline and cordon line "on-board" ridership to observed ridership	• Modeled on-board ridership should deviate from the observed ridership by less than the percentages shown in Figure 5.4 for highway screenlines	Level 3 (Debugging Test – Depends on time required to process detailed boarding and alighting count data)

5.5 Temporal Validation and Sensitivity Tests

5.5.1 Temporal Validation

Temporal validation (see Section 3.4) involves using the estimated/calibrated model to make forecasts for a year other than the base year, and requires data for (at least) one other year. Validation tests for the forecast year could include any tests discussed above, but are usually limited to system level tests.

The extent of temporal validation that will be performed will be based partially on the resources available. At a minimum, system level validation to the backcast year will be performed, which will be considered priority level 1 for temporal validation. Priority level 2 temporal validation tests to the backcast year will include aggregate validation tests of the model components described in the previous subsections. Of course, results from these temporal validation checks are not likely to be as good as the tests performed for the base year of the model. Differences in survey or sampling methods, random error associated with surveys, and changes in travel behavior over time (which the model may not be sensitive to) can all create inconsistencies in temporal validation.

5.5.2 Parameter / Variable Sensitivity Testing

One goal of activity-based models is an increased sensitivity to model inputs. Sensitivity testing (see Section 3.5) 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.).

Like temporal validation, sensitivity tests can be performed for any of the validation measures described in the previous subsections. Typically, however, data are not available for which to compare the results of sensitivity tests. Instead, sensitivity tests should be reviewed for reasonableness, with expected outcomes of the tests shaped beforehand. Any unexpected outcomes observed from the tests should be explainable. Specific sensitivity tests will be developed in consultation with FDOT District 4 staff.

5.6 System Integrity Tests

The system integrity tests involve running the software to ensure that it works as documented. Actual tests cannot be designed until the software design and coding are complete. However, the following general tasks will be performed:

- Verify that the software can be successfully installed on microcomputers running the Windows operating system.
- Verify that program modules can be invoked as documented.
- Perform selected calculation verifications.
- Review the output database for unexpected missing information.

6.0 References

Cambridge Systematics, Inc. (2010). *Travel Model Validation and Reasonableness Checking Manual – Second Edition*, Travel Model Improvement Program, Federal Highway Administration.

Cambridge Systematics, Inc. (2016). *Model Design Plan for SERPM 8.0 Model*. Prepared for Florida Department of Transportation District 4.

Federal Transit Administration (2006/2007/2009). Travel Forecasting for New Starts Proposals. Workshop materials provided at http://www.fta.dot.gov/12304_7275.html (Retrieved April 12, 2013).

Madre, J-L, Axhausen, K., and Brög, W. (2007). Immobility in travel diary surveys. Transportation, Volume 34, Number 1, pp. 107-128(22).

Richardson, A. J. (2006). Fandangles and Other Measures of Incidental Trips. Transportation Research Record 1985, Transportation Research Board of the National Academies, pp. 106-114.

Smith, M.E. (1979). Design of Small Sample Home Interview Travel Surveys. *Transportation Research Record 701*, Transportation Research Board.