

## Executive Summary

Southeast Michigan Council of Governments (SEMCOG) has built a travel demand forecast model using TransCAD. The model calibration and validation require traffic count data input from numerous sources including the Michigan Department of Transportation (MDOT), county road commissions, and local communities. While there is a wealth of data, its wide variety and mixed formats pose a great challenge to create a dataset for periodic model calibration and validation. Another major issue is the reconciliation of the link definition difference between the TransCAD model network and the current traffic count locations. Therefore, an efficient process is needed to automatically retrieve, adjust, and format count data from available sources to meet the requirements of model calibration and validation. A statistically sound spatial analysis process is also necessary to select and associate traffic count data to corresponding TransCAD network links. Furthermore, it is necessary to identify the exact needs of additional traffic counts for future model validation and calibration.

To accomplish these objectives, this project includes three main parts:

1. Data Review and Model Calibration/Validation State-of-the-Practice Assessment
2. Data Specifications for Travel Demand Model Integration
3. Validation File Development

The key findings and recommendations from the project are as follows:

### **Traffic Count Data**

SEMCOG has been a model of best practice in regional traffic count data management. All major road agencies in the region, except MDOT, currently use the Traffic Count Database System (TCDS) software. The standardization allows easy transfer of traffic count data into the SEMCOG's Regional Traffic Count Database (RTCD), which is the region's central data repository. The database provides a wealth of data that can be used for model calibration and validation. The RTCD is also linked to a GIS map (Michigan Geographic Framework) for easy display of count locations and count data. This linkage can be used as a basis for matching to the TransCAD-based travel model developed for SEMCOG, because its model highway network is derived from the same geographic map.

Building upon its success in the regional traffic count data management, SEMCOG has a great potential in moving up to the next level with the focus on the following areas:

- **Data Detail**: Significant amount of the counts in the region have been collected using a 1-hour time interval. The interval should be reduced to 15-minute, as it will provide sufficient details for most planning and engineering applications. The existing counts also lack of vehicle classifications and speed data. The modern "raw data" counting technology should be considered for acquiring all data types in one setup.
- **Data Quality**: Even though there are some QA/QC routines built in the TCDS and RTCD, additional procedures need to be implemented to further improve the data quality.
- **Other Data Sources**: There are vast amount of traffic count data available that can be gathered to supplement the regular counting programs. The data are being collected in various planning and engineering projects, such as traffic signal timing optimizations.

SEMCOG should consider setting up a regional data portal that allows more agencies and consultants to upload the data to the central data depository.

### **Ideals for Validation Dataset**

A number of ideals should be considered for validation datasets created from the RTCD:

- **Coverage**: The validation dataset should provide adequate coverage in terms of number of links for the primary validation strata for network links. These strata include geographic strata such as county; area type; and functional class.
- **Randomness**: Randomness in link selection for the validation dataset would reduce the likelihood of inappropriate model adjustments.
- **Screenlines**: Screenline validation datasets should be separate from the “general” validation dataset. By nature, screenline validation data violate the randomness ideal.
- **Consistency with the modeling network**: Since the RTCD does not necessarily use the same link definition as the modeling network, care must be exercised in link selection to ensure that the counted location reflects the same location in the modeling network.
- **Time-of-day counts**: To the maximum extent possible, the validation dataset should include time-of-day counts in addition to the daily counts. The RTCD maintains counts on links for periods as small as 15 minutes and the model might use up to 10 time periods for assignment. The short time periods used for traffic assignment should not be considered for model validation, however, due to the normal variation in counts. At the most, validation to three or four time periods (which are aggregations of the periods used for traffic assignment) should be performed.
- **Classification counts**: To the maximum extent possible, the validation dataset should include vehicle classification counts in addition to the vehicle counts.
- **Consistency with forecasting period**: Links selected for the validation dataset should have counts that were collected on a weekday, not a weekend day. While factors can be developed to adjust weekend day counts to Average Annual Daily Traffic (AADT) or Average Weekday Daily Traffic (AWDT), weekend travel patterns and characteristics are so different from those on weekdays to make the selection of links with weekend counts a questionable practice for validation datasets. This ideal becomes even more important if time-of-day or classification counts are included on the validation dataset.

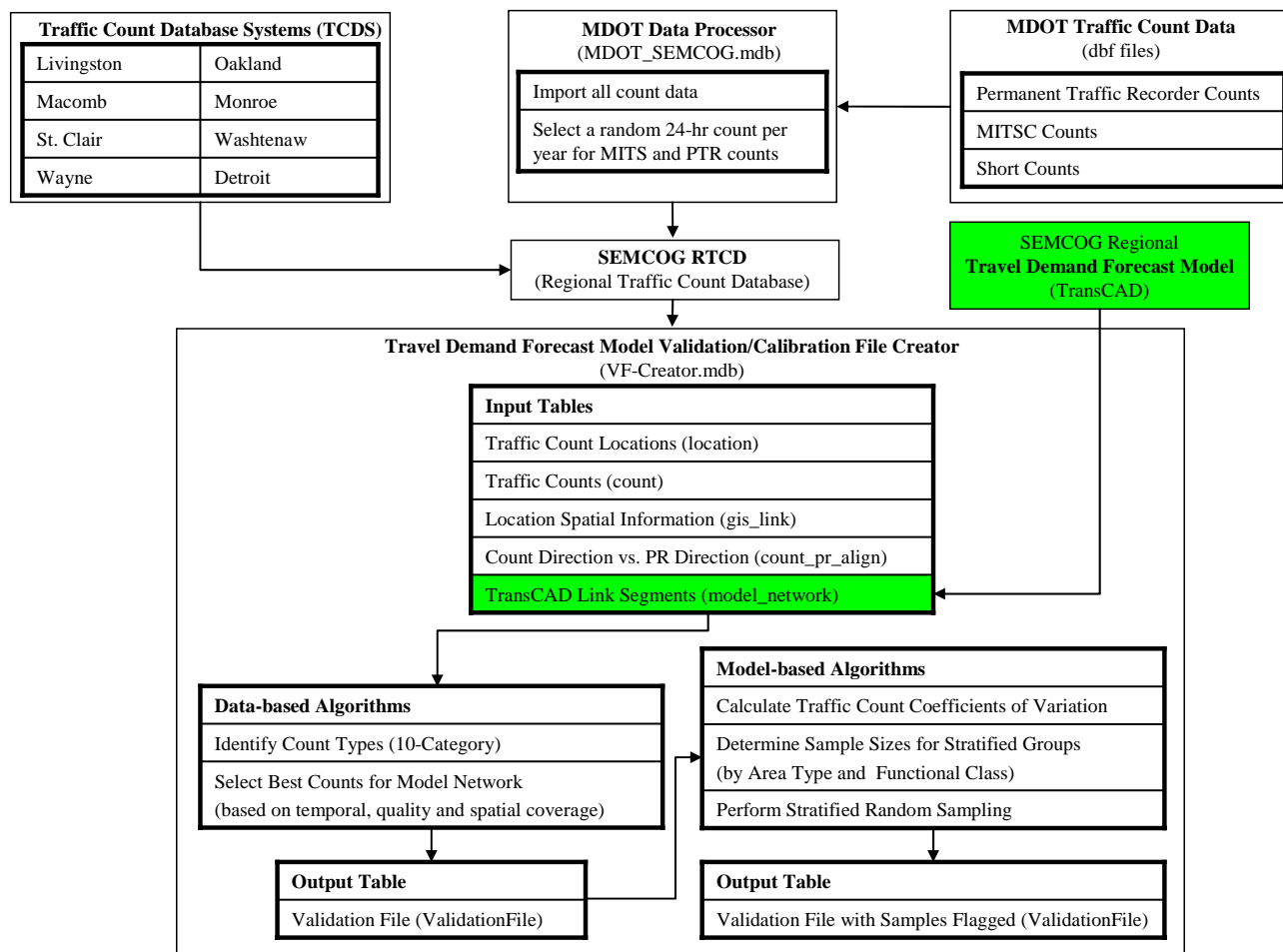
### **Sampling Methods**

The cost of collecting traffic count data can be substantial. Statistical sampling methods can be used to create a validation dataset that uses as few data points as possible. There are two desirable sampling plans:

- **Data-based Sampling Plan**: The plan would only sample the model links with existing count data. It will create a feasible dataset based on the available traffic count data in the RTCD. This plan can be used for past year validation. While this plan would not require additional traffic counts, it may produce a biased dataset as the region is noticeably lacking count data for certain geographic areas and functional classes. This plan is comparable to procedures traditionally used throughout the country.
- **Model-based Sampling Plan**: Based on the required confidence level and precision rate, the plan would sample links from the whole universe of the TransCAD model network. It is the preferred procedure for developing a validation file. However, the model-based sampling plan can identify TransCAD links without associated counts in the RTCD and, for this reason, is primarily suited for the development of validation datasets for anticipated future year validations. Once the sample has been specified, appropriate actions should be taken by SEMCOG to ensure that the selected links are included in the regional counting program.

**Validation File Development Process**

As a final product of this project, a suite of MS Access database programs has been developed to automatically process traffic count data from MDOT and RTCD and generates traffic count input file for TransCAD. The system is based on the data specifications, business rules, and sampling methodology developed throughout the project. The system flow is illustrated in the figure below.



For further information about the project, please contact the following:

Tom Bruff  
**SEMCOG**  
 535 Griswold Street, Suite 300  
 Detroit, MI 48226  
 Tel: (313) 324-3340  
 Fax: (313) 961-4869  
 Email: bruff@semcog.org  
<http://www.semco.org>

Ben Chen, PE, PTOE  
**Midwestern Consulting, LLC**  
 3815 Plaza Drive  
 Ann Arbor, MI 48108  
 Tel: (734) 995-0200  
 Fax: (734) 995-0599  
 Email: bc@midwesternconsulting.com  
<http://GIS.MidwesternConsulting.com>