Estimating Annual Average Daily Traffic (AADT) for Local Roads for Highway Safety Analysis

by

Tao Wang, Ph.D.
Research Associate
Lehman Center for Transportation Research
Florida International University
10555 West Flagler Street, EC 3680
Miami, FL 33174
Phone: (305) 510-3331
Email: twang001@fiu.edu

Albert Gan, Ph.D.*
* Corresponding Author
Professor
Department of Civil and Environmental Engineering
Florida International University
10555 West Flagler Street, EC 3680
Miami, FL 33174
Email: gana@fiu.edu

Priyanka Alluri, Ph.D.
Research Associate
Lehman Center for Transportation Research
Florida International University
10555 West Flagler Street, EC 3680
Miami, FL 33174
Email: palluri@fiu.edu

This paper contains 5,000 words, 7 figures, and 2 tables.
(7,250 equivalent words)

Submitted for:
Presentation and Publication
The 92nd Annual Meeting of the Transportation Research Board

November 15, 2012
ABSTRACT

Annual average daily traffic (AADT) is a required input to the newly released SafetyAnalyst software application. Further, AADT is also required to calculate crash rates. Traditionally, AADTs are estimated using a mix of permanent and temporary traffic counts collected in the field. Because field collection of traffic counts is expensive, it is usually performed for only the major roads. The mandate by the Federal Highway Administration (FHWA) to report the top 5% of high crash locations on all public roads underscores, for the first time, the need for state Departments of Transportation to acquire AADTs for also the non-state local roads. However, many local jurisdictions either do not have any AADT data, or they do not have them in sufficient quality. This paper presents a method to estimate AADTs for local roads using the travel demand modeling method. A major component of the method involves a parcel-level trip generation model that estimates the trips generated by each parcel. The generated trips are then distributed to existing traffic count sites using a parcel-level trip distribution gravity model. The all-or-nothing trip assignment method is then applied to assign the trips between the parcels and the traffic count sites onto local roadway network to yield estimates of AADTs. The estimated AADTs were compared with those from an existing regression-based method using actual traffic counts from Broward County, Florida. The results show that the proposed method produces significantly lower mean absolute percentage errors.

Key words: Annual Average Daily Traffic, Tax Parcel Records, Trip Generation, Travel Demand Modeling.
INTRODUCTION

Annual average daily traffic (AADT) is the average 24-hour traffic volume at a roadway location over an entire year. AADT is required for many transportation analyses including economic evaluation of highway safety projects, estimation of highway user revenues, computation of highway statistics such as vehicle miles traveled (VMT), development of improvement and maintenance programs, etc.

One well-known application of AADT is in the calculation of crash rates. As part of the new Highway Safety Improvement Program (HSIP) of the Federal Highway Administration (FHWA), states are required to submit an annual report describing no less than 5% of their highway locations on all public roads that exhibit the most severe safety needs (1). To submit the annual 5% reports (also known as Transparency Report), the Florida Department of Transportation (FDOT) needs to have AADTs for all roads in Florida; however, FDOT estimated AADTs for only its state roads but not the local roads.

In addition to the need for 5% reporting, AADT is also a required input to SafetyAnalyst (2), a new safety analysis software released by the American Association of State Highway and Transportation Officials (AASHTO). The software aims at providing state and local highway agencies with a comprehensive set of tools to enhance their programming of site-specific highway safety improvements. FDOT is interested in deploying SafetyAnalyst for all roads in Florida. Because AADT is a required input to SafetyAnalyst, FDOT was, again, faced with the problem of not having AADTs for local roads.

The most accurate method for obtaining the AADT of a roadway segment is to install an Automatic Traffic Recorder (ATR) to count the total volumes continuously throughout the entire year. However, because the installation and maintenance of permanent counters are expensive, the number of permanent counters is typically limited. Therefore, it is not economically feasible to apply this method of AADT estimation on a widespread basis.

An alternative approach to estimating AADT is to use portable counts, also called short-term, seasonal, or coverage counts. The collected short-term volumes on the interested roads are then used to calculate Average Daily Traffic (ADT) which is then converted to AADT by applying some adjustment factors. This factor approach is more economically feasible than the permanent count method, but is still too costly to cover the local roads, which number over two million segments in the state of Florida.

In 2007, FDOT contracted with the University of South Florida to develop regression models to estimate AADTs for the local roads (3). The project attempted to improve upon a set of regression models developed earlier by Xia et al. (4). However, a preliminary evaluation based on data from the Miami-Dade and Broward counties, Florida showed that the estimation errors of the models exceeded 100% and 200%, respectively.

This paper presents an improved method of estimating AADTs for local roads in Florida. The method applies travel demand modeling techniques at the tax parcel level. In this method, trips are generated based on parcel data using the trip generation rates from the trip generation report of the Institute of Transportation Engineers (ITE) (5). A gravity trip distribution model is then developed to distribute the parcel trips to the traffic counts sites on the major roads. All the trips are finally assigned to local roads by applying all-or-nothing assignments based on free-flow travel times.
LITERATURE REVIEW

Several existing AADT estimation methods have been reported in the literature. Among the methods, the regression analysis has been the most widely used. A non-exhaustive list of states that has developed regression models to estimate AADTs for their highway networks would include Kentucky (6), Alabama (7), Minnesota (8), and Indiana (9). In Florida, regression models were first developed to estimate AADTs for local roads in Broward County (4). The models were modified and improved in multiple follow-up studies (10, 11, 12). A more recent attempt to estimate AADTs for local roads in Florida was performed by Lu et al. (3). As aforementioned, they attempted to improve the regression models developed by Xia et al. (4). However, the estimation errors from their models were found to be quite high.

The availability of high-resolution satellite images and aerial photos had encouraged some researchers to explore the potential of estimating AADTs with image-based data. McCord et al. (13) first proposed an image-based method for AADT estimation using satellite images and aerial photos. Later, a related method combining image-based and ground-based estimations was implemented for Ohio road segments (14, 15). One problem with image-based methods is that it is difficult to estimate traffic volumes on local roads as traffic on such facilities is usually sparse and infrequent.

Another method attempted for AADT estimation was based on machine learning algorithms. For example, a multi-layered, feed-forward, and back-propagation neural network with supervised learning was developed to compare the Artificial Neural Network (ANN) approach for AADT estimation to the traditional factor approach with 48-hour short-term count data (16). The ANN approach was later applied by the same researchers to estimate AADTs on low-volume rural roads (17, 18). Another effort that involved ANN models for AADT estimation was conducted by Lam and Xu (19). A more recent study was performed by Li et al. (20) who combined a K-Nearest Neighbor (K-NN) algorithm with Geographic Information System (GIS) technology to aid in AADT estimation. In another recent study, a modified version of support vector regression machine (SVR) named SVR-DP (SVR with Data-dependent Parameters) was used to forecast AADT one year into the future based on historical AADTs (21).

Studies using the travel demand modeling approach for AADT estimation are relatively rare. One known study that used this approach was conducted by Zhong and Hanson (22). The study developed travel demand models to estimate AADTs on low-class roads for two regions in the province of New Brunswick in Canada (22). While the study showed that the method has the potential to improve AADT estimation, the results from the study were negatively impacted by the available census geographic unit data, called dissemination area (DA), which was the smallest census unit available in Canada. However, a DA by definition is a geographic unit composed of one or more neighboring dissemination blocks with a population of 400 to 700 persons. Accordingly, data based on DAs are clearly too coarse for applications involving local roads, which typically attract most trips generated from the abutting land uses. In this research, this travel demand modeling approach is applied based on data at the parcel level.

METHODOLOGY

A parcel-level travel demand analysis model was implemented in this study. Similar to the traditional travel demand forecasting model, the parcel-level model also involves a series of mathematical models to simulate human travel behavior. The model consists of the following
Wang, Gan, and Alluri

four steps: network modeling, parcel-level trip generation, parcel-level trip distribution, and parcel-level trip assignment.

Different from the traditional travel demand model, which attempts to simulate the choices that the travelers may make during the entire trip from the origination to the destination, the parcel-level model attempts to simulate choices that travelers may make in response to the given local street system. The model does not include the mode choice step as the transit trips and other modes are insignificant on local roads. The functionalities of each step involved in the parcel-level travel demand analysis model are as follows:

- **Network Modeling** defines the boundaries of the study area, prepares and preprocesses the roadway network, parcel, and traffic counts data, and sets up the network representation of the roadway linked with parcels and traffic count sites.

- **Parcel-level Trip Generation** estimates the number of vehicle trips generated by each parcel in the study area. The estimation is calculated based on the land use type of each parcel and its corresponding ITE trip generation rate.

- **Parcel-level Trip Distribution** determines where the trips generated by each parcel will go. It determines the number of trips between a parcel and a traffic count site based on traffic count data (or AADT estimated from the count data) and the shortest travel time between them.

- **Parcel-level Trip Assignment** predicts the routes the travelers will take to reach the traffic count sites on major roads, resulting in the estimated AADTs of local roads in the study area.

In the network modeling step, the study area boundary, commonly called the cordon line, is defined. Unabridged roadway network data, detailed parcel data, and traffic count sites data were used in the analysis. Similar to the traditional travel demand model, centroids and centroid connectors were created. Each parcel was assigned a centroid and a centroid connector to represent access to a neighboring road segment. An example of a subarea with the added centroid connectors connecting the parcels and the closest roads is shown in Figure 1. The green polygons represent the parcel boundaries; the blue lines represent the roadway; and the gray lines are the added centroid connectors.

Parcel-level trip generation is the process to estimate and quantify the number of trips each parcel will generate. In this application, this step was implemented using both the Department of Revenue (DOR) parcel data and the trip generation rates and regression equations provided by ITE Trip Generation Report (5).

The DOR parcel data has 100 land use types for the parcels, and the ITE Trip Generation Report provides trip generation rates and/or equations for 10 main land use categories and 162 sub-categories. The two types of land use categories were matched to estimate the parcel trips. In the case that a parcel land use type encompasses multiple ITE land use categories, the results were averaged and weighted by an estimate of the relative presence of each ITE land use category in the study area. The calculation can be expressed as follows:

\[
T = \sum_{i=1}^{n} F_i(X) \times P_i
\]

where,

\[
T = \text{number of vehicle trips generated by a parcel};
\]
\[ F_i = \text{ITE trip generation function (either regression equation or average trip rate) for ITE land use type } i; \]
\[ X = \text{independent variable such as dwelling units, gross floor area, etc.}; \]
\[ n = \text{number of ITE land use types}; \]
\[ P_i = \text{proportion of ITE land use type } i \text{ in the study area}. \]

FIGURE 1 Example of study area with parcels (green lines), centroid connectors (gray lines), and existing roads (blue lines).

For most of the land use types, the ITE trip generation rates/equations are based on dwelling units or areas, which are also the attributes of a parcel in the parcel database. Hence, for a majority of parcels, the trip generation can be calculated directly by using the parcel data. However, for land use types of which the corresponding independent variable used by ITE rates
is a size attribute that is different from the parcel data, they are adjusted by the ratio between the ITE and parcel based mean values using the following equation:

\[ T = F(X \times R) \]  

(2)

where,

\( T \) = number of vehicle trips generated by a parcel;
\( F \) = the ITE trip generation function (either regression equation or average trip rate, which can be regarded as a special linear regression);
\( X \) = independent variable such as dwelling units, gross floor area, etc.; and
\( R \) = the ratio of mean values between the ITE independent variable and the parcel size attribute.

Due to the lack of other demographic and land use data for a few uncommon land use types, the proportion of ITE land use type \( P_i \) in Equation (1) and the ratio of mean values between the ITE independent variable and the parcel size attribute \( R \) in Equation (2) were determined by human judgment.

Due to the different travel patterns among the weekday, Saturday, and Sunday, the number of trips for weekdays and weekend were averaged using the following equation:

\[ T_{\text{average}} = \frac{T_{\text{weekday}} \times 5 + T_{\text{Saturday}} + T_{\text{Sunday}}}{7} \]  

(3)

where,

\( T_{\text{average}} \) = the final estimated number of daily trips generated by a parcel,
\( T_{\text{weekday}} \) = average weekday trips generated by a parcel,
\( T_{\text{Saturday}} \) = Saturday trips generated by a parcel, and
\( T_{\text{Sunday}} \) = Sunday trips generated by a parcel.

Similar to the traditional travel demand model, parcel-level trip distribution is also derived from Newton's law of gravity. However, the trips are distributed between the parcels and the traffic count sites instead of among the zones. To distribute trips, all the parcels have only productions and no attractions, and all the traffic count sites have only attractions and no productions. Productions of a parcel are the trips generated in the parcel-level trip generation step, and the attractions of a traffic count site are either the traffic count data or AADT estimated from the traffic count data. In addition, while traditional travel demand model usually includes the effects of multiple travel impedance factors, such as travel time, cost, etc., parcel-level trip distribution considers only the shortest free-flow travel time. This is expedient because travel time is the major factor that determines trips on local roads, and travelers will most likely choose the fastest path to access major roads to reach their destinations. The parcel-level trip distribution can be expressed as follows:

\[ T_{ij} = \frac{A_j / D_{ij}}{\sum_{k=1}^{n} (A_k / D_{ik})} \times T_i \]  

(4)
where,

\[ T_{ij} = \text{daily vehicle trips between parcel } i \text{ and traffic count site } j, \]
\[ T_{i} = \text{total daily vehicle trips generated by parcel } i, \]
\[ A_{j} = \text{AADT estimated from traffic count volume at traffic count site } j, \]
\[ D_{ij} = \text{the shortest free flow travel time between parcel } i \text{ and traffic count site } j, \]
\[ n = \text{number of traffic count sites}. \]

It should be noted that, since the parcel-level trip distribution step distributes the trips from the parcels to the nearby traffic count sites, enough traffic count data are required to evenly cover the entire study area. Uneven coverage of traffic count sites may cause inaccurate distribution of trips.

For traffic assignment, the traditional travel demand model commonly assigns trips using an all-or-nothing with capacity restraint method, also known as the equilibrium assignment method. In this application, only the simple all-or-nothing assignment method is needed because congestion seldom happens on local roads. In other words, for the same reason that affects the trip distribution step, the travelers’ route selection is based mainly on the free-flow travel time to the nearby major roads. The simple all-or-nothing assignment also improved the efficiency of the model as it involves only one iteration.

After the trips for all of the parcels were assigned to the local roads, the AADT is estimated as the sum of trips in both the directions for a roadway segment. Finally, data post-processing was performed to calculate the final AADT values for all local roads.

MODEL DEVELOPMENT

Two development tools were applied to implement this model: ArcGIS 10.0 from ESRI and Cube 5.1.3 from Citilabs. An application of ArcGIS called ModelBuilder was used to perform the data preprocessing and post-processing. Scripts were then developed in Cube Voyager to call and customize the library programs to implement the model. Figure 2 shows the system components and procedure used to estimate AADT. The procedure includes the following sub-steps:

- ArcGIS is used to preprocess the input data for the model including the DOR parcel data, unabridged highway network data, and traffic count sites data.
- The preprocessed input data are imported into Cube, and the highway network is built from the unabridged roadway shape file.
- The built highway network is used by the network modeling step to calculate the free flow travel time skim matrix.
- The parcel-level trip generation step is performed by using the merged DOR parcel data and traffic count site data as well as the trip generation rates and regression equations provided by the ITE Trip Generation Report.
- A parcel-level trip distribution gravity model is used to distribute the generated trips between the parcels and the nearby traffic count sites.
- The distributed trips are assigned to the network using all-or-nothing assignment method in the parcel-level trip assignment step.
- The traffic volume data of the loaded network are exported, and ArcGIS is used again to calculate the final AADTs, which are then joined to the original roadway network to obtain the roadway network with AADTs.
ArcGIS ModelBuilder, which provides a visual programming environment allowing users to graphically link geoprocessing tools into models, was used to implement the ArcGIS component. While the models built with ModelBuilder can be executed directly in ArcGIS, they can also be exported to scripting language such as Python. The Python scripts can be called with the Cube Voyager Pilot program, so theoretically all the steps of the ArcGIS part can be integrated into Cube to simplify the execution of the entire model. However, because this part of the procedure requires some geo-processing tools that are supported only by ArcGIS 10.0, which is not compatible with the current version of Cube (5.1.3), integration of ArcGIS into Cube could not be realized. Nevertheless, this incompatibility would not affect the results of the entire model.

Figure 3 shows the model steps and the input and output files for each steps implemented in Cube. When the model is run, the four steps are executed in sequence, and the output files of a preceding step become the input files of a later step. It is noted that if there were no compatibility problems as mentioned above, the ArcGIS part of the procedure could have been combined with Cube, and the steps shown in Figure 3 would be all the steps involved in the entire model.

Table 1 summarizes the input and output files for each step. There are two input files for Cube: the network file preprocessed by ArcGIS, and the DBF file for the merged parcels and traffic count sites shape file which is also generated by ArcGIS. There is one output file generated by the Cube part, i.e., the DBF link file with the traffic volume information exported from the loaded network assigned in the parcel-level trip assignment step. Among the steps, the output files of a preceding step become the input files of a later step.
TABLE 1 Input and Output Files

<table>
<thead>
<tr>
<th>Model Step</th>
<th>Input File</th>
<th>Output File</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Modeling</td>
<td>Preprocessed Network File</td>
<td>Free Flow Time Skim Matrix File</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modified Network File</td>
</tr>
<tr>
<td>Parcel-level Trip Generation</td>
<td>Merged Parcels and Counts DBF File</td>
<td>Vehicle Trips DBF File</td>
</tr>
<tr>
<td>Parcel-level Trip Distribution</td>
<td>Free Flow Time Skim Matrix File</td>
<td>Distributed Trips Matrix File</td>
</tr>
<tr>
<td>Parcel-level Trip Assignment</td>
<td>Distributed Trips Matrix File</td>
<td>Link DBF File with Volume Exported from Loaded Network</td>
</tr>
</tbody>
</table>

MODEL EVALUATION

Because the actual AADT values for local roads were not available, the AADTs estimated from short-term traffic count data had to be used as the ground truths for model evaluation in this study. To quantify the estimation error, the mean absolute percentage error (MAPE), calculated as follows, was used:

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{F(i) - G(i)}{G(i)} \right|
\]

where,

\( G(i) \) = ground truth AADT at location \( i \),

\( F(i) \) = estimated AADT at location \( i \), and

\( n \) = total number of locations.
As aforementioned, regression modeling has been the most widely used AADT estimation method. Based on the method’s popularity and also the availability of the experimental results, the regression method proposed by Lu et al. (3) was compared with the proposed parcel-level travel demand model. Lu et al. developed multiple regression models that estimate AADTs as functions of socio-economic and roadway characteristics.

Broward County in Florida was selected as the study area for evaluation. The County includes over 4,000,000 parcels. However, Cube 5.1.3 can only process a maximum of 32,000 zones (or parcels in this application) at a time. A total of 10 subareas with a total of 78 evaluating count sites were selected to cover diverse areas in this evaluation. An example map of the subareas is shown in Figure 4. Two types of count sites are shown in this map. One is the estimation count sites on the major roads (indicated by circular dots in the map) which are used to estimate the AADTs, and the other is evaluation count sites on the minor roads (indicated by the triangles in the map) which are used to evaluate the estimation results.

FIGURE 4 An example of the subareas.

Figures 5 and 6 compare the overall performance of the regression method and the proposed method, respectively. As shown in the figures, the maximum AADT is lower than 30,000 vehicles/day and within a reasonable range, since all the testing locations were on local roads. From Figure 5, it can be seen that the regression models tend to overestimate AADT. The plotted dots in Figure 6 show that the results of the proposed method are distributed more evenly on both sides of the 45-degree sloping line than in Figure 5.
FIGURE 5 Comparison of estimated AADTs from regression models and ground truth AADTs.

FIGURE 6 Comparison of estimated AADTs from parcel-level demand model with ground truth AADTs.
Table 2 quantifies the accuracy of the two methods using the MAPE measure. It can be seen that the proposed method has consistently lower estimation error than the regression models for all study areas. The overall estimation errors for all the 78 evaluating count sites in the 10 study areas were also calculated. The results show that the proposed model generated 52% MAPE, which is 159% lower than the corresponding value of 211% from the regression models.

TABLE 2 Errors Comparison of Regression and Demand Models

<table>
<thead>
<tr>
<th>Subarea #</th>
<th>MAPE from Regression Models (%)</th>
<th>MAPE from Demand Model (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>217</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>314</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>177</td>
<td>48</td>
</tr>
<tr>
<td>4</td>
<td>345</td>
<td>66</td>
</tr>
<tr>
<td>5</td>
<td>1756</td>
<td>49</td>
</tr>
<tr>
<td>6</td>
<td>405</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>114</td>
<td>62</td>
</tr>
<tr>
<td>8</td>
<td>186</td>
<td>49</td>
</tr>
<tr>
<td>9</td>
<td>181</td>
<td>56</td>
</tr>
<tr>
<td>10</td>
<td>157</td>
<td>39</td>
</tr>
<tr>
<td>All Subareas</td>
<td>211</td>
<td>52</td>
</tr>
</tbody>
</table>

Depending on the availability of traffic count data, most of the traffic count sites used for this evaluation are located on local roads that are directly connected to the state roads. The lower-level local roads such as the community roads were not used in this evaluation because of the lack of traffic count data. However, the proposed method is expected to perform better even for lower-level local roads as its trip generation is based on detailed parcel level data. To verify this assumption, the AADTs estimated using the two methods for the available lower-level local roads were checked and compared. Figure 7 shows the estimation results for the roads in a community of approximately 160 houses. In this figure, the AADTs estimated from the proposed method and the regression method are displayed in red and green, respectively. It can be seen that the AADTs estimated by the regression method were unreasonably high. One reason that the regression method significantly overestimated the AADTs in this case was due to their inability to recognize that the layout of the local roads was designed such that through traffic, if any, was minimal.

CONCLUSIONS

Annual average daily traffic (AADT) data are needed to identify high crash locations on all public roads for 5% reporting as mandated by FHWA. They are also required for input to the newly released SafetyAnalyst software application. The lack of AADT data, especially for the vast local road networks, prevents such application from being deployed to improve safety on local roads. This paper described a parcel-level travel demand model that was developed to estimate AADTs for local roads. Compared to results from a set of existing regression models, the parcel-level demand model improved the accuracy of AADT estimation for local roads by 159% based on the MAPE measure. A major component of the proposed model involves a parcel-level trip generation model that estimates the trips generated by each parcel. The model makes use of the tax parcel data together with trip generation rates from the Institute of Transportation Engineers (ITE) Trip Generation Report. One advantage of using tax parcel data
is that the data is updated at least annually, making it possible to update the AADTs in response to lane use changes. One disadvantage of the proposed model is the large number of parcels that have to be preprocessed to improve the model’s efficiency. For applications involving large areas, the study area has to be divided into multiple subareas and run the model separately. In addition, the model also requires that there not only be sufficient traffic count data to cover the entire study area, but they should be relatively evenly spaced, as uneven coverage of traffic count sites may result in inaccurate distribution of trips.

FIGURE 7 Example of AADT estimation for local roads in a community.
(Note: AADTs estimated by the proposed method and the regression method are shown in red and green, respectively.)

ACKNOWLEDGEMENTS

This research was funded by the Research Center of the Florida Department of Transportation (FDOT). The authors are thankful to the Project Manager, Mr. Joseph Santos, of the FDOT Safety Office for his guidance and support.
REFERENCES


