DESIGN OF FLORIDA'S TURNPIKE'S STATE INTEGRATED LAND-USE TRAVEL DEMAND MODEL

Thomas Adler
Resource Systems Group
tadler@rsginc.com

Youssef Dehghani
PB Americas
dehghani@pbworld.com

Michael Doherty
URS Corporation
michael_doherty@urscorp.com

1. INTRODUCTION AND BACKGROUND

This paper describes the design of the Turnpike State Model (TSM) for the Florida Department of Transportation (FDOT) Turnpike Enterprise. The TSM is being used to evaluate the feasibility of potential intercity highway projects that extend beyond the boundaries of available urban and regional travel models.

A combination of tools and databases has been used over the last 20 years to evaluate FDOT's proposed intercity projects, including the Florida Standard Urban Transportation Structure (FSUTMS) regional models and other available databases as well as the statewide model developed by FDOT in the early 1980s. The obsolescence of the initial statewide model created a need for a new and improved version. For this next generation of Turnpike's statewide model (the TSM), a modern database version was desired that would provide several notable improvements over the original.

- It would be based on readily available and updatable databases supported by FDOT including the GIS-based GDT/TeleAtlas roadway network, the Roadway Characteristics Inventory (RCI), funded and scheduled Work Program roadway improvements, and Florida's Traffic Information database of annually updated traffic counts and characteristics.
- Incorporate land use data for existing population and employment from the US Census and the InfoUSA databases in a way that is independent of the model's zone structure.
- Rely on independent and updatable forecasts of population and employment provided by Florida's Bureau of Economic and Business Research (BEBR) and US Department of Commerce, Bureau of Economic Analysis (BEA).
- The allocation of forecasted population and employment would be made more sophisticated by the incorporation of a Land Use Allocation Model (LUAM) within the model structure.
- Be independent of local data availability, work schedules and data quality.
- It would have a sharper focus on inter-city travel than the previous SWPM by incorporating more traffic analysis zones and it would have a simplified structure optimized for toll applications.
- Accuracy and simplicity would be enhanced for intercity levels of analysis by the incorporation of a powerful Matrix Estimation (ME) procedure.

The resulting TSM created under this framework is an integrated transportation and land use model that is built on a database framework in which positional data, network attributes, land use information, and other data are maintained in a consistent format that can be easily updated with future year information. Positionally accurate GIS data are associated with a more traditional link network in a way that allows the network to be viewed and edited with full GIS shape details for the roadways and zones, but allows the model calculations to be conducted using the much more efficient link network. Accompanying this work was an extensive validation effort directed at both the model itself and the input data sources. The validation effort used automated tools and extensive manual checks to ensure integrity of the model and underlying data.

The TSM includes a land use allocation module (LUAM) to allocate user-defined land use control totals (e.g., at county level) to zonal level. The allocation model takes into consideration three sets of effects:

- Current households and employment densities
- Developable land (excluding water and other managed land)
- Measure of zone-to-zone transportation accessibility, produced from the travel component of the TSM

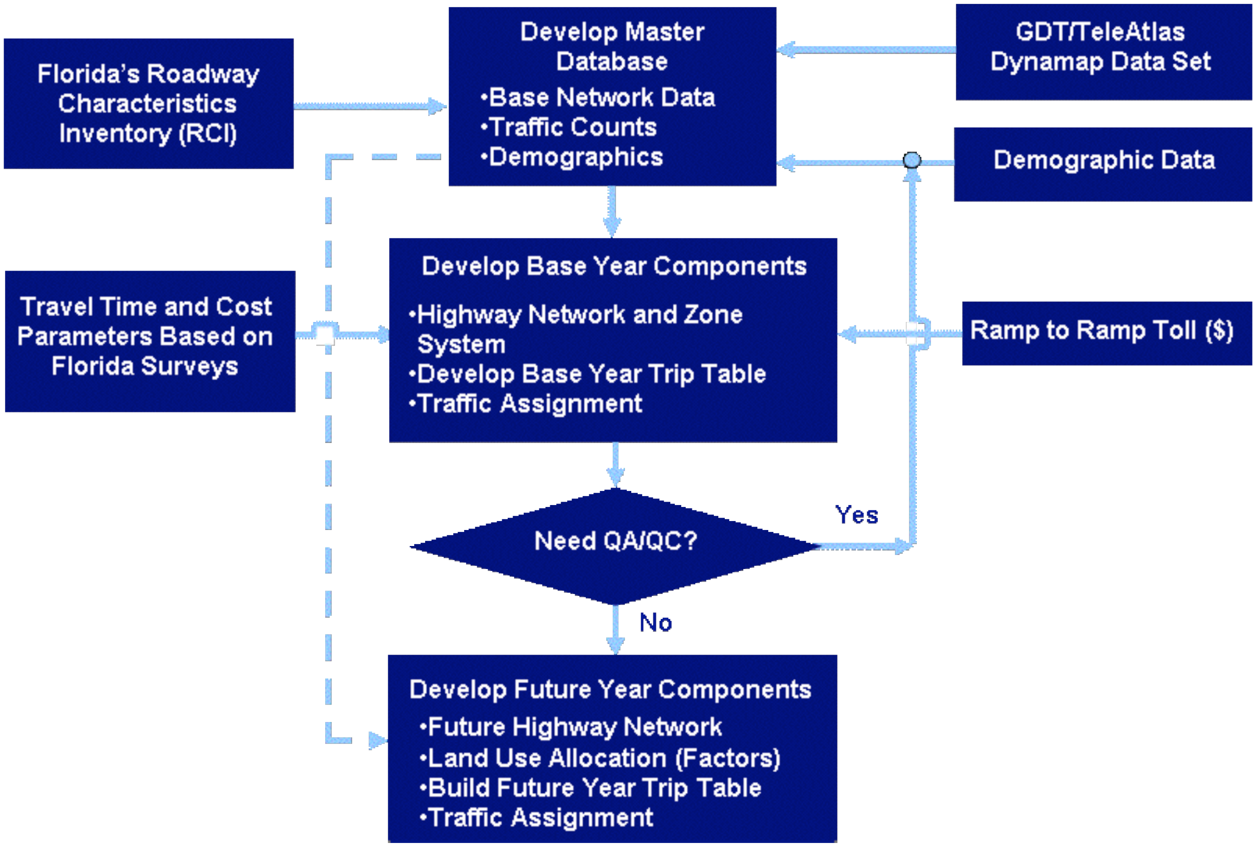
Another unique feature of the TSM is the application of matrix estimation (ME) for the development of a statewide base-year trip table. This key feature simplifies the model structure while improving its volume to count and trip length simulation accuracy.

The TSM development was staged in two phases to meet the immediate needs of Turnpike Enterprise projects and allow the additional time required to develop toll mode choice model capabilities. This paper describes the modeling work completed during the first phase, which involved developing a daily statewide travel model with a traditional toll modeling capability to meet the schedule for current projects. Key features of the current TSM components include the following:

- Automated procedures for generating a highway network with necessary data fields from FDOT's GDT/TeleAtlas dataset
- Automated procedures for developing traffic zone boundaries reflecting natural barriers and population/employment densities
- Easily updatable land-use Census and Bureau of Economic and Business Research (BEBR) estimates
- Reliance on a fully integrated database model
- Reliance on survey data and traffic counts to establish baseline travel patterns
- Reliance on state-of-practice procedures for toll forecasting analysis for auto passenger and truck traffic
- Reliance on a land use allocation model for estimation of zonal-level growth in demographic variables
- Automated procedures for displaying model results on a GIS-based network

A simplified schematic representation of the key TSM components is highlighted in Figure 1.

Figure 1: Schematic Flowchart of Florida's Turnpike State Model (TSM)



2. MODEL DATA

The first phase of the statewide database model development relied on readily available data. The most salient data items that were assembled include:

- U.S. Census Journey-to-Work (JTW) data collected in 2000
- U.S. Department of Transportation's National Household Travel Survey (NHTS) collected in 2001
- Zonal-level demographic data for 2000 was obtained from the U.S. Census and InfoUSA®
- County-level demographic future forecasts obtained from University of Florida's Bureau of Economic and Business Research (BEBR) databases and the U.S. Bureau of Economic Analysis (BEA)
- Databases maintained by FDOT Central Office, which were provided to the Turnpike Enterprise for this model development effort, include:
  - ✓ GIS-based GDT/TeleAtlas network database (2004) for all 67 counties.
  - ✓ Florida's Roadway Characteristics Inventory (RCI) and Work Program.
  - ✓ Traffic counts database available in the FDOT's Traffic Information System for all roadway segments contained in that database.

A brief description of each data item, including how they are used in the TSM, is presented below.

2.1 Survey Data

The U.S. Census Journey-to-Work and U.S. Department of Transportation's National Household Travel Survey are two key data items that were used to establish current travel patterns and provide inputs for creating a "seed" trip matrix. The NHTS data include the full range of trips made by a household on a given day, including long-distance trips. While the NHTS dataset includes the full range of trips made by about 1,100 resident households, the overwhelming majority of trips are 50 miles or less in length: trips over 50 miles in length constitute only 2.4 percent of all trips, over 100 miles are only 0.6 percent of trips and over 200 miles are only 0.3 percent of trips.

A large-scale statewide toll choice survey was also conducted to support estimation of values of time and a toll choice model to be used in the toll modeling.

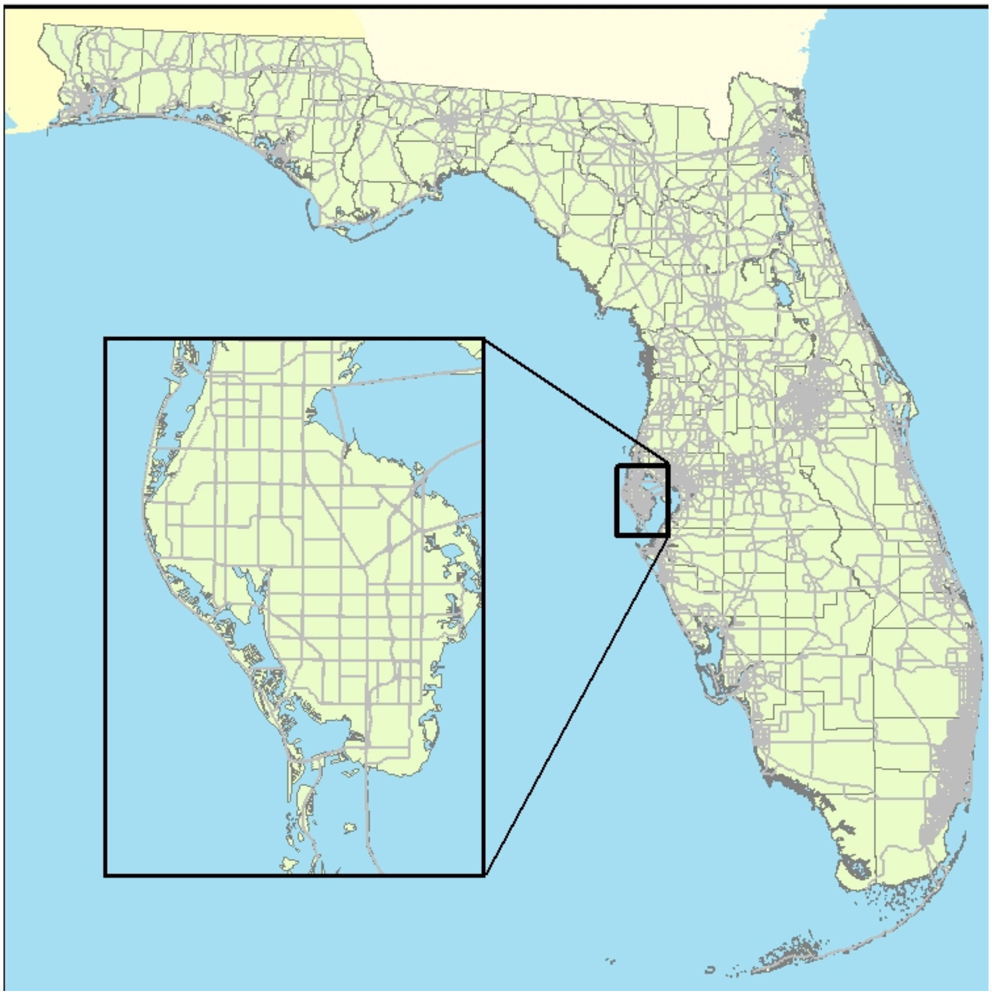
2.2 Roadway Network Data

Available travel demand models generally covered the state's urban areas and metropolitan regions but left rural areas untouched. One possible approach for building this new state-level model would have been to merge the networks and data from those existing models, as has been done in other states. However, even though the models were all based on a standardized approach known as the Florida Standard Urban Transportation Modeling Structure (FSUTMS), there are enough inconsistencies (e.g., different base years, validation parameters, and levels of detail) that simply merging the networks would have created a state network with a number of important and irresolvable inconsistencies. Therefore, a completely new network was constructed. The model network was specified to include all roads with a state-level function. This includes all roadways that currently serve other than local functions, as well as those local roads that provide primary access to significant land parcels.

A good way to construct such a network is to use the data compiled by one of the major digital map providers. Those data generally contain basic information about the road network: road alignments accurate to within a meter or so, road classification, roadway names, directions of flow, and the basic topology of the road intersections. However, they do not currently contain information about the numbers of lanes, posted speed limits or other details needed to calculate traffic capacities. Fortunately, the Florida Department of Transportation (FDOT) maintains a complementary dataset, the Roadway Characteristics Inventory (RCI), which contains those necessary details. As a result, the information needed to construct a transportation model network already exists and is actively maintained (updated and corrected as necessary) by others for purposes other than travel forecasting.

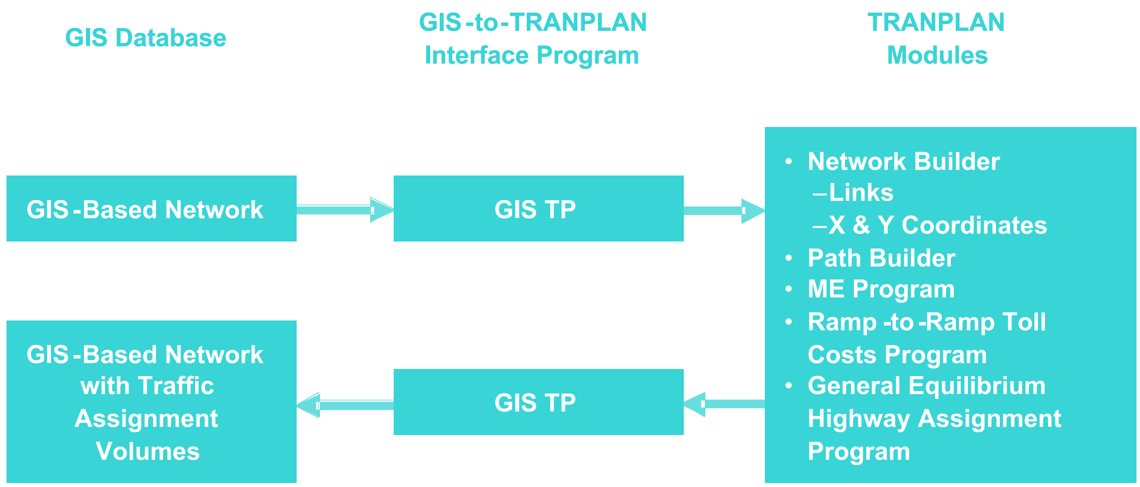
The TSM's roadway network was built using these existing data sources. The road alignments and basic roadway features were drawn from the most recent GDT/TeleAtlas Dynamap® coverages for the 67 Florida counties. Those coverages include a total of 1.9 million roadway arcs (segments). Screening out local road segments reduced the total to just over 415,000 arcs ("arc network"). Figure 2 shows the resulting base year network used in the TSM.

Figure 2: Florida's Turnpike State Model (TSM) Base Year Network



A set of network-builder tools, GISTP (developed by Fennessy Associates, depicted in Figure 3) was used to create a topologically correct TRANPLAN-format network consisting of approximately 100,000 model links ("link network"), each of which is associated with a set of arcs that define the true shape of the segment. The network can be viewed with the full alignment details using a GIS file viewer, and the smaller link version can be used in network modeling software. The factor of four in size reduction from the arc network to the link network is very significant in the performance of the model, affecting model run times and file sizes in almost direct proportion. Given the size of the model, these savings are very significant; they allow the model forecasting process to be computationally tractable with current computer hardware and will allow the model to be run considerably more efficiently in "production mode" as computer hardware improves.

Figure 3: Schematic Process between GIS-Based and TRANPLAN-Based Network

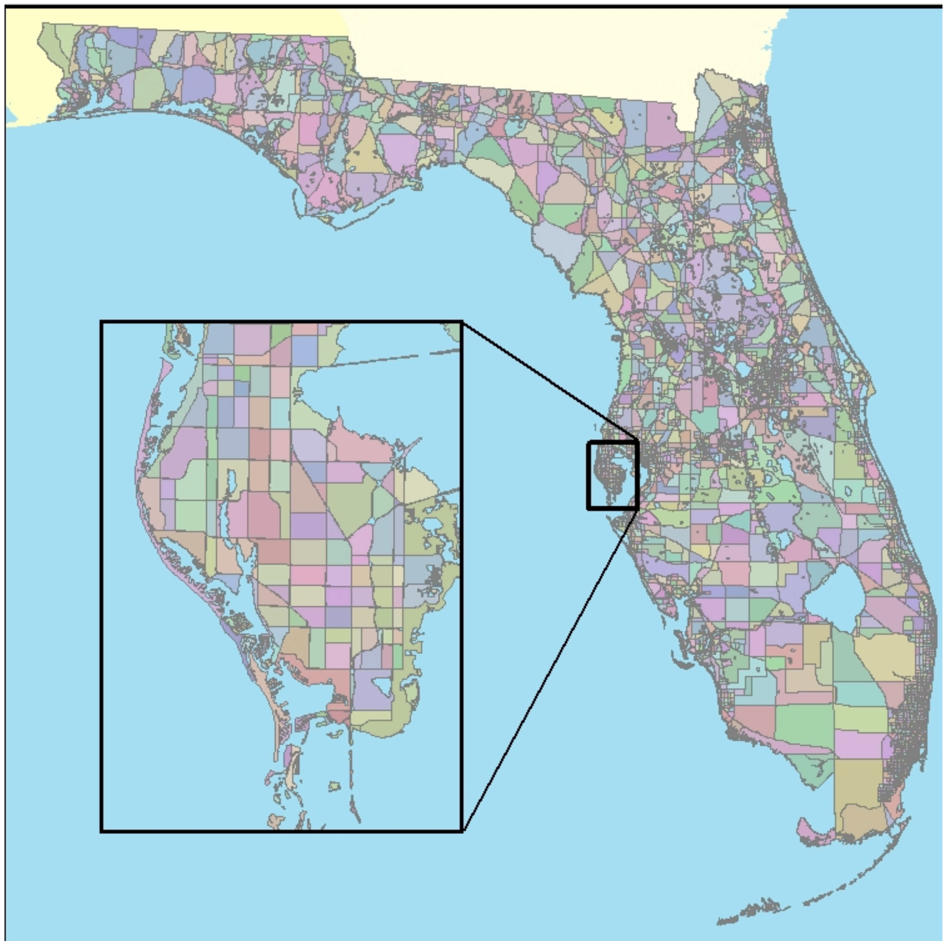


Each roadway section has accompanying data describing the roadway classification/function, name/route, location, lane geometry and speed limit. These data were derived by conflating the FDOT's RCI database with the GDT/TeleAtlas Dynamap® data.

Traffic counts, and the locations of those counts, were taken from FDOT's Traffic Information System for all roadway segments contained in that database. The database is annually updated and reflects Annual Average Daily Traffic (AADT) and percentage of truck traffic. Over 11,500 roadway segments (links) on the state roadway facilities have associated traffic counts, representing approximately 11 percent of the model's network. Traffic capacities were calculated for each link based on facility types, area types (urban vs. rural) and numbers of lanes. Florida currently has a total of 41 tolled facilities operated by a number of separate agencies, the largest of which is Florida's Turnpike. For each of these facilities, point-to-point toll matrices were also assembled, and these are overlaid on the network in the trip assignment process.

Transportation analysis zones (TAZs) were defined using a unique approach that is enabled by the fact that the model's roadway network is geographically accurate. The arc network was used as a template for the design of transportation analysis zones. Zones were formed from the boundaries created by the road alignments so that no roads fell in the interior of the zones. The result was a total of almost 27,700 zone polygons. About 23,100 of these were determined to be "undevelopable" because they are waterways, parks, highway medians, enclosed ramp areas or other locations that are clearly not suitable for traffic-generating land uses. The remaining 4,600 TAZs include 37 external zones connecting to Georgia and Alabama. Figure 4 shows the model's zone structure. Approximately 25,000 connector links were added to connect the zones to the roadway network, using a combination of automated and manual methods.

Figure 4: Florida's Turnpike Enterprise State Model (TSM) Zone Structure



3. MODEL VALIDATION

The focus of many model validation efforts is on modifying/refining the travel demand equations and other parameters in a model to ensure that they accurately reflect actual behavior (e.g., traffic counts). However, the travel demand modeling process in the Florida's Turnpike Enterprise State Model was designed around a matrix estimation procedure that ensures the model will closely replicate observed traffic counts. The matrix estimation approach was chosen so that a significant fraction of the initial effort could be devoted to validation of the data that underlie the model. In many other modeling efforts, the bulk of the validation effort is spent on calibrating the travel demand model parameters rather than on improving the data from which the models are built. With matrix estimation, validation can instead shift to ensuring integrity of the input data.

The TSM design concept was, wherever possible, to: 1) use data that had already been assembled (mostly for other purposes), which are internally consistent and actively and reliably maintained and 2) use automated procedures, wherever possible, to create and update the model using those data. The primary data sources that were selected – GDT/TeleAtlas Dynamap®, RCI and the Traffic Information System – all met these criteria, and automated procedures to create the model – GISTP – existed at the time work on the model began. There were a number of initial challenges that involved significant labor effort in cleaning the Dynamap® files sufficiently to create closed polygons throughout the state so that the automated TAZ-generation process could be used. However, the more difficult validation tasks were to verify the internal validity of each of the individual datasets and to ensure that they were mutually consistent.

An approach known as "matrix estimation" (ME) was used to create the base year trip table. Matrix estimation determines zone-to-zone trip volumes based on observed traffic counts. The method takes as inputs:

- A transportation network with observed traffic counts, where applicable
- Zone-to-zone paths – for equilibrium assignments, the paths are determined for each iteration of the assignment
- Iteration fractions from the equilibrium assignment
- Weights by facility type – representing the reliability of certain counts
- An initial estimate of the zone-to-zone traffic flows (this is referred to as the "seed" trip table)

ME assign trips to the network paths from the network assignment model, totals the resulting trips at each link on which a count is located, and compares those totals to the observed counts. It then adjusts the zone-to-zone traffic volumes up or down depending on whether the totals at the links through which those flows pass are below or above the observed traffic counts. The ME program performs several user-controlled passes to reassign the "new" trip table and to continue adjusting to the counts. The basic technique is an iterative process: network assignment, ME program, reassign the new trip table, rerun the ME program, etc., until the resulting zone-to-zone traffic volume table, when assigned to the network, very closely matches observed traffic volumes. The primary advantages of the matrix estimation approach are that it requires very limited input data compared to conventional approaches and it results in a travel demand estimate that very closely replicates observed traffic counts.

The initial matrix estimation runs using a seed matrix of all ones, and with the issues listed above, nevertheless converged to a solution with an RMSE of only 33 percent and with errors by facility type well within commonly accepted guidelines. However, that solution was far from correct, insofar as the underlying data forced a solution that was clearly affected by the various data errors that were identified and corrected. These errors resulted in clearly distorted estimates for a proposed new major roadway when the model was applied in its uncorrected form.

After the data issues were resolved and the matrix estimation procedure was applied, the root mean squared error was reduced to 15 percent, after a calculation of new paths and re-assignment of the final trip table (12 percent before the re-assignment). Matrix estimation adjusts the trip table to match a given set of paths that are, in this case, generated using a user equilibrium assignment. Because of the size of the path files generated by the model (~20 GB), it is more efficient to run several matrix estimation iterations for each assignment. Each of these matrix estimation iterations results in improvements for the given set of paths files, but some of this improvement is lost when the trip table is applied to a new set of paths. A new set of paths was generated after every five matrix adjustments.

Figure 5 shows the RMSE resulting from the user equilibrium assignment of the final adjusted trip table produced from the matrix estimation procedure. As noted earlier, both the overall RMSE and the RMSEs by volume category were substantially improved by the corrections made as part of the validation effort. With the exception of the lowest-volume roadways, which are deliberately not fully represented in the model network, all of the RMSE values are well below maximum recommended levels.

Figure 5: Matrix Estimation Results by Volume Category

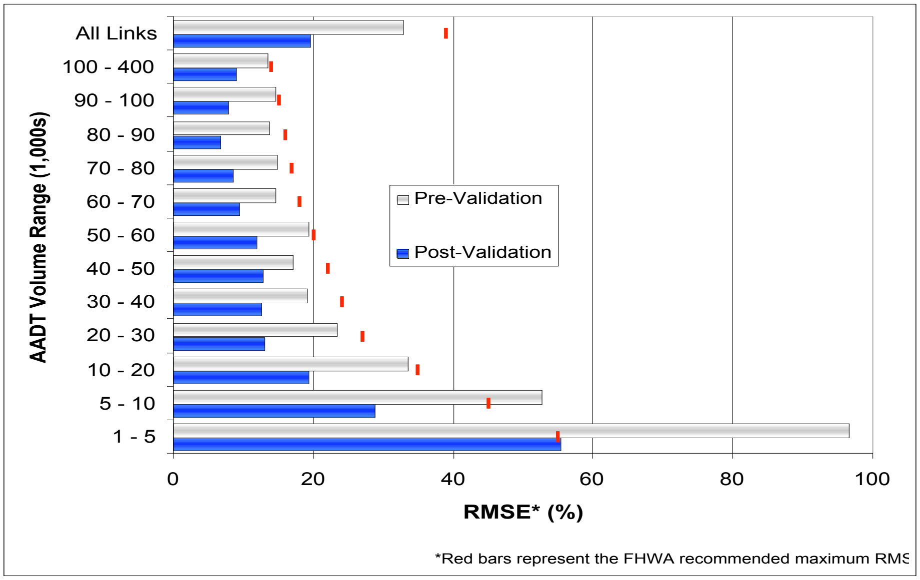
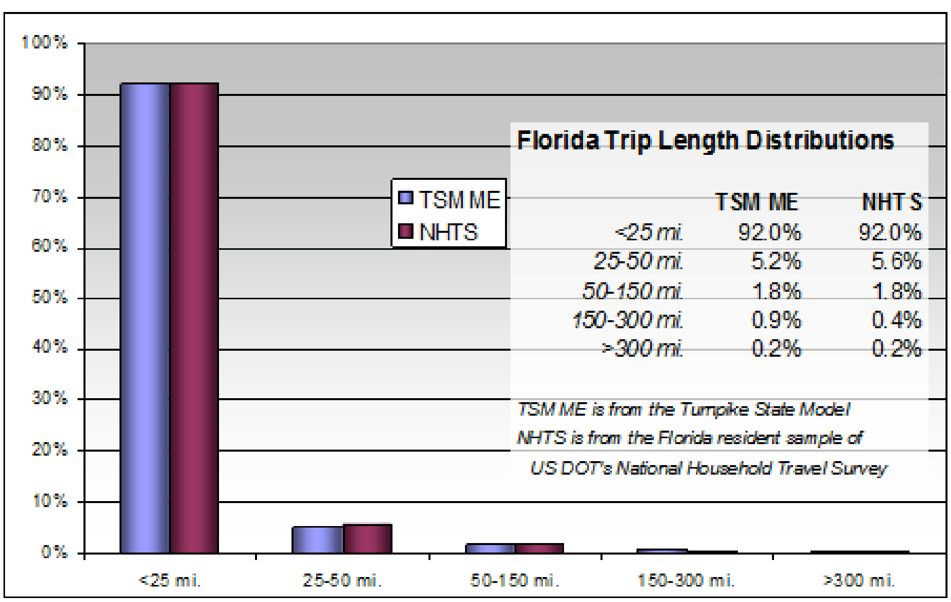


Figure 6 compares the trip length distribution by distance category, indicating a close match between estimated and observed values.

Figure 6: Trip Length Distributions Comparison (in miles)



4. TOLL MODELING PROCEDURE AND APPLICATIONS

The current version of the TSM relies on a ramp-to-ramp toll application during the equilibrium assignment process. The ramp-to-ramp procedure is superior to other toll modeling procedures, which tend to approximate the toll structures on many toll facilities. The EQUILIBRIUM HIGHWAY ASSIGNMENT program in TRANPLAN was modified to perform the ramp-to-ramp toll methodology. As in other toll modeling procedures, the toll costs are converted to toll time equivalents using an estimate of "willingness-to-pay." Estimates of willingness-to-pay were derived from the statewide toll choice survey. The conversion factor is expressed as "CTOLL" in minutes per toll costs in cents. The program reports all ramp-to-ramp movements and provides an estimate of revenue. Also, the program was modified to accept ramp-to-ramp inputs for each of the 41 toll entities (or companies) in the State of Florida.

The first set of applications involved feasibility analyses of proposed new intercity toll highways. These analyses used the model in a conventional way, forecasting future traffic volumes based on future scenarios that included other already-planned transportation facilities, future growth as projected at the county level by other state and federal agencies (with allocation of growth within the counties using an integrated land use model) and pricing comparable to other new toll facilities. The results of these runs were useful in establishing ranges of volumes and revenues that could be expected under these scenarios. However, the land use scenarios and forecasts did not encompass the potential effects of concerted public-private partnerships to develop the areas served by the new corridors. One of the key questions was how much new growth would be required to support the new facilities. To address this question, the model was used in a very different way. A financial analysis was conducted to determine the revenues and traffic volumes required to finance the new facilities. Those traffic volumes were then treated as control counts, and a matrix estimation procedure was used to determine the origin-destination volumes required to produce those counts. Finally, these volumes were then translated into the population and employment growth that would support the facilities.

5. CONCLUSIONS

The Florida's Turnpike State Model (TSM) effort successfully developed and implemented a database-driven statewide modeling system encompassing an integrated land use and travel demand model. The current model databases need to be maintained and periodically updated as base GIS network, traffic counts, population and employment updates become available.

Future phases of the Turnpike's state modeling effort will include development of behaviorally based travel demand modeling procedures, including trip distribution and toll mode choice modeling capabilities. This work could replace the CTOLL process and matrix estimation processes and further refine the other key model components. It is expected that these efforts will improve the ability of the Turnpike State Model (TSM) to support the feasibility assessment of future Turnpike projects and initiatives.