MINIMUM TRAVEL DEMAND MODEL CALIBRATION and VALIDATION GUIDELINES FOR STATE OF TENNESSEE

By:

Dr. Fred Wegmann
Jerry Everett
The University of Tennessee
Center for Transportation Research
Knoxville, Tennessee

Calibration and Validation Guidelines

Introduction

The calibration and validation of travel demand models is essential to accurately model current/future travel for a metropolitan area. In many areas there are now specific requirements that must be met. For example, the transportation conformity guidelines require that the travel model set used to prepare the conformity analysis have a validation that is not more than 10 years old.

A major shortcoming of many travel demand forecasting (TDF) models is the lack of attention and effort placed on the model evaluation and reasonableness-checking phase. There is a trade off between increasing the level of accuracy of a model and the cost associated with additional data collection, calibration, and validation. There are, however a number of inexpensive evaluation and reasonableness checks that can be performed to enhance TDF model's forecasting ability.(1)

It is important that these checks be performed after each step of the four-step modeling process. Too often the modeler proceeds through four steps and then conducts tests based on the overall results of the travel model. The preferred approach is to apply evaluation and reasonableness checks during the process of calibrating each individual modeling step. After each step has been validated, the overall model is validated. If this incremental validation approach is not followed the problem of "error propagation" will occur. Error propagation occurs when errors made in each step are compounded and thereby increase the overall modeling error.

Travel modes rely on several primary data sources for model calibration and validation.

1. Accurate estimates of base year traffic analysis zone (TAZ) household characteristics and employment information
2. An accurate representation of the base year highway and transit network
3. An accurate base year travel survey and
4. Accurate base year ground counts are all needed for model calibration

The household survey data are used to develop mathematical relationships or functions, which relates observed travel behavior to variables that can be more easily and directly forecasted. In general, the model parameters that rely on travel survey information are assumed to remain constant between the model year and the model forecast year.

Travel survey information is sometimes lacking, as financial resources often are not available to conduct a full set of travel surveys. When this is the case, part of the model may have to be developed using "borrowed" travel survey data from another similar urban area that has conducted a travel survey or by relying on information from a potentially

outdated travel survey or by having to prepare model specifications with limited or incomplete travel survey information.
Network information for the base year is generally available and with the GIS and other software now available, base year highway and transit networks can be developed with a high degree of accuracy, provided care and quality checks are employed during preparation.

Base year ground counts in sufficient numbers of locations and accuracy are required for base year model validation. The MPO or city traffic-engineering department often collects these data locally. Often the local traffic count data will be supplemented by data collected by the state DOT.

Calibration and validation occur **after** the original "estimation" or development of the model set. This report is NOT intended to be a manual on travel demand model development. This report assumes that the final model structure including the selection of relevant variables and the specification of the initial parameters has already been determined. Even the techniques used for model calibration are complex and for the most part are beyond the scope of this report. We will only highlight some of the more fundamental calibration and validation checks that can be performed by the analyst. Special attention will be given to the validation of the traffic assignment step.

Calibration and Validation

Once the model parameters have been estimated the process of calibration and validation begins. Model calibration adjusts parameter values until the predicted travel matches the observed travel within the region for the base year. For purposes of forecasting it is assumed these parameters will remain constant over time. Calibration is conducted in all four steps of the modeling process and normally occurs after establishing model parameters.(2)

Model validation tests the ability of the model to predict future behavior. Validation requires comparing the model predictions with information other than that used in estimating the model. Validation is typically an iterative process linked to calibration.
If the analyst finds that the model output and the independent data are in acceptable agreement, the model can be considered validated. (2)

There are normally two types of validation checks - reasonableness checks and sensitivity checks. Reasonableness checks are tests that include the comparison of rates, checking of the total regional values and logic tests, etc. The analyst evaluates the models in terms of acceptable levels of error, ability to perform according to theoretical and logical expectations, and consistency of model results with the assumptions used in generating the results.

Sensitivity checks are tests that check the responses to transportation system, socioeconomic, or political changes. Sensitivity often is expressed as the elasticity of a variable. For example, one might examine this impact on travel demand if parking or toll fees are doubled.

In practical application the meanings of calibration and validation have changed over the years. This is especially true in areas where funding limitations preclude the conduct of travel surveys. In many of these locations the processes of calibration and validation have become a single exercise. Without current Origin Destination (O-D) studies, many of the models are calibrated by using default values derived from other studies and transferred to the local environment. In this process calibration and validation are merged since an independent database is not available for calibration. In order to develop a good match between model link volumes and ground counts, model parameters are modified to provide "reasonable" agreements. This is clearly the situation in many Tennessee MPOs where there is high if not complete reliance on transferable parameters for many planning variables.

The ultimate test of a travel demand model set is its ability to accurately predict traffic volumes on the transportation system. Therefore, in many areas traffic counts are the primary data parameter used for model validation. As will be discussed in this report, a number of checks are used to compare the model's simulated link values with the traffic counts. In any accuracy check it must be recognized there are errors associated with the ground counts. These errors are due to equipment malfunction, the inappropriate use of daily and seasonal factors to estimate Annual Average Daily Totals (AADT), and the absence of good classification data to correct axle counts to vehicles. (2) Likewise, while validation involves the results from the assignment phase of the four-step travel demand process, the errors can be attributed to all phases in the modeling process. Errors from previous steps may either compensate or be additive.

Various accuracy checks have been established as part of validation. The guidelines proposed by FHWA indicate that the model needs to be accurate enough as to not affect the number of lanes required. (2) However, given the scrutiny that may accompany conformity analyses, MPOs are encouraged to strive for a higher degree of accuracy when resources permit.

The remainder of this report will discuss some of the basic steps that can be taken to calibrate and validate the individual steps of the travel demand model set. Given the complexity of mode split models and their limited application in Tennessee no discussion of their calibration is provided in this report. However, an appendix has been provided that describes reasonableness checks for each step in the modeling process that the analyst can utilize as part of the calibration/validation process. (1)

Trip Generation Calibration/Validation

Trip generation calibration/validation involves three components: evaluation of the base year zonal socioeconomic estimates, development and application of a trip production (P) model, and the development and application of trip attraction (A) model. In a trip generation step, the trip production rate or model is applied using zonal socioeconomic estimates to produce trip productions. Similarly, the trip attraction rate or model is applied using zonal socioeconomic estimates to produce an estimate of zonal trip attractions.

The analyst is concerned with the reasonableness of the base year zonal socioeconomic estimates, the appropriate functions or mathematical expressions used for the trip production model and trip attraction model, and the parameters associated with these models. The analyst is also concerned with the trip P&A balance.

The census transportation planning package (CTPP) provides a breakdown of households by TAZ, household size, automobile ownership, and income group. These household and automobile ownership data and the income data can be used to verify the reasonableness of the base year zonal estimates. A comparison of observed (CTPP data) and estimated (base year zonal estimates) households, by size subgroup, will help identify serious bias in the base year zonal estimates.

Household travel surveys are used to estimate trip rates. Surveys do not provide direct estimates of zonal trip ends. Therefore, there is not an observed number of trip ends to compare with model-estimated trip ends. The analyst can compare summary statistics such as vehicle trips per household, person trips per household by trip purpose, and the percentage of person trips by trip purpose for the study area with values for similar study areas.

Typical values are provided below:

- Person trips per household: 8.5 to 10.5
- HBW person trips per household: 1.7 to 2.3
- HBO person trips per household: 3.5 to 4.8
- NHB person trips per household: 1.7 to 2.9
- HBW trips 18% to 27% of all trips
- HBO trips: 47% to 54% of all trips
- NHB trips: 22% to 31% of all trips

Trip Distribution Calibration/Validation

Calibration of the trip distribution model is often performed using two methods:

- Comparison of regional trip length frequency distribution (TLFD) and mean trip lengths by trip purpose, and
- Evaluation of the area-to-area flows of trips

A TLFD is a plot of the percent of total trips that occur for each separation in miles or in minutes as shown in Figure 1, which will be discussed in more detail later in this section.

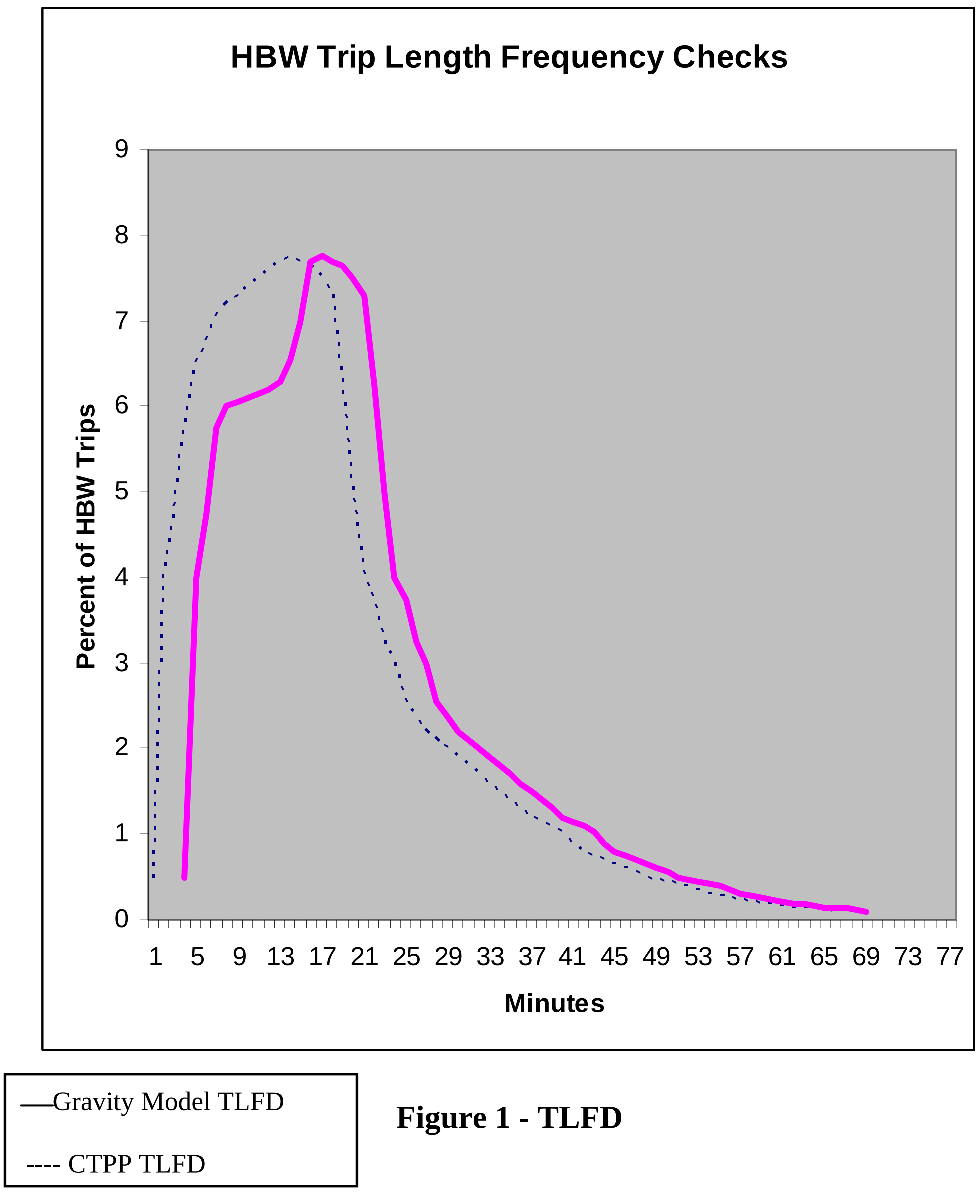


Figure 1 - TLFD

Household travel survey data are used to estimate TLFD and mean trip lengths in miles and minutes for each trip purpose. A sample as small as 500 households will provide a stable estimate for the TLFDs and mean trip lengths for the three internal trip purposes. These observed TLFDs and mean trip lengths from the travel survey are compared with the TLFDs and mean trip lengths from the application of the gravity model.

The Census Journey to Work (JTW) can be used to develop an observed TLFD for the HBW trip purpose. Note that there are differences in the definition of trips between travel surveys and JTW data that an analyst must consider.

Observed patterns of TAZ-to-TAZ flows of trips cannot be derived from small-scale household travel surveys. Commute trip area (areas or sectors are groups of TAZs) flows from the Census JTW tables, based on one in eight urban households sample, can be used to obtain reliable estimates of small area-to-area flows of trips. A comparison of area-to-area flows for the HBW trip purpose gravity model output with the JTW output can provide a reasonableness check on the performance of the gravity model.

Calibration of the gravity model continues by adjusting the friction factor values until a satisfactory agreement is achieved between the modeled and observed mean trip lengths and TLFDs for each trip purpose.

Figure 1 shows a typical comparison between the TLFDs for HBW trips, as estimated by the gravity model with observed trip TLFDs as reported in the Census JTW data. The analyst will consider differences in the definitions of work trip and trip chaining conventions in evaluating the differences between the two curves. Trip chaining refers to how intermediate stops made as part of a home-to-work trip are considered in the analysis of travel survey data. For example, a stop to purchase coffee on the way to work may be considered as one trip or two trips.

Figure 1 shows separations in minutes on the X axis. Similar graphs are prepared comparing the TLFDs for all internal trip purposes (HBW, HBO, and NHB) as estimated by the gravity model with observed TLFDs estimated from the household travel survey.

To calculate a TLFD, the number of trips made at one minute of travel time, two minutes of travel time, three minutes of travel time, etc., are counted. The number of trips at each minute of separation is divided by the total number of trips, and the result is multiplied by 100. The resulting percentages are plotted with the minutes of separation on the X-axis and the percent of trips on the Y axis.

Trip Assignment Calibration and Validation

Calibration of the trip assignment steps builds upon the successful calibration of each of the previous steps. If significant errors persist in these earlier steps, calibration of trip assignment will be meaningless. The analyst should proceed to a comparison of assigned volumes with counted volumes only after being satisfies that:

- The base year TAZ household characteristics and employment estimates are reasonable;
- The trip production and trip attraction estimates by trip attraction estimates by trip purpose produced;
- The gravity model is distributing the trips appropriately and the resulting mean trip lengths and TLFDs by trip purpose are agreeing with observed values and;
- The mode split model is providing a reasonable share of trips to each mode by trip purpose

Trips assignment validation involves comparing model generated link volumes compared to traffic counts. These comparisons can be made at various levels:

1. Geographic Areas of the Study Area
2. Functional Classification of Roadway
3. Screenlines
4. Link Specific Comparison

Statistics used to make the comparisons include:

1. Absolute Difference in Volumes
2. Percent Difference in Volumes
3. Average Error
4. Average Percent Error
5. Standard Deviation
6. R Square
7. Root Mean Square Error
8. Correlation Coefficient

Additionally, for specialized analyses such as emissions estimation comparisons the quantity of VMT estimated can be made. Here model estimated link volumes multiplied by distance are compared to link estimates of VMT produced from the HMPS database.

For both highways and transit, the analyst must recognize the day-to-day inherent variability in traffic volumes and transit ridership. The analyst must be comfortable with the reliability in traffic volumes and transit ridership used for travel model validation. As a part of this, the analyst must verify how the highway counts were derived and whether the counts are AADT, annual average weekday traffic (AAWT), axle counts divided by two, or some other measure.

This guidelines proposal for Tennessee will address specific statistical measures to be applied at specific levels of analysis. Table 1 provides an overview of where the appropriate tests will be defined in this document.

Guidelines

Guidelines provided by various agencies and the achieved level of accuracy are presented in this document. The first item to consider is the number of observations to be included in the verification, it is suggested that:

"The regional agency should strive to obtain traffic counts on 10 percent or more of the regionwide highway segments being analyzed, if resources allow. This 10 percent goal also applies to the distribution of counts in each functional classification (freeways and principal arterials, (at a minimum). Validation for groups of links in a screenline should include all highway segments crossing the screenline."(4)

1.0 Percent Differences in Volume

Percent error is more widely used than absolute numerical values because it better reflects the volume of the roadway. Typically, errors of 10% or less are established for screenlines and cutlines. Maximum desirable deviations in total screenline volumes have been provided by FHWA (see Figure 2). (4)(9)

FHWA also suggested all simulated links volumes should be within 5-10% of observed traffic volumes on the highway networks. Further it is suggested counts be obtained for 65% of the freeway and primary links in the networks. (4)

2.0 Percent Difference in Volume – By Functional Classification

Regionwide Validation Criteria Suggested Values have been proposed by both FHWA and the Michigan DOT by Functional Classification as Noted in Table 2:

Table 1:Reference to Statistical Test Proposed in Guidelines

Level of Analysis***

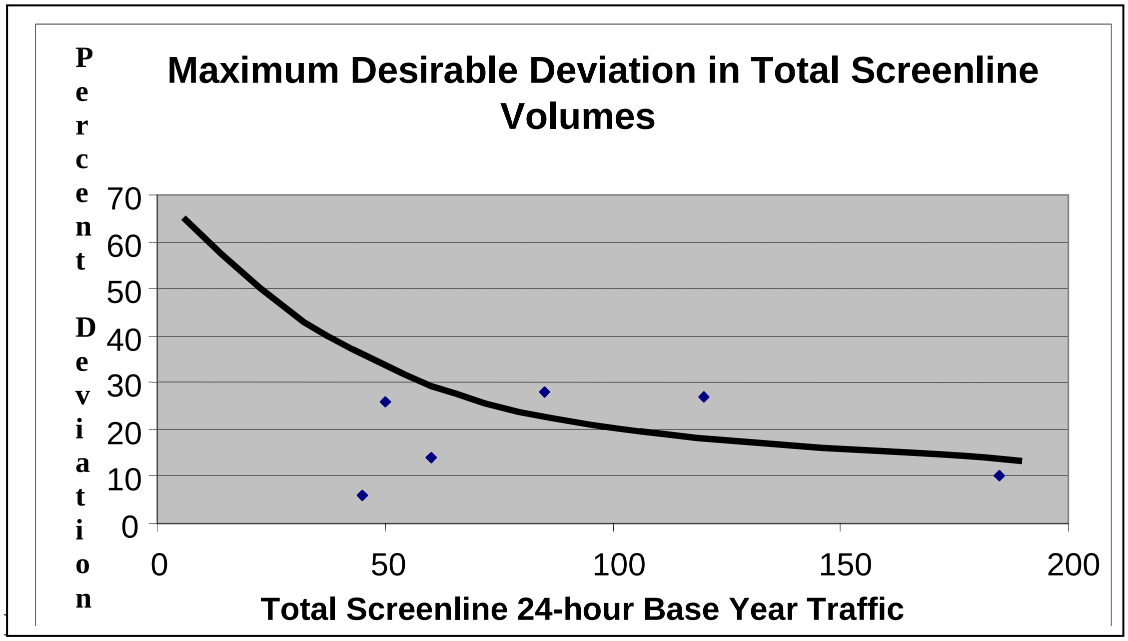
Statistical Test	Screenline/ Cutline Volumes	Functional Classification Volumes	Study Area Link Volumes	Links Volumes	Links Peak Hour
Percent Difference In Volume	1.0*	2.0	2.0	1.0 2.0	--
Correlation Coefficient	--	--	--	3.0	--
Root Mean Square	--	--	--	4.0	5.0**
Regionwide VMT	--	6.0	6.0	--	--

*Report Section Number where guidelines are discussed
**Optional
***The comparison may be made on vehicle-miles of travel rather than link volumes

Table 2: Suggested Region wide Validation Criteria

Functional Classification	Percent FHWA Error	Michigan DOT Percent Error
Freeways	+/- 7 percent	+/- 6%
Principal Arterials	+/- 10 percent	+/- 7%
Minor Arterials	+/- 15 percent	+/- 10%
Collectors	+/- 25 percent	+/- 20%
Frontage Road	+/- 25 percent	-----

Source (3) (5)



"This test will provide insight into whether the assignment model is loading trips onto the functionally classified systems in a reasonable manner. The percent error by functional classification is the total assigned traffic volumes divided by the total counted traffic volumes (ground counts) for all links that have counted volumes, disaggregated by functional classification."(4)

A percent deviation can be expressed as a function of link volumes. A desirable error range for link volumes has been presented by FHWA. (3)(8) In table form they are provided in Table 3. For comparison purposes, values provided by the Michigan DOT are also included in Table 3.

3.0 Correlation Coefficient:

The correlation coefficient estimates the correlation between the actual ground counts and the estimated traffic volumes and is produced by most software packages. A suggested regionwide correlation coefficient of more than 0.88 has been proposed. (3)

A sample scattergram is provided below (Figure 3). Further, it is suggested any links that lie outside of a reasonability boundary should be reviewed.

4.0 Percent Root Mean Square Error (RMSE)

The computation of RMSE is as follows (6):

$$RMSE = \sqrt{\frac{\sum_j (Model_j - Count_j)^2}{(NumberofCo\ unts - 1)}}$$

The Montana Department of Transportation (MDT) suggests that an appropriate aggregate Percent RMSE is less than 30%. An analysis by Oregon DOT offered Percent RMSE by link volume and is presented in Table 4.

Table 3: Percent Difference Targets for Daily Volumes for Individual Links

Average Annual Daily Traffic	**Desirable Percent Deviation**	
	FHWA	MDOT
< 1,000	200	60
1,000 – 2,500	100	47
2,500 – 5,000	50	36
5,000 – 10,000	25	29
10,000 – 25,000	20	25
25,000 – 50,000	15	22
> 50,000	10	21

Source: (3) (4)

Table 4: Percent RMSE by Link Volume

Link Volume	% RMSE
0 to 4999	115.757
5000 to 9999	43.141
10000 to 19999	28.272
20000 to 39999	25.383
40000 to 59999	30.252
60000 to 89999	19.199

Source (8)

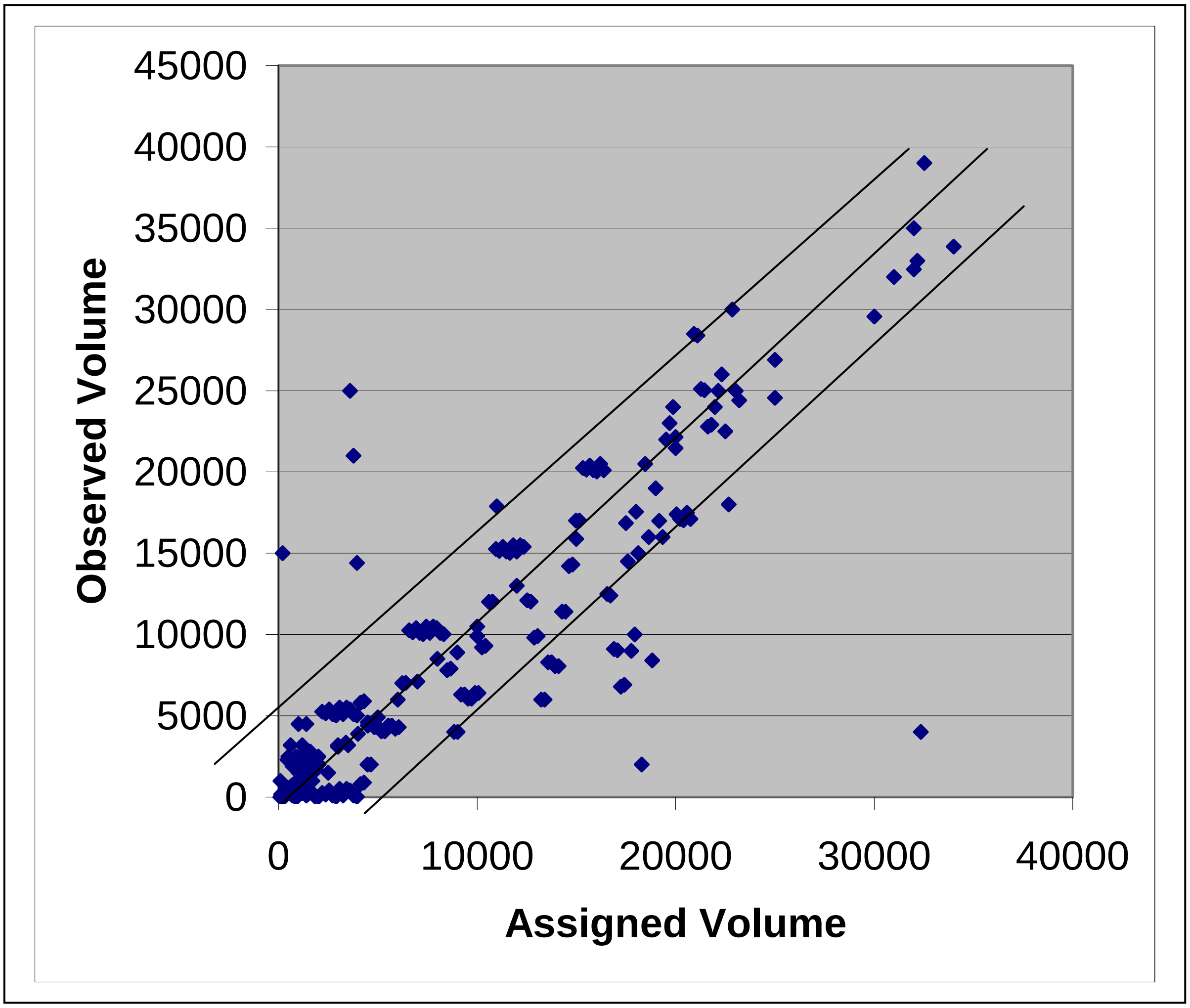


Figure 3- Scatter Plot of Assigned vs. Observed LinkTraffic Volume

5.0 Peak Hour Validation Targets:

Some agencies have focused on a model's peak hour volume estimate. Some typical validation criteria are presented in Table 5 for selected metropolitan areas where data was available.

The total AM and PM peak-hour volumes crossing each cordon or screenline should be within 10 percent of the AM and PM peak-hour counts, respectively. On an individual link basis, the following tolerances should be applied (8):

- 75% of freeway link volumes within +/- 20%
- 50% of freeway link volumes within +/- 10%
- 75% of major arterial link volumes with 10,000 vehicles per day within +/-30%
- 50% of major arterial link volumes with 10,000 vehicles per day within +/-15%

6.0 Regionwide VMT

It is suggested the assigned VMT should agree with the estimate VMT within 5%. VMT can also be compared by functional classification as noted in Table 6.

Highway Performance Monitoring System. If a (HPMS) data base is being used as part of the calibration process, sections 1.0 and 2.0 may be applied with a vehicle-mile comparison. In areas that are designated by the Environmental Protection Agency to be non-attainment for moderate and serious carbon monoxide, there is specific published guidance for forecasting and monitoring VMT.

Table 5: Percent Root Mean Square Error Comparisons

	Reno		**Phoenix**	**Concord**	
Facility	PM	ADT	ADT	PM	AM
Freeway	18.3	18.6	25.4	NA	NA
Arterial	39.2	36.8	38.5	NA	NA
Collector	76.1	77.5	62.7	NA	NA
Total	39.9	36.8	40.6	31.1	36.8

Source (6)

Table 6: Urban Area VMT by Facility Type

Facility Type	Urban Area Population		
	Small (50-200K)	Medium (200K-1M)	Large (> 1M)
Freeways/Expressways	18-23%	33-38%	40%
Principal Arterials	37-43%	27-33%	27%
Minor Arterials	25-28%	18-22%	18-22%
Collectors	12-15%	8-12%	8-12%

Source (6)

Recommendations:

For application in Tennessee, it is suggested the following tests be conducted:

1. Percent difference in value for screenlines and link volumes
2. Percent difference in volume by classification
3. Correlate coefficient by link volumes
4. Root mean square for link volumes

Where FHWA and Michigan criteria are provided it is suggested the FHWA criteria be selected.

References

1. FHWA, NHI, Introduction to Urban Travel Demand Forecasting, Participant Workbook. NHI course number 152054, Publication # FHWA-NHI-02-040, March 2002.

2. N. Pedersen and D. Sandahl. Highway Traffic Data for Urbanized Area Project Planning and Design. NCHRP Report 255. Transportation Research Board. Washington, DC 1982.

3. Model Validation and Reasonableness Checking Manual 7.0 Assignment Procedures Travel Model Improvement Program. U.S. Department of Transportation. http://tmip.fhwa.dot.gov

4. J.H.K. and Associates. Travel Forecasting Guidelines for California Department of Transportation. Emeryville, California. November 1992.

5. L. Tekiele. Michigan Department of Transportation Urban Model Calibration Target. Michigan Department of Transportation. Lansing, Michigan. June 1993.

6. P. Baso et al. Travel Forecasting Guidelines. Texas Transportation Institute. Texas A&M University System. College Station, Texas. October, 1994.

7. Ismart, D. Calibration and Adjustment of System Planning Models. Federal Highway Administration, Washington, DC, 1990.

8. Harvey, G. et al. A Manual of Regional Transportation Modeling Practice for Air Quality Analysis for the Natural Association of Regional Councils, Washington, D.C. July 1993.

9. Travel Demand Model Development and Application Guidelines. Parsons Brinckerhoff Quede and Douglas for the Oregon Department of Transportation. Portland, Oregon, June 1995.

APPENDIX

REASONABLENESS CHECKS
Source (1)

Socioeconomic Data Reasonableness Checks

- Review sources of TAZ base year estimates and forecast year projections.
- Graph 1950 to 2000 county or metro population data from census: add forecast population used for travel model and evaluate trend for reasonableness.
- Graph 1950 to 2000 average household size from Census: add forecast household size used for travel model and evaluate trend for reasonableness.
- Graph 1989 to 1999 regional median household income from Census after converting to constant dollars for the same year. Add forecast median household income used for travel model to the graph and evaluate trend reasonableness. The Census household income is for the year preceding the Census. The 2000 Census shows the household income for 1999.
- Determine the constant dollar year that was used to develop the trip production rates, and verify that the forecast median household income is for that same constant dollar year.
- If average automobile availability per household was used in place of household income for trip production, or if automobile availability was used for the mode split model, plot average automobile availability against time and evaluate trend for reasonableness.
- Calculate the distribution of employment by basic, retail, and service employment, and compare the base and forecast year distributions. Do the changes in the distribution of employees by type of employment agree with what is known about the changing character of the urban area? Nationally, service employment is increasing relative to basic and retail employment.
- Plot average employees per household per capita for the region against time; evaluate for reasonableness, Nationally, employees per household have been increasing over time, but the rate of increase is decreasing.
- Plot average employees for the region against time checks reasonableness.
- Compare base year and forecast year number of households by zone or by sector (groups of zones), and evaluate for reasonableness. Do the zones or sectors with growth agree with your expectations for the urban area?
- Compare base year and forecast year employment by zone or sector and evaluate for reasonableness. Do the zones or sectors agree with your expectations for the urban area?
- Ask for or plot the area types for the base year and for each analysis year. The area type is a measure of urban density. Have the area type boundaries been revised to reflect increasing density as the urban area grows? Do the changes in area types appear reasonable?

Travel Survey Data Reasonableness Checks

- Determine the source of the travel survey information used for model calibration. Were travel surveys conducted? If yes, in what year? What kinds of travel surveys were conducted (household, workplace, external station, truck, special generator, etc.)?

- If surveys were not conducted, what was the source of the trip rates, mean trip lengths, and trip length frequency distributions used in the model calibration? If the survey data were "borrowed" from another urban area, is that urban area similar in terms of geographical area, population and household characteristics, employment characteristics, urban density, and transportation system characteristics?

- What definition was used for trip rates by trip chaining? These definitions are important as they directly affect the trip rates by purpose. It will be difficult to compare trip rates by trip purpose with those from other urban areas if the trip chaining definitions are different.

- How do the trip rates by trip purpose compare with those for other urban areas with similar characteristics? Were person or vehicle trip rates used? How do the rates compare with earlier rates for the same urban area?

- How do the mean trip lengths by trip purpose compare with those from other urban areas with similar characteristics? Are the HBW mean trip lengths the longest and the HBO mean trip lengths the shortest?

- How do the TLFDs by trip purpose compare with other urban areas with similar characteristics?

Network Data Reasonableness Checks

- Examine the plot of a tree from a major activity center, such as the regional airport, the CBD, or a regional shopping mall. A tree is a minimum time path from one zone to all other zones. Evaluate the routes for reasonableness. Are these the routes that a knowledgeable traveler would probably take? Repeat for two or three major attractions in the region.

- Examine the definitions of the facility types used to code the networks. What is the level of detail?

- Examine the speed and capacity look-up table. Was LOS C, D, or E used to code the capabilities? The link volume that must be achieved before the trip assignment algorithm starts decreasing the speed on the link is a function of the coded link capacity. And, the interpretation of the link volume to capacity ratio in the trip assignment output is directly dependent on the LOS capacity used for coding the network. The numerical values selected for the speed and capacity used for coding the network. The numerical values selected for the speed and capacity tables are part of the model calibration process. Once calibrated, these values do not change between the base and forecast year networks.

- The travel mode should include a narrative or table description of all significant transportation projects added to the network for each analysis year. Examine this narrative or table description, and evaluate if these are still viable projects.

- Determine the major additions to the forecast networks in each analysis year. Are these transportation system improvements consistent with the most recent MTP? Are the project scopes reflected in the network (location and number of additional lanes, location of LRT, location of HOV, etc.) consistent with the MTP?

- Most travel demand software packages have plotting capability tied to a GIS platform. The analyst can then prepare network plots with network characteristics coded in different. Plots of facility types, number of lanes, freeflow speeds, and area types can be used to detect coding errors in facility type codes, number of lanes, freeflow speeds, and link area type. The software can also be used to compare the coded link area types. The software can also be used to compare the coded link lengths with lengths computed from the digitized node coordinates.

Trip Generation Reasonableness Check

- Examine the form of the trip production rate models. Is a cross classification or regression model used? Are the production rate models sensitive to household size, household income, and household automobile availability?

- Examine the form of the trip attraction rate models. Is a cross classification or regression model used? Are the attraction rate models sensitive to the total employment and retail employment? Are the models sensitive to basic, service, and retail employment and the number of households?

- What trips purposes were used? Were HBNW trips divided into additional trip purposes, such as retail or school trips?

- How were external-through and external-local trips accommodated in the travel model?

- Were the trip P&A rates for vehicle trips or for person trips?

- What was the P&A balance by trip purpose before scaling? Scaling factors of 0.9 to 1.1 are considered good. The scaling factor for the work trip purpose should be better than 0.9 to 1.1 and the scaling factor for NHB trips may be worse.

- Were special generators used in the travel model? If yes, what was the basis for the special generator trip rates? Were additional special generators of the same type added to the future year analysis? For example, if regional shopping malls were treated as a special generator were new shopping malls added to the forecast year? If yes, in what TAZ?

Trip Distribution Reasonableness Checks

- Examine the mean trip length for each trip purpose for the base year and for each forecast year. Is the mean trip length for HBW trips stable or increasing? Why? Is the change reasonable? The mean trip length for HBO trips is generally stable. Is it? If not, why not?

- Examine the base and forecast year plots of the TLFDs by trip purpose. A TLFD plot shows the percentage of trips occurring for each one minute separation in trip time. Do the plots look reasonable? Are there any significant changes in the shape of these distributions between the base and forecast years? If yes, why?

- Were the friction factors used in the gravity model kept constant between the base and forecast years? This is general practice. If the friction factors were changed, what was the basis for the change?

- Were terminal times used for the trip distribution step? What logic was used to calculate terminal times? Was the same logic used for the forecast years?

- Were socioeconomic adjustments factors (k-factors) used in the base year calibration? If yes, what was the basis for calculating the k-factors? What distribution model deficiency did the k-factors correct?

- How does the HBW mean trip length resulting from the trip distribution step compare with the 1990 and 2000 Census journey-to-work (JTW) mean trip length?

- How do the sector interchange volumes or percentages resulting from the trip distribution step compare with the JTW sector interchange volumes or percentages? It is practical to conduct this comparison only at a sector level and not at the zone level.

Mode Split Reasonableness Checks

- Were automobile occupancy factors by trip purpose used to convert person trips to vehicle trips? If yes, what was the basis for these factors? Were the automobile occupancy factors kept constant between the base year and the forecast years?

- Was a mode split model used? If yes, what form of the logit model was used? What modes were included in each nest of the model?

- What variables were included in the utility function for each trip purpose?

- Were the coefficients of the variables used in the utility functions logical?

- Does the mode split model documentation include a discussion of the independence from irrelevant alternatives (IIA) properties of the model?

- What was the value of time assumption used in the utility equation?

- What was the cost of parking assumption used in the utility equations?

- If a new mode was added to the model in a forecast year, how was the mode specific constant for that mode estimated?

- What network model coding logic was used to code access to transit in the transit network?

- How do the shares by mode compare between the base and forecast year?

- How do mode shares compare with the shares in other cities with a similar urban form?

Trip Assignment Reasonableness Checks

- Examine the all-or-nothing assignment. These are the volumes that would result if there were enough capacity for everyone to travel on their minimum time path. A comparison of the all-or-nothing volume with the equilibrium assignment volume for a given link provides an indication of the amount of traffic that might divert to that link if additional capacity were provided.

- Examine the equilibrium assignment. This assignment is the travel model's estimate of the amount of traffic forecasted to occur on a particular network, given the zonal household and employment forecast (the demand of transportation) and the coded transportation network (the supply of transportation).

- What volume delay function was used? Was the BPR function? If not, what function was used? At what volume to capacity ratio does diversion begin to occur with the function used?

- For model validation, how did the assigned screen line volumes compare with the counted screen line volumes?

- For model validation, how did the assigned cut line volumes compare with the counted cut line volumes?

- Was a time-of-day assignment performed? If yes, what were the time periods? What was the source of directional split factors by time-of-day? What was the source of the time-of-day factors? Were time-of-day factors kept constant between the base and the forecast year or were peak spreading factors used? If peak spreading factors were used, what was the source of the factors?

- If time-of-day modeling was not performed, how were the 24-hour volumes converted to hourly volumes for the traffic assignment step? What conversion factor(s) was used? Or, what was the relationship between peak hour volume and 24-hour volume?

- Intrazonal and centroid connector trips can be used to estimate local travel not assigned to the transportation network. What fraction of total VMT was assigned to intrazonal and centroid connector trips? Is this value reasonable?

Travel Model Performance Reasonableness Checks

Several simple checks can be performed to evaluate the overall reasonableness of an urban travel demand model. Travel model traffic assignment software provides summary statistics that are helpful for this application. Summary statistics are provided on VMT and vehicle hours of travel (VHT), and average speed for the entire system and by facility types.

The following table provides example values for selected U.S. cities.

Table 7: VMT, VHT and Average Speed Statistics

City	System VMT (X1000)	System VMT per Capita	Freeway VMT (x1000)	Principal Arterial VMT (x1000)	Average Peak Period Freeway Speed (mph)	Average Peak Period Principal Arterial Street Speed (mph)
Boston	58,285	19	21,800	16,110	45	26
Detroit	87,620	22	29,355	28,365	43	27
Louisville	23,800	28	9,475	3,995	53	30
New Orleans	17,900	16	5,470	5,250	47	29
Laredo	2,185	13	360	850	59	30
Omaha	11,480	21	2,995	4,135	53	28
Sacramento	27,300	22	10,470	8,335	47	28
Salem	3,145	17	1,060	1,345	52	31
Kansas City	39,130	29	17,310	5,730	52	28
Baltimore	43,245	29	20,775	8,915	46	28
Boulder	1,725	16	475	53	59	30
Tucson	12,460	19	1,775	5,090	51	27
Atlanta	95,110	37	38,650	14,575	43	27

Source: 1999 Texas Transportation Institute Annual Mobility Study

The travel model user will be particularly interested in how these summary statistics change between the travel model calibration or validation year to each of the forecasted years. A modest decrease in system and facility type speeds can be expected if the forecasted growth in demand exceeds the forecasted growth in the highway and transit system supply (center lane miles, lane miles, and route miles).

VMT per capita calculated by dividing the validation year VMT by the validation year population and the forecasted VMT by the forecasted year population provides an excellent reasonableness check. Typical values of VMT per capita are provided in the table 7. Significant changes in VMT per capita between the validation year and the forecasted year are not expected.