PB99-164634

Estimation of Annual Average Daily Traffic for Off-System Roads in Florida

FINAL REPORT

Prepared for

Florida Department of Transportation

STATE OF FLORIDA
DEPARTMENT OF TRANSPORTATION

by

L. David Shen, Fang Zhao and Diana I. Ospina
Principal Investigators
Lehman Center for Transportation Research
Department of Civil & Environmental Engineering
Florida International University
Miami, FL 33199

July 28, 1999

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

We would like to thank Ms. Qing Xia, Mr. Adong Xu, and Mr. Caijun Luo, all former graduate research assistants at the Lehman Center for Transportation Research for their dedicated work on data processing and modeling.

Dr. Zhenmin Chen from the FIU Department of Statistics provided technical advice on statistics modeling. His assistance is acknowledged and appreciated.

Dr. Albet Gan reviewed this report and provided many constructive comments. His effort is greatly appreciated.

We would like to extend our special thanks to Mr. Albert Dominguez, the project manager from District 6, and Mr. Gordon Morgan, the project manager from the Transportation Statistics Office of the Florida Department of Transportation (FDOT), for providing various reports and information throughout this study and for their constructive comments and suggestions.

This research was supported by a grant from the FDOT. The execution of this project would not have been successful without the assistance and cooperation from the representatives of each district's planning office who handled off-system road data, district planning managers, modeling coordinators, traffic engineers, and statistics department administrators that participated in the surveys and provided significant information.

Many people from the different cities, counties, metropolitan planning organizations, public work offices and regional planning councils in the State of Florida helped us to collect the information needed for this research. They are too many to list here individually. We would like to extend our appreciation and acknowledgment in general to each of these persons for their time and patience to answer our telephone calls, respond to our surveys, and their effort to provide accurate information.

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the Department of Transportation (FDOT) or U.S. Department of Transportation.

This publication is prepared in cooperation with the State of Florida Department of Transportation and the U.S. Department of Transportation.

EXECUTIVE SUMMARY

Estimation of Annual Average Daily Traffic (AADT) is extremely important in traffic planning and operations for the state departments of transportation (DOTs), because AADT provides information for the planning of new road construction, determination of roadway geometry, congestion management, pavement design, safety considerations, etc. AADT is also used to estimate state wide vehicle miles traveled on all the roads and is used by local governments and the environmental protection agencies to determine compliance with the 1990 Clean Air Act Amendment. Additionally, AADT is reported annually by the Florida Department of Transportation (FDOT) to the Federal Highway Administration.

In the past, considerable efforts have been made in obtaining traffic counts to estimate AADT on state roads. However, traffic counts are often not available on off-system roads, and less attention has been paid to the estimation of AADT in the absence of counts. Current estimates rely on comparisons with roads that are subjectively considered to be similar. Such comparisons are inherently subject to large errors, and also may not be repeated often enough to remain current. Therefore, a better method is needed for estimating AADT for off-system roads in Florida.

This study investigates the possibility of establishing one or more models for estimating AADT for off-system roads in Florida. It is intended by the FDOT that once such models are developed, important variables that contribute to AADT may be identified and their impact on AADT may be quantified. As a result, a matrix or a classification method may be developed for grouping roads into different categories. Roads in each category would possess similar characteristics. AADT may then be estimated for roads that do not have traffic counts based on those on other roads in the same category.

The issues to be addressed in this research include data availability, data collection and processing, suitability of various data for model development, choice of modeling techniques, model accuracy, model applications, and model improvements. This document presents the results of the research efforts addressing the above-mentioned issues.

STATE-OF-THE-ART

Literature on estimating AADT for roads that do not have traffic counts is limited. Most of the relevant literature is on either estimating traffic volume for freeways or extrapolating 24-hour or 48-hour traffic count data to obtain AADT. One attempt to estimate AADT for off-system roads was made by a research group at Purdue University, Indiana, in which a multiple regression method utilizing data aggregated at the county level was employed. Several other studies have been reported on estimating AADT for state roads using statistical methods. The predictors used include county population size, total number of through lanes, state/non-state code indicating jurisdictions, location type (rural/urban), personal income level, vehicle registrations, etc. A common problem reported in a number of publications is the lack of data for some potentially important predictors. For example, in a Minnesota study, it was proposed to use the population within a certain distance of the roadway as a predictor, but the attempt was abandoned due to the unavailability of data.

DATA COLLECTION

Data availability is critical for both model development and for future model applications. Generally speaking, two groups of data are relevant to this research. They are AADT estimated from traffic counts and other non-traffic data that may be used to build models to estimate AADT.

The objective of the data collection effort in this research is to collect, document, and summarize all required data elements to support subsequent and future research efforts. The data collection effort has been carefully planned to obtain dependable, quality information from all agencies handling data on off-system roads. It includes a comprehensive literature review, telephone interviews with districts' personnel responsible for traffic data collection, and a survey.

The FDOT generally has good traffic data collection programs that meet most of the Department's needs, such as planning, design and operation assistance and administration functions. The Department conducts traffic data collection on the highway network of Florida to estimate volumes, types of vehicles, and weight of trucks. While the FDOT maintains traffic data and roadway inventory information on state roads, it only collects a small amount of data on off-system roads every three years for Highway Performance Monitoring System (HPMS) reporting. The availability, completeness, and file format of data pertaining to off-system roads are also uncertain.

Collection, maintenance and processing of data related to off-system roads that are not part of the HPMS are primarily the responsibilities of MPOs, counties, cities and other agencies. These local agencies usually limit the data collection process to short period counts, typically for 24 hours on a midweek day (Tuesday, Wednesday, or Thursday) once every calendar quarter or as requested for specific projects. Estimations of AADTs are rarely made using the available counts or on a consistent network basis.

Some districts exchange data with local agencies. For example, District 4 exchanges data once every two years with each of the counties in the district. It occurs more frequently in District 5. District 6 updates and checks for accuracy of the data collected by MPOs for off-system HPMS samples and exchanges system-wide traffic count data yearly. However, for many others there are no formal mechanisms for such information exchanges.

Data collected from the FDOT districts and counties include traffic counts, traffic count locations, base maps, TAZ structures, road classifications, and FSUTMS files. One county also provided a future land use map. The data formats vary from county to county and include hard copies, GIS files, Microstation files, and AutoCAD files.

Aggregated data at the county level are also collected for use in two model investigations. The data are obtained from the Division of Economic Development, Florida Department of Commerce. They include such information as county population, lane miles of state highway system, registered vehicles, municipal population, labor force, per capita income, and taxable sales for each of the 67 counties in Florida. This information is updated annually.

MODEL DEVELOPMENT

Multiple linear regression is chosen as the modeling method based on considerations of the FDOT's preference, the method's simplicity, and support of its appropriateness for application to AADT estimates in existing literature.

Because counties generally do not convert their traffic counts into AADT, the traffic data used for modeling are average daily traffic (ADT) instead of AADT. Therefore, the model variable is ADT. The use of ADT instead of AADT causes some potential problems, which will be discussed in the conclusions and recommendations.

Recognizing the distinct characteristics of different Florida counties/cities, which range from large metropolitan areas to rural areas, it is expected that different areas will have different sets of variables that contribute to AADT. Therefore, based on data availability and types of areas, four regression models have been developed:

(5) State wide model;
(6) Rural model;
(7) Small-medium urban model; and
(8) Large metropolitan area model (Broward County).

The state wide model and rural county model were developed using data aggregated at the county level. The data are obtained from the 1995 *Florida County Profile* published by the Florida Department of Commerce, Division of Economic Development, Bureau of Economic Analysis. This effort is motivated by the Purdue University's work in which a regression modelcovering all counties in the State of Indiana was developed. The objective was to investigate the feasibility of identifying a set of variables that are applicable state wide and determine how such a model will perform.

State Wide Model

The state wide model utilizes data aggregated at the county level. The sample size for model development is 107. The final functional form of the model is

$$ADT = 9643.704161 + 0.014645 \times POP - 0.155037 \times LABOR - 0.181236 \times INCOME \\ + 0.000005139 \times TAXABLE + 0.058710 \times VEHICLES$$

where

POP = the total population within a county;
LABOR = the total labor force within a county;

$INCOME$ = the per capita income of a county;
$TAXABLE$ = the taxable sales of a county; and
$VEHICLE$ = the total number of automobile registrations in a county.

The model performance is given in **Table E.1**.

Table E.1 Performance of the State Wide Model

R^2	Adj. R^2	Min. % |Error|	Max. % |Error|	Avg. % |Error|
0.2938	0.2439	8.00	1096.00	188.00

The model is apparently inadequate. The main cause is that some important predictors are missing in the model, resulting in a low R^2. The counties included in the model also have diverse characteristics resulting in large variances in the data that cannot be explained by the few variables used. All variables are also correlated as shown by more analyses. For instance, a model with two variables, population and per capita income, has a low R^2 of 0.1295. However, the variables are still slightly correlated.

Rural Area Model

A threshold value of 100,000 for county population is used to select eight counties for the development of the rural area model. These counties include Bradford, De Soto, Walton, Sumter, Gadsden, Jackson, Columbia, and Highlands Counties. The sample size is 27. The model has the following form:

$$ADT = 4853.49 + 0.12 \times POP + 0.26 \times LABOR - 18.93 \times LANEMILE - 0.0032338 \times VEHICLES$$

where
POP = the total population within a county;
$LABOR$ = the total labor force within a county;
$LANEMILE$ = the total lane miles of county roads in a county; and
$VEHICLES$ = the number of automobile registration a county.

The model performance is given in **Table E.2**.

Table E.2 Performance of the Rural Area Model

R^2	Adj. R^2	Min. % |Error|	Max. % |Error|	Avg. % |Error|
0.4525	0.2508	6.87	83.96	35.59

The rural area model shows a significant improvement over the state wide model mainly because the counties being modeled are more uniform in their characteristics. However, the model still suffers similar problems as the state wide model and is not useful due to its relative low R^2 and large predication errors. Further analyses also showed that all variables are correlated.

Small-Medium Urban Area Model

This model includes counties that have a population greater than 100,000, but excludes major metropolitan areas. A sample of 270 is randomly selected to generate the model. The model has the following form:

$$ADT = -13418 + 6770.23 \times LANES + 1580.14 \times ATYPE1 + 2.85 \times COM\_EMP + 1.78 \times HOT\_OCC$$

where

$LANES$ = number of lanes at the count station location in both directions;
$ATYPE1$ = area type of the count station location;
COM_EMP = commercial employment in a TAZ; and
HOT_OCC = number of hotel/motel occupants in a TAZ.

The model performance is given in **Table E.3**.

Table E.3 Performance of the Small-Medium Urban Area Model

R^2	Adj. R^2	Min. % |Error|	Max. % |Error|	Avg. % |Error|
0.6937	0.6856	0.48	59.29	27.75

This model represents a significant improvement over the state wide and rural area models. The main reason is that more detailed information is available to allow the inclusion of more relevant variables that can explain the variances in the traffic data to a larger degree.

Large Urban Area (Broward County) Model

In a large urban area, the transportation system consists of many different types of facilities, and the land use patterns are more complex. Consequently, the travel patterns are more varied. As a case study, Broward County has been chosen because of the availability of data in digital format. Digital data significantly reduce efforts required for data processing and provide an excellent opportunity to perform more sophisticated analyses using GIS tools. Data from 443 count stations are used for model development. The final model has the following form:

$$ADT = -12886 + 4689.86 \times LANES + 5227.57 \times FCLASS1 + 1388.27 \times AREA1 + 0.15 \times AUTO - 1224.06 \times ACCESS2$$

where

LANES	=	the number of lanes in both directions;
FCLASS1	=	functional classification of a roadway;
AREA1	=	land use type;
AUTO	=	the estimated total number of automobiles within a certain distance of a count station; and
ACCESS2	=	presence of other county roads nearby.

Table E.4 Performance of the Broward Model

R^2	Adj. R^2	Min. % |Error|	Max. % |Error|	Avg. % |Error|
0.6120	0.6069	0.86	61.99	23.73

It first appears that the Broward County model has similar performance as the small-medium urban model. However, a careful examination of the error distribution and percentage of testing points falling within a certain level of errors reveals that the Broward County model is superior even though its R^2 and adjusted R^2 are slightly lower than those of the small-medium urban model.

CONCLUSIONS

The advantage of the state wide and rural models is that the data used in these models are easily obtained and continually updated. However, the models do not perform well as indicated by their low R^2 and large errors in the testings. The possible reasons for the poor performance of the models include omission of important independent variables, especially information specific to local conditions, small sample size, and the large variances in the size and characteristics of the counties being modeled. However, the rural county model is an improvement over the state wide model because variance in the data is reduced when the counties being modeled have more similar characteristics. One serious problem suffered by both models is the correlation among all the variables, which suggest that the current variables set is inadequate and more variables must be considered.

For the small-medium urban area model, data other than ADT are obtained from the FSUTMS ZDATA files and are aggregated at the TAZ level. Compared to the county level data, TAZ level data are much more detailed and can account for data variation within the same county. The model performance is significantly improved compared to the first two models.

The Broward County model is unique in several ways. Firstly it is a large urban area with complex urban forms, transportation systems, and travel patterns. No literature on estimating ADT for such a large urban area using statistical or other techniques has been found. The model has five independent variables: number of lanes, function classification, area type, automobile ownership, and access to county roads. Except for automobile ownership, the values of other variables are easily obtained. However, some problems exist in the functional classification systems and they may be a source of error.

The Broward County model demonstrates that the characteristics of the roadways themselves explain the variation in ADT more than the socioeconomic factors, which do not appear to be significant as expected. The "buffer zone" method used to aggregate the data does not appear to be useful.

The variable sets in the small-medium urban model and the Broward County model are somewhat different in that commercial employment and hotel population are included in the small-medium urban model but not in the Broward County model. There may be several reasons including the different characteristics of the areas included in the models and the different methods for aggregating data used to develop the models.

The models generally have a rather large negative intercept. While the intercepts themselves do not have physical meanings, their large negative values tend to make the model underestimate ADT more often. This is likely a result of not including all relevant variables that have a significant impact on ADT.

Both the small-medium urban model and the Broward County model include number of lanes and area type as important factors, which indicates that these factors should be included if similar models are to be developed. However, area type definitions are not well understood and need further investigation.

Finally, the traffic data used are ADT instead of AADT. The data are not consistent since traffic counts are collected during different times throughout the year. This may introduce errors into the models especially considering Broward County has a significant seasonal resident population.

RECOMMENDATIONS

This research has identified sets of variables that may be used as predictors of ADT. These sets of variables, along with others that may be identified in future studies, may be used to classify roads into categories of similar roads for the estimation of ADT for roads that do not have traffic counts. However, before this classification is attempted, more study is needed to further investigate the issues arising from this research. The following research is recommended for consideration in future efforts to improve ADT prediction methods:

(1) Further investigate the development of a rural county model that incorporate more variable such as land use and accessibility that reflect local conditions.

(2) Further improve the small-medium urban area model by studying different data aggregation methods and by including county level data in the model.

(3) Test the use of a standard roadway functional classification system in models. The functional classification system that has been used in Broward County is different from the Federal Highway Administration's functional classification system. The county is currently completing the reclassification of all county roads according to the FHWA system. Therefore, the FHWA

system should be tested in the model to determine how much variance in ADT may be directly attributed to the functional classification system.

(4) Develop a better measurement of accessibility by considering intersection types and connectivity to expressways. For the rural model, include accessibility information which is missing from the current model.

(5) Investigate area type definitions and evaluate the appropriateness of its use in urban areas where land use patterns may be complex. Include land use information in the rural model.

(6) Study the impact of economic activities on AADT of a road segment by defining a service area for the corridor that contains that road segment. This larger service area will account for some of the through traffic that is not originated or destined locally.

(7) Test the Broward model on other large urban areas including the Miami-Dade County. Also develop a model based on Miami-Dade County data, since much of the needed data are available in GIS format including detailed land use information, which is not available in Broward County. The main benefits will be to compare the models in order to determine a common set of variables that may be used for establishing categories for estimating ADT, and to determine to what degree differences in geographic characteristics would affect the model structure.

(8) Develop tables that may be conveniently used to determine ADT for a roadway based on its characteristics as identified in the regression models.

(9) Test model performance over time to study the stability of model coefficients.

(10) Explore other potential techniques such as neural networks, which are good at capturing nonlinear relationship between ADT and variables, insensitive to correlations between independent variables, easy to develop, and may be continually updated with new data.

(11) Study seasonal factors for off-system roads. Broward County, for instance, has begun to convert ADT to AADT starting for 1998 data. The seasonal factors applied are developed based on traffic counts from permanent count stations on the state roads. The applicability of these seasonal factors to off-system roads needs to be carefully examined since many off-system roads possess rather different characteristics from those of state roads.

1. INTRODUCTION

1.1 Problem Statement and Research Objectives

Estimation of Annual Average Daily Traffic (AADT) is extremely important in traffic planning and operations for the state departments of transportation (DOTs), because AADT provides information for the planning of new road construction, determination of roadway geometry, congestion management, pavement design, safety considerations, etc. AADT is also used to estimate state wide vehicle miles traveled on all the roads and is used by local governments and the environmental protection agencies to determine compliance with the 1990 Clean Air Act Amendment. Additionally, AADT is reported annually by the Florida Department of Transportation (FDOT) to the Federal Highway Administration.

In the past, considerable efforts have been made in obtaining traffic counts to estimate AADT on state roads. However, traffic counts are often not available on off-system roads, and less attention has been paid to the estimation of AADT in the absence of counts. Current estimates rely on comparisons to roads that are subjectively considered to be similar. Such comparisons are inherently subject to large errors, and also may not be repeated often enough to remain current. Therefore, a better method is needed for estimating AADT for off-system roads in Florida.

Literature on estimating AADT for roads that do not have traffic counts is limited. Most of the relevant literature is on either estimating traffic volume for freeways or extrapolation of 24-hour or 48-hour traffic count data to obtain AADT. One attempt to estimate AADT for off-system roads was made by a research group at Purdue University, Indiana, in which a multiple regression method was employed utilizing aggregated data at the county level. Several other studies have been reported on estimating AADT for state roads using statistical methods. The predictors include county population size, total number of through lanes, state/non-state code indicating jurisdictions, location type (rural/urban), personal income level, vehicle registrations, etc. A common problem reported was the lack of data for some potentially important predictors. For example, in a Minnesota study, it was proposed to use the population within a certain distance of the roadway as a predictor, but the attempt was abandoned due to unavailability of the data.

This study investigates the possibility of establishing one or more models for estimating AADT for off-system roads in Florida. It is intended by the FDOT that, once such models are developed, important variables that contribute to AADT may be identified and their impact on AADT may be quantified. As a result, a matrix or a classification method may be developed for grouping roads into different categories. Roads in each category would possess similar characteristics. AADT may then be estimated for roads that do not have traffic counts based on those from other roads in the same category.

To accomplish the goal of developing a practical method for estimating AADT for off-system roads the development of a good model or models is required that will relate contributing factors to AADT. The success of the model development effort depends not only on the modeling techniques chosen, but also on the data availability and methods of data aggregation. While FDOT maintains a wide range of different data on state roads, there is not a central database available that provides adequate information about off-system roads such as roadway condition inventory, traffic counts, and many other types of potentially relevant data. Therefore, data collection, data compilation and data processing are as significant as modeling.

The issues to be addressed in this research include data availability, data collection and processing, suitability of various data for model development, choice of modeling techniques, model accuracy, model applications, and model improvements. This document presents the results of the research efforts addressing the above-mentioned issues.

1.2 Research Methodologies

The research involves a literature review to learn about the current state-of-the-art in techniques for estimating AADT in the U.S. as well as in other countries. The literature review also provides information about the data used in AADT estimation methods and their processing. Based on the results of literature review and preference of FDOT, multiple regression has been chosen as the modeling technique. The choice is based on the simplicity of regression analysis and existing literature that supports its appropriateness for application to AADT estimation.

This research also includes a major data collection effort. Various levels of governments in the state were contacted to gather information about their data collection programs and data needed for this project. Problems concerning data availability and data format were identified through survey and telephone interviews.

Recognizing the distinct characteristics of different Florida counties/cities, which range from large metropolitan areas to rural areas, it is expected that different areas will have different sets of variables that contribute to AADT. Therefore, based on data availability and types of areas, four regression models have been developed:

(3) State wide model;
(4) Rural model;
(5) Small-medium urban model; and
(6) Large metropolitan area model (Broward County).

This report documents each of the research activities and the models.

1.3 Organization of the Report

In the remainder of this report, a literature review is first provided in Chapter 2. Chapter 3 describes the data collection effort and presents the result of that activity. Four different regression models are developed for different types of geographic areas. These models are presented in Chapter 4.

Finally, Chapter 5 provides the conclusions derived from this research, and Chapter 6 presents the recommendations regarding future research directions.

2. LITERATURE REVIEW

The motivation behind the study of AADT estimation is to develop an effective data collection plan, to reduce the data collection cost, and to produce AADT estimation with a better accuracy. In the following sections, literature related to estimation of AADT is reviewed. Two types of problems are associated with AADT estimation. One is to convert coverage counts to AADTs at given locations and the other is by using available coverage counts, usually already converted to AADTs to predict AADTs at locations where such counts are not available. Literature concerning these two types of problems is reviewed in Section 2.1 and Section 2.2, respectively.

2.1 Conversion of Coverage Counts to AADTs

The objective of highway traffic counting program is to obtain estimates of average daily traffic volumes, which is a basic parameter used in transportation planning, design, and operation. It is not generally feasible to conduct year-round traffic counts on every highway segment in a jurisdiction. Therefore the estimates are most often computed from short period coverage counts from portable Automatic Traffic Recorders (ATRs) located at different strategic points to provide continuous spot traffic data. Coverage counts are short period count and are often collected for a 48-hr period (or a minimum of 24 hours) once a year on weekdays (Erhunmwunsee 1991). These short coverage counts need to be adjusted for the estimation of AADT to remove the bias caused by the seasonal fluctuation.

All methods for estimating AADT from ATR data are based on the assumption that the traffic volume on a highway section exhibits a regular pattern according to the season of the year and the day of the week. The most commonly used approach in obtaining AADT for a particular roadway is to first conduct a short-term traffic count (usually over a period of 48 hours) and then convert the short-term traffic volume into AADT (Faghri and Hua 1995). The conversion procedure may be a simple multiplication of the average daily traffic volume (ADT), obtained from the short-term counts, with a reasonable coefficient. The key to this type of problems then becomes the determination of the coefficient, which is often a combination of a number of factors that account for, for example, traffic volume variation due to different weekdays, different months, and differences in the number of axles of vehicles.

Ritchie (1986) summarized the basic model of estimating AADT for a particular highway segment based on a single, short-duration count as follows:

$$AADT = VOL \times F_S \times F_A \times F_G$$

where

VOL = average 24-hour volume from a standard 72-hour Tuesday-Thursday short count;
F_S = seasonal factor for the count month;
F_A = weekly alxe correction factor if VOL is in axles, equals to 1 if VOL is in vehicles;

F_G = growth factor if VOL is not a current year count, equals to 1 otherwise.

The seasonal factor is the ratio of ADT to AADT. It indicates the changes of traffic on the roadway during different seasons in a year.

The axle correction factor is given by the reverse of the average number of axles per vehicle in a given factor group (usually highway functional class):

$$F_A = \left(\sum_C (Axles_c) \times (P_C)\right)^{-1}$$

where

F_A = the reverse of the average number of axles per vehicle in a given factor group;
$Axles_C$ = the number of axles per vehicle in class C;
P_C = proportion of vehicles in class C (system level estimate).

The growth factor often represents a relatively minor part of the factoring process that obtains AADT from short counts. When historical data are used to estimate AADT, growth factors are needed. Several methods for estimating growth factors exist. Normally, a factor is obtained by computing the ratio of AADT in recent years to that in the previous years for each ATR in a group and applying the regression analysis procedure.

Another approach to estimating AADT is to first classify streets by their basic functions, such as expressway, arterial, collector, and local street (Erhunmwunsee 1991). It requires seasonal count data to compute the seasonal factors. The 24-hour coverage count data are used to estimate AADT for each of the various classes of streets. Albright (1990) also recommended this method because of its simplicity in application after a comparison of the precision of AADT estimation using clustering analysis. However, Erhunmwunsee pointed out the fact that the result was based more on judgement rather than objective measurement as a disadvantage of the method.

A third method for estimating AADT is the averaging method (Erhunmwunsee 1991). Using this method, monthly counts from each ATR are collected and averaged to produce AADT for each ATR station. Erhunmwunsee states that "an advantage of this method is that, from the standpoint of inductive statistics, averaging is mathematically tractable and has many desirable properties for sound testing. A disadvantage, however, is that the cost of traffic data collection has been increasing steadily over the years". Broward County uses this method to convert quarterly 24-hour counts to AADTs.

In recent years, with progress made in artificial intelligence, neural networks have become a popular technique for control, pattern recognition, and prediction problems. They have also been applied to AADT estimation from coverage counts. Neural networks are a computational model fashioned after human brains, which consist of interconnected neurons and work by getting input from sensors, then propagating the input through the network of neurons to produce a response.

During the propagation of the input, each neuron in the network accepts input from other neurons and sends its output to other connected neurons. Building an artificial neural network involves determining a network configuration (see **Figure 2.1**) and training the network by feeding it with a large set of input-output pairs. A neural network will automatically adjust the weights of each neuron based on some algorithm and eventually produce a model that is able to produce the expected output given an input. The neural network has several advantages including the ability to model nonlinear functions, the ability to learn and handle situations that have not been encountered during its training, and fast response, which is important to real-time applications.

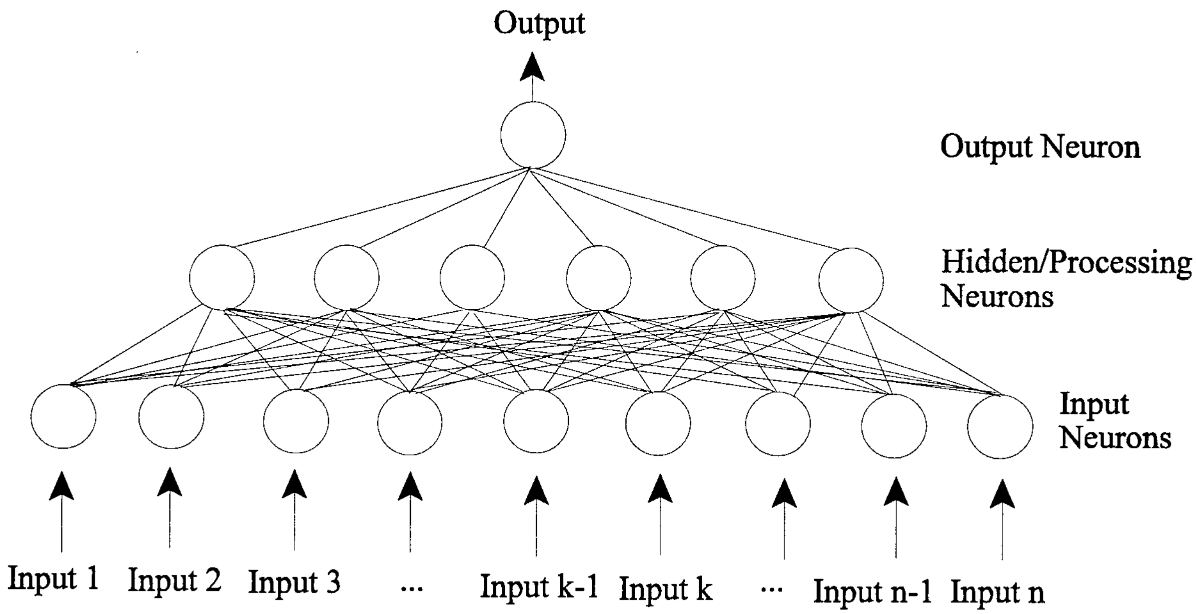

Figure 2.1 Artificial Neural Network Model

Applications of neural networks for estimating AADT from traffic counts have been reported in (Sharma *et al.* 1999), in which the traditional factor approach is compared with a neural network model. Accurate estimates by the factor approach require that ATR and sample sites be appropriately classified into one of the several ATR groups, which in practice is difficult. A neural network model that did not impose such requirements was developed and it was found that the performance of the model based on two 48-hour counts was comparable to that of the traditional factor method based on a single 48-hour counts. The 95th percentile error range for the neural network model was 14.14% to 16.68%. It was concluded that due to the difficulty in correctly assigning ATR and sample sites to ATR groups, the neural network approach was easier to use and performed better than the traditional factor method.

In (Faghri and Hua 1995), a neural network model that determines the monthly seasonal factors for roadway groups such as urban, rural, and arterial, was used to estimate AADT from traffic counts. The neural network model was compared with cluster analysis and regression analysis methods, and was found to outperform the latter two. The average errors from the three methods were 0.044, 0.24, and 0.233, respectively.

2.2 AADT Prediction Models

AADT prediction is a travel forecast problem. There exist many travel demand forecasting models that predict a future year travel based on a model calibrated for a base year. Such forecasting is usually for the purpose of long range transportation planning, usually set for a future year of 10 to 20 years from the base year. These models require substantial resources for the collection of base year data, including transportation facilities, traffic volumes, transit services, transit ridership, socioeconomic data, demographic data, etc., and future socioeconomic and demographic data. Modeling efforts normally take a few months to over a year depending on the size of the urban area and complexity of the system being modeled. Because of the resources required to carry out such a forecast, travel demand models are updated infrequently, for example, every 10 years. Even though the base year forecast may be quite good estimate of AADT, and AADT may be estimated using the base year forecast using a growth factor, as time elapses, the model accuracy will deteriorate. Therefore, relying on long term travel forecast models for determining AADT on a yearly basis is problematic. When a region is not within the coverage area of a urban travel demand model, other alternatives have to be found. In this section, alternative methods to travel forecast models are described.

2.2.1 Elasticity Based Approach to Estimation of AADT

An elasticity-based approach is the most commonly used method in the development of prediction models (Mohamad *et al* 1998). A study by Neveu (1983) resulted in models to forecast future-year AADT as a function of base-year AADT, modified by various demographic factors. In this research, multiple linear regression was used to identify factors that best estimate AADT and their respective elasticities. The factors examined include population, number of households, automobile ownership, and employment. These data were collected at town, county and state level. Each of Neveu's model is relatively simple, with only one or two independent or predictor variables. The study producd a set of nomographs that gave quick estimates of the growth factor, i.e., the elasticity portion of the model. The number of households was found to be a better determinant of travel than population. For example, for his interstate traffic prediction model, the AADT is a function of county automobiles and town households. For his principal arterial model, the AADT is a function of county households and town population, and for his minor arterial and major collector model, the AADT is a function of town households. Additionally, the functional classification was identified as the only factor that had a significant effect on traffic growth rates.

2.2.2 Regression Models

Time series and multiple regression are two regression models for forecasting AADT. Time series model is based on historical data. Trend lines drawn through prior year data observations are extrapolated to the target year. Time series model only deals with the traffic data, while multiple regression model is concerned with obtaining traffic value from socioeconomic and other measures, i.e., a relationship between traffic and associated factors. There has been research on the application of both methods for AADT estimation. Because of the lack of historical data for off-system roads, time series model are not considered this research.

Kentucky Study

In a Kentucky study conducted by Deacon (1987), a two-step modeling process was developed to forecast highway traffic volumes on the state highway systems. In the first stage, influences of state wide economic conditions on the overall growth in travel on Kentucky roads and streets were quantified. A regression model that includes personal income, vehicle registration, fuel price and miles of highway was developed in this stage. In the second stage, local effects were investigated to explain how traffic grew on a particular facility. The formulas used in the first stage are:

$$\textit{Vehicle Registration} = a + b\ (\textit{Personal Income})$$
$$\textit{Vehicle Miles} = c\ (\textit{Vehicle Registrations}) + d\ (\textit{Fuel Price})$$
$$\textit{AADT} = (\textit{Vehicle Miles}) / (365 \times \textit{Miles of Highway})$$

In the formulas, a, *b, c and d* are calibration constants. Here all variables except fuel price were aggregated at the state level.

During the second stage, a "cross tabulation analysis" related to site-specific traffic growth was developed. In the cross-tabulation models, sites grouped in the same cell were considered to be identical, and the average entry for the group of sites was used to represent the most likely estimate of AADT for any particular site. The analyses were performed using two models. The first enabled estimates of current or base-year volume; and the second described how the volume was expected to grow in future years. When the base year traffic count was not available, an interpolation or extrapolation, or both, were made based on a weighted, least-squares calibration of the linear growth curve:

$$\textit{ADT} = a + b\ (\textit{Year})$$

where

ADT = average daily traffic;
Year = date of calendar year; and
a and *b* = calibration constants.

To minimize the extent of erroneous estimation, estimates were not made unless at least four years of volume data were available, and extrapolations and interpolations were limited to within six years of an actual count.

The initial independent variables chosen included functional class, base-year volume, membership in federal-aid system, administrative class, route signing, development density (rural, small-medium urban, and urbanized), access control, urbanization, geographic area, and county population growth. These variables are arbitrarily chosen from potential variables.

Minnesota Study

A Minnesota study by Cheng (1992) conducted a regression analysis for predicting AADT. The potential predictors of traffic volume were chosen from variables currently available in the road-log (RLG)database. The following 13 data items are considered to be useful in estimating traffic volumes:

- Route system: the ownership of the road section, including state roads and non-state roads;
- City population: the population size of the city where the road section is located;
- Population size of the county;
- Location (urban or rural);
- Functional classification: the usage of a road section. There are five functional classes for rural and eight for urban road sections, respectively. The five functional classes for rural sections are: interstate, other principal arterial, minor arterial, major collector, and local roads. The eight functional classes for urban sections are: interstate, other connecting freeway, other non-connecting freeway, other connecting link, other non-connecting link, minor arterial, collector, and local roads;
- Intersection category: the route system of any intersecting road sections;
- Special road sections: a variable indicating whether a road section has a special status such as "national forest highway" or "great river road";
- Federal-aid system: a variable indicating whether a road section receives federal aid, and if yes, what type of federal aid;
- Control of access: a variable indicating whether access to a road section is uncontrolled, partially controlled, or fully controlled;
- Through lanes on the road section: total number of through lanes (in both directions) on the road section;
- Type of truck-route: there are eight truck-route classifications;
- Width of the road sections: the width of the road sections, in feet, including sidewalks if any, and non-traffic-carrying lanes such as space for parking;
- Surface type: the type of pavement on a road section. There are twenty-five categories.

Variables regarded as not usable were dropped and statistical methods were used to reduce the number of predictors to four: route system, populations size of the county, total number of through lanes, and rural/urban identification code. The final model is:

$$AADT = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \beta_4 X_3 + \beta_5 X_4 + \beta_6 X_2^2$$

where

X_1 = county population size;
X_2 = total number of through lanes;
X_3 = state/nonstate code; and
X_4 = rural/urban code.

When selecting data items for regression analysis, it was noticed that either some of the variables were not usable or they added significant complexity to the model. For example, city population size was missing for the majority of the road sections and functional classification had a large number of categories. Functional classification was discarded because of the difficulty of deriving a consistent coding system due to the diversity of classifications. Another limitation of regression analysis noted was that some important predictors of traffic-volumes were not available, such as the population size within certain distances of a road section. Other predictors that were potentially useful but were difficult to obtain included: major intersections within a given radius, peak hour volume, and geographic location of a road section.

Indiana Study

A recent study in Indiana presented the following multiple regression model to predict AADT on county roads (Mohamad *et al.* 1998):

$$AADT = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

where

X_1 = location (urban/rural);
X_2 = access (easy access or close to the state highway or not);
X_3 = county population; and
X_4 = total arterial mileage of a county.

The initial independent variables were chosen as:

- county population
- county households
- county vehicle registration
- county employment
- county per capita income
- county state highway mileage
- location (urban or rural)
- presence of interstate highways
- accessibility

In the study, 89 count stations in 40 selected counties in Indiana were used. The R^2 of the selected model is 0.75. Several variable selection methods in multiple regression analysis were employed to select the best model in the building process. It was found that multiple linear regression models provided reasonable statistical results.

It is noted that in the Purdue study, the model was developed based on data from 89 count stations from 40 counties in Indiana, with an average of about 2.5 stations per county. This raised the question as to how representative these data points are and if the model is biased toward the selected data set.

3. DATA COLLECTION

3.1 Introduction

Data availability is critical for both model development and for future model applications. Generally speaking, there are two groups of data relevant to this research: AADT estimated from traffic counts and other non-traffic data that may be used to build models to estimate AADT. There are three types of traffic count data: data obtained from continuous counts with permanent traffic recorders (PTRs) devices, control or seasonal counts, and short period or coverage counts. PTRs are permanent installations that continuously monitor hourly traffic volumes year-round. Control or seasonal counts are taken two to 12 times a year, for periods of time ranging from 24 hours to two weeks. Short-period counts or coverage counts are collected manually for periods of time ranging from six hours to seven days. They are usually collected for specific projects and adjusted using predetermined seasonal factors.

The FDOT generally has good traffic data collection programs that meet most of the department's needs, such as planning, design and operation assistance and administration functions. The department conducts traffic data collection to determine the volumes, type of vehicles and the weight of trucks using the highway network of Florida. While the FDOT maintains traffic data and roadway inventory information on state roads, the availability, completeness, and file format of data pertaining to off-system roads are less certain.

The seven FDOT districts have essentially the same program to monitor periodically sections of off-system roads or *"active off-system roads"* functionally classified as a collector or above to be included in the Department's Highway Performance Monitoring System report, which requires AADT information. The AADT estimates are based on data that districts collect every three years for a small number of roads including those that have a railroad crossing., which do not represent the current conditions of the entire off-system road network. For example, during 1997-1998 District 6 collected data on 70 off-system segments of roads totaling about 80 miles out of 574 miles of principal and minor arterial that compose the District 6 road network.

Collection, maintenance and processing of data related to off-system roads are primarily the responsibilities of MPOs, counties, cities and other agencies. These local agencies usually limit the data collection process to short period counts, typically for 24 hours on a midweek day (Tuesday, Wednesday, or Thursday) once every calendar quarter or as requested for specific projects.

This chapter documents the information obtained from all 67 counties, seven FDOT Districts, cities, metropolitan planning organizations (MPOs), public work offices, and regional planning councils in the State of Florida.

3.2 Data Collection Methodology

The objective of the data collection effort is to identify the data collection programs of various state and local agencies, assemble data that may be pertinent to this project, and document, summarize, and evaluate all required data elements. The data collection effort was carefully planned to obtain dependably, quality information from all offices handling data on off-system roads. It has involved a comprehensive literature review, telephone interviews with districts' personnel responsible for traffic data collection, and a survey.

The literature review has included mainly the following documents:

- Manual on Uniform Traffic Studies, FDOT.
- Planning and Field Data Collection, FHWA.
- Highway Performance Monitoring System (HPMS) Field Manual, US DOT, April 1994.
- Highway Capacity Manual, TRB, 1994.
- Florida's Level of Service Standards and Guidelines Manual for Planning, FDOT 1995.
- Roadway Characteristics Inventory Manual, FDOT.
- Manual of Transportation Engineering Studies, ITE, 1994.
- Manual on Uniform Traffic Control Devices (MUTCD), 1997.
- Evaluation of Traffic and Highway Data Collection Survey, Parsons Brinckerhoff, 1995.
- Guidelines for Traffic Data Programs, AASHTO.
- Traffic Monitoring Guide, FHWA.
- Traffic Data Collection, FDOT.
- General Interest Roadway Data/RCI Features and Characteristics Handbook.

Literature research and telephone interviews with the representative of each district's planning office who handles data on off-system roads such as district planning managers, modeling coordinators, traffic engineers, and/or statistics department administrators resulted in a database of appropriate contacts (see **APPENDIX B**) for actual data collection. Conversations with the representatives also provided important directions for the preparation of the surveys that were used as the main instrument for the data collection effort.

The survey was prepared based on the findings from the literature review and interviews with the representatives from the district planning offices. The survey asked the district offices to provide general information regarding the districts and detailed information of the state and off-system roads including roadway characteristics, frequency and methodology of data collection, data maintenance, and data processing standards. The survey form is included in **APPENDIX A**.

The survey form was sent to district offices together with a letter explaining the goal and importance of the research. Follow-up calls to the contact persons were made to confirm the receipt of the surveys. Telephone communication was used to promptly answer any questions raised by the respondents at the district offices.

Since preliminary investigation revealed that the districts did not have a comprehensive data collection program for off-system roads, the district offices were also requested to provide information about other agencies, firms and/or persons handling off-system roads data in the survey form. Survey responses provided by the districts indicated that MPOs and regional planning councils (RPCs) were the primary agencies that handled most of the data of off-system roads.

The second step of the data collection involved obtaining information from the MPOs and RPCs state wide. The results of the data collection from district offices and from local agencies are summarized in the following two sections, respectively.

3.3 Information Available at District Level

According to the *State Highway System Report 1: All Roads* issued by the FDOT, the Highway System consists of 11,929.2 centerline miles as of December 31, 1997.

The on-system state road network is classified according to the Federal Classification as urban and rural roads. The urban road networks consist of 3,386.8 centerline miles of urban principal arterials, 1,391.4 centerline miles of urban minor arterials, and 141.7 centerline miles of total collectors. The rural roads include 4,633.7 centerline miles of principal arterials, 1,953.9 centerline miles of minor arterials, 404.9 centerline miles of major collectors, and 13.9 centerline miles of minor collectors. **Figure 3.1** shows the centerline mileage for the state road category. The total mileages of roads by category and district are summarized in **Table 3.1**. The table shows that the off-system road milage makes up the major portion of the district roadway system. For instance, in Districts 2 and 4, state road mileage is only about 15% of that of the off-system roads.

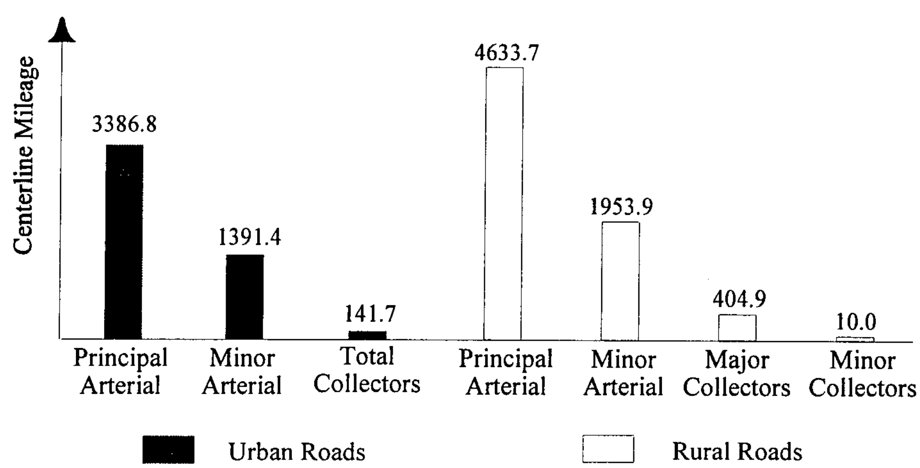

Figure 3.1 Centerline Mileage of State Roads by Categories

Table 3.2 summarizes the survey responses from the districts regarding data collection frequency for state and off-system roads, and existence of any coordination between the district offices and local agencies. In summary, District 2 does not collect data on off-system roads; Districts 1 and 6 occasionally collect data for specific projects. District 1 also collects data at rail crossings every three years if the data is not available from other local agencies. District 5 reported that 100% of the network was being surveyed by the district and that traffic counts were being collected year round as time permitted. This data collection effort is a part of a process to implement a GIS at the district, which also included the development of a GIS base map. The number of locations and segments to be surveyed are determined after a detailed analysis of the intersection data, as reported by the county and cities, and a field inventory. Counts are usually available by the county and/or city. If the information is not available, permanent stations and/or portable equipment are installed to obtain the information.

Table 3.1 Centerline Mileage of Off-System Roads and State Roads by Districts
(Source: *State Highway System Report, 1997* and Survey Responses)

Districts	*Total State Roads Mileage*	*Off-System Roads*	*Total Mileage of the Different State Road Categories by District*								
			Rural Principal Arterial	*Rural Minor Arterial*	*Rural Major Collector*	*Rural Minor Collector*	*Urban Freeway*	*Urban Expressway*	*Urban Principal Arterial*	*Urban Minor Arterial*	*Urban Collector*
1	1846.2	n/a	921.8	278.6	39.1	0.0	n/a	n/a	383.7	194.6	28.4
2	2530.3	17628.3	1115.8	657.6	75.2	0.0	n/a	n/a	418.5	256.3	6.9
3	2370	n/a	838.1	671	286	13.9	n/a	n/a	324.7	207.0	28.7
4	1386.4	8403.7	393.6	40.2	2.0	0.0	n/a	n/a	625.5	252.2	72.9
5	2075.1	n/a	843.4	275.3	0.4	0.0	n/a	n/a	773.0	180.8	2.2
6	691.4	574.0	176.4	11.8	0.0	0.0	n/a	n/a	342.9	160.3	0.0
7	1027.4	n/a	4,633.7	19.3	2.2	0.0	n/a	n/a	518.5	140.3	2.6

Notes: n/a = not available.
- = not reported

Some districts exchange data with local agencies. For example, District 4 meets with the counties twice a year and exchanges traffic data with the appropriate county agencies at these biannual meetings. District 5 exchanges data once every two years with each of the counties in the district. District 6 updates and checks for accuracy the data collected by MPOs for off-system HPMS samples. District I reported that it has been found a big difference between data reported by counties and cities and those obtained by the district. However, for many others there are no formal mechanisms for such information exchanges.

Table 3.2 Data Collection Frequency by District and Coordination between Districts and Local Agencies

Districts	*Frequency of Data Collection on State Roads*	*Frequency of Data Collection on Off-System Roads*	*Other Agencies Collecting Off-system Data*	*Off-system road data collection coordination and data sharing*
1	annually	data collected for special projects if not available at other agencies; at rail crossings every three years if unavailable	counties, cities, MPO and contractors	n/a
2	annually	data collected for specific projects; at rail crossings every three years if not available	counties and cities	data is exchanged between Jacksonville MPO and Gainesville MTPO
3	once per year, 24 hrs in urban areas and 48 hrs in rural areas	once per year, 24 hr in urban areas and 48 hr in rural areas	n/a	n/a
4	twice per year	exchanges traffic data with the appropriate county agencies twice a year	counties	data exchange once every two years with each county in district
5	annually at established stations	annually at established stations	counties and cities	data exchanged with counties and cities
6	annually	data collected for specific projects	MPO, Public Works	collection by MPO, input and accuracy by the district
7	annually	-	n/a	n/a

Notes: n/a = not available
\- = not reported

Table 3.3 lists for each district the type of data collected on state roads. The major problem reported by the districts regarding data collection on off-system roads is that many of these roads have never been inventoried in the past. As a result, the collected data cannot be verified. Another problem, as pointed out by District 6,is the lack of a standard procedure for performing data collection for off-system roads, in contrast with the sophisticated program established for data collection on the state roads.

Table 3.3 Data Collected by District Offices

Districts	*State-road data collected, complied, analyzed and reported*
1	Volume, % of trucks, peak hour, weight, speed, vehicle type, directional, etc
2	Traffic count, vehicle type, peak hour volumes
3	15- minute interval volume counts, LOS and peak hour, 24 hour trucks
4	24- and 48-hour ADT, Classification using FDOT SPS software
5	features specified in roadway inventory checklist
6	All types of features
7	Volume and classification data for traffic counts

Table 3.4 provides information on data collection methods used at districts to collect traffic data on both state and off-system roads. **Table 3.5** summarizes information regarding methods used by districts to estimate AADT from traffic counts and level of accuracy required for the estimation. It

needs to be pointed out that this estimation involved converting traffic counts to AADT for locations where the counts were conducted. It did not involve estimating AADT for roads where counts were not available. The districts commonly use growth factors to estimate current traffic from historical traffic data. District 4 reported to estimate AADT for roads that do not have traffic counts by comparing them to similar roads. The actual procedure and criteria used were not reported.

Table 3.4 Methods for Traffic Data Collection for State and Off-System Roads

Districts	*State road data collection method*	*Off-system roads data collection method*
1	permanent stations, portable equipment, manual counts	portable stations, portable equipment
2	permanent stations, portable equipment	portable stations portable equipment
3	permanent stations, portable equipment, manual counts, video recording	portable equipment, manual counts
4	permanent equipment, portable equipment	n/a
5	permanent equipment (section # & milepost) portable equipment	n/a
6	permanent stations, portable equipment, manual counts	portable equipment
7	permanent stations, portable equipment	permanent stations, portable equipment and video recording.

The districts were also requested to provide GIS coverage in Arc/Info export format with their related databases for TAZ, road network, and land uses, digital files of ZDATA1 and ZDATA2, and historical traffic counts for state and off-system roads in digital or printout format. The ZDATA1 file contains socioeconomic data such as single- and multi-family dwelling units, percentage of vacant dwelling units, single- and multi-family population, percentage of households of different types with no vehicle, one vehicle, and two or more vehicles, respectively, as well as the hotel data, such as number of hotel rooms, percentage of rooms occupied, and number of hotel guests. The ZDATA2 file contains data of industrial employment, commercial employment, service employment, total employment, and school enrollment by TAZ. The FSUTMS network information includes the XY file that defines the network nodes and the LINKS file that defines the network links used to model the transportation network. FSUTMS data were used to obtain information about road facility type, number of lanes, area type, and socioeconomic conditions for regression analysis.

All districts provided hard copies and electronic files with historical traffic counts for state roads, and two districts provided hard copies with traffic counts and count sation maps. District 2 provided GIS coverage for road network as well as ZDATA and TAZ files for Alachua, Clay, St. Johns, and Putnam counties and Jacksonville MPO. District 4 provided GIS coverage for road network as well as ZDATA 1 and ZDATA 2 and TAZ files for Broward county. District 3 provided ZDATA1 and

ZDATA2 files for all four urban areas within the district, including Tallahassee, Panama City, Fort Walton Beach, and Pensacola. District 5 reported that the information was not available.

Table 3.5 Methods for AADT Estimation for State and Off-System Roads

Districts	*Data and method used to estimate AADT*		*Accuracy level required for AADT estimates/forecast data*	
	state roads	*off-system roads*	*Off-System Roads*	*State Roads*
1	*Data*: vehicle counts, vehicle	n/a	n/a	n/a
2	*Data*: vehicle counts	n/a	n/a	10%
3	Data: vehicle counts, statistical records Method: for stations not surveyed AADTs are estimated based on growth factors from previous years	Same as for state roads	10%	10%
4	counts are not estimated on state roads.	RCI, count are estimated, if not available from the county, by comparing to a similar roadway	n/a	n/a
5	*Data*: vehicle counts, vehicle classification, statistical records *Method*: growth factors are used to estimate AADT statically that not collected	*Data*: Vehicle counts, *Note*: AADTs are not estimated where counts are unavailable.	n/a	n/a
6	*Data*: vehicle counts, vehicle classification, vehicle weights, statistical records *Method*: adjustment factors from permanent stations and data from vehicle classification	*Data*: vehicle classification, vehicle weights, statistical record *Method*: same as for state roads *Note*: AADT estimations only in Monroe, not in Miami-Dade	10%	10%
7	*Data*: vehicle counts, vehicle classification, RCI *Method*: volumes data are processed at the central office	n/a	5%	5%

3.4 County Level Data

Because FDOT dose not have extensive traffic data on off-system roads, attempts were made to identify availability of data from local agencies that may be involved in traffic data collection. The agencies and the corresponding contact persons were identified from the survey forms returned by the FDOT district offices. A list of contact persons from counties in each FDOT district is provided in **APPENDIX B**. These people were contacted. There were 36 responses, representing 53 percent of the counties. It was found that the data collection on off-system roads and the data processing procedures are inconsistent. There is a lack of standards for recording, maintaining, and processing the data colleted, especially for off-system roads. The formats of the data obtained varied from the latest version of digital files (GIS, CAD, EXCEL, and LOTUS) to illegible hard copies. It was reported by some of the agencies that they were in the process of updating the database files. **Table 3.6** summarizes the information received from the counties, cities and other agencies that handle off-system road data. The data received included AADTs, socioeconomic data from ZDATA1 and ZDATA2 files of the FSUTMS, FSUTMS network, and other information such as count station maps, historical traffic counts, and roadway improvement projects. In the table, the numeric entries indicate the year(s) for which the data are available.

Table 3.6 Off-system Data Provided by Counties

District	*County*	*AADT (Year)*		*ZDATA (Year)*	*FSUTMS Network (Year)*	*TAZ*		*Other Information*
		Digital	*Hard Copy*			*Digital*	*Hard Copy*	
1	Charlotte	90	-	90	90	√	-	-
	Collier	97	97	90	-	√	√	Transportation Plan Year 2004-14 &24 Historical Traffic Count 1973-97
	Lee	-	90-91	-	90	√	-	Traffic Count Location
	Manatee	91-96	-	00-20	90	-	-	Count Station Map
	Polk	93	93	90	90	-	-	Count Station Map
	Sarasota	91-96	-	00-20	90	√	-	Road Classification Map
2	Alachua	90,97	90,97	00-20	90	√	-	-
	Clay	-	96,97	90,20	-	√	-	-
	Duval	91,97	-	90,20	90	√	-	-
	Putnam	-	-	90,20	-	√	-	-
	St. Johns	-	-	92,20	-	√	-	-
3	Bay (Panama)	-	89,05	93,20	93	√ (CAD)	-	Count Station Map
	Escambia (Pensacola)	-	89-05 90,96	92,20	97,20	-	-	Count Station Map
	Jackson	-	90	-	-	-	-	Count Station Map
	Jefferson	-	-	-	-	-	-	Count Station Map
	Leon	90,96, 97	-	90	90	√ (CAD)	-	-
	Liberty	-	-	-	-	-	-	Count Station Map
	Okaloosa	90,96, 97	-	95	95	-	-	-
	Santa Rosa	-	90,96	-	-	-	-	-
	Wakulla	-	-	92-20	-	√	-	-
	Walton	-	-	92-20	-	√	-	-
	Washington	-	-	-	-	-	-	Count Station
4	Broward	96	94	-	-	-	-	Hwy Imp. & Road Class.
	Indian River	-	90-94	90	90,20	-	-	-
	Martin	96	-	90	90	-	-	Count Station & Road
	Palm Beach	95,96	-	90	-	√	-	-
	St. Lucie	93-97	-	90	-	-	-	Road Map Classification Future Land Use Dsg

District	County	AADT (Year)		ZDATA (Year)	FSUTMS Network (Year)	TAZ		Other Information
		Digital	Hard Copy			Digital	Hard Copy	
5	Brevard	-	-	-	-	-	-	-
	Flagler	-	-	-	-		-	-
	Lake	-	-	-	-	-	-	-
	Marion	90	89-93	90,95	90	-	-	Count Station Map
	Orange	94	94	90	90	√	-	Count Station Map
	Osceola	97	-	-	-	-	-	-
	Seminole	-	-	-	-	-	-	-
	Sumter	-	89-98	-	-	-	-	Count St. & Road Class.
	Volasia	96,97	-	90	90	√	-	Count Station Map
6	Dade	-	90	90	90	-	-	Count Station & Road
	Monroe	-	90	-	-	-	-	-
7	Citrus	-	87-97	90	90	-	-	Count St. & Road Class
	Hernando	-	-	90	-	-	-	Count Station Map
	Hillsborough	-	-	90	-	-	-	Count Station Map
	Pasco	-	90	90	-	-	-	-
	Pinellas	-	97,98	90,91,	90	√	-	Count Station & Road

It is noted that the FSUTMS data used in the analysis are of 1990, since these are the most recent data available from most counties.

In order to determine the locations of the AADT counts, some counties provided count maps in hard copy, while others provided digital files. For instance, the facility type, area type and TAZ number of Palm Beach County can be shown in ArcView in GIS files. Bay County and Leon County provided TAZ numbers in AutoCAD format. St. Lucie County provided a MicroStation file in DGN format.

Aggregated data at the county level are also collected and used in two model investigations. The table in **APPENDIX D** provides the 1995 county level data in 67 counties in Florida, including county population, lane miles of state highway system, registered vehicles, municipal population, labor force, per capita income, and taxable sales in each county. The counties are ordered according to population in 1995 from the lowest (6,043 in Lafayette) to the highest (2,031,336 in Dade). However, the 1995 AADT data are available in only 25 counties, which are denoted by "S" in the "Model" column to indicate that they are included in the state wide model.

4. ADT MODELS

Because counties generally do not convert their traffic counts into AADT, the traffic data used for modeling are average daily traffic (ADT) instead of AADT. Therefore, the model variable is ADT. The use of ADT instead of AADT causes some potential problems, which will be discussed in the conclusions and recommendations.

Four regression models, including a state wide model, a rural model, a small-medium urban area model, and a large urban area model based on data from Broward County, were developed and tested. The efforts began with a somewhat crude model and subsequently attempted to improve the initial model by taking into consideration in increasingly levels of detail of the unique aspects of different areas. The basic approach for all the models is the regression analysis. However, depending on the data availability and data format, the actual data used in different models vary. In this chapter, the first section describes the result of an effort to use data aggregated at the county level to develop a state wide model and a rural model. The remaining two sections present the two models that use TAZ level data. These models include the small-medium urban area model and the Broward model. Each of these models is described in terms of the data set used, the model development, the regression analysis result, and the testing result.

4.1 State Wide Model and Rural County Model

The state wide model and the rural county model were developed by using data aggregated at the county level. The data were obtained from the 1995 *Florida County Profile* published by the Florida Department of Commerce, Division of Economic Development, Bureau of Economic Analysis. The profile was the most recent publication at the time the work was done. The data obtained for all 65 Florida counties are given in **APPENDIX D**.

This effort was motivated by Purdue University's work in which a regression model that covered all counties in the State of Indiana was developed. The objective is to investigate the feasibility of identifying a set of variables that are applicable state wide and to determine how such a model will perform.

4.1.1 State Wide Model

Based on ADT data availability, 25 counties were to be included in the state wide model. They are Alachua, Bay, Bradford, Brevard, Clay, Charllot, Collier, Columbia, Dade, De Soto, Escambia, Gadsden, Highlands, Jackson, Lee, Leon, Manatee, Okaloosa, Osceola, Polk, Walton, St. Johns, Seminole, Sumter, and Sarasota.

The data set used included the 1995 ADT data collected at 118 count stations in the 25 counties and the following six types of data:

(1) Population (*POP*): the total population within a county.
(2) Municipality Population (*MUNICI*): the total population in incorporated areas within a county.
(3) Labor Force (*LABOR*): the total labor force within a county.
(4) Per Capita Income (*INCOME*): the per capita income of a county.
(5) Taxable Sales (*TAXABLE*): the taxable sales of a county.
(6) Lane Miles (*LANEMILE*): the total lane miles of state roads in a county.

Two types of data that are desired but not included due to the effort required to collect the information are area type and accessibility to major highways. Area type information is not directly available for some counties that do not have FSUTMS models. Obtaining the accessibility information will involve determining the county roads on a map and examining the proximity of the count station locations to major highways.

A sample size of 107 out of the 118 is randomly selected to generate the state wide ADT model. **Figure 4.1** shows the results of the multiple regression analysis using all six independent variables. There is a weak relationship between *ADT* and the independent variables. The coefficient of determination R^2 is only 0.2938. In other words, only 29.38% of the variation in ADTs can be explained by these independent variables. The overall *F*-test is significant ($Prob > F = 0.0001$) while some of t statistics for individual variables are insignificant (e.g., *MUNICI*: $Prob > |T| = 0.6916$, etc.). The prob.-values related to some individual t statistics are large, indicating that some of the individual independent variables in the model are not as important as others. The large Variance Inflation values (*VIF*) for *POP*, *MINICI*, *LABOR*, and *TAXABLE* indicate a substantial correlation between these variables.

The results of R^2 and C_p are listed in **Figures 4.2** and **4.3**, respectively. Of the single variable subsets, the total vehicle registrations in a county explains the 15.32% of variation and *POP* (the county population) accounts for 12.64% of variation in ADT, respectively. Adding labor force, per capita income, taxable sales, and state road lane miles increased R^2 by only 5.61%, 6.21%, 1.36% and 0.38%, respectively.

The result of the C_p test is shown in **Figure 4.3**. It indicates that a model with five variables appears to be a good choice. The five variables are the county population (*POP*), county labor force (*LABOR*), county taxable sales (*TAXABLE*), per capita income (*INCOME*), and county registered vehicles number (*VEHICLES*). The regression result using these five variables is given in **Figure 4.4**. The model has the following form:

$$ADT = 9643.704161 + 0.014645 \times POP - 0.155037 \times LABOR - 0.181236 \times INCOME + 0.000005139 \times TAXABLE + 0.058710 \times VEHICLES$$

The R^2 of the regression model with five variables is 0.2890. To detect the existence of multicollinearity, or correlation among independent variables, the variance inflation factors (*VIF*) are computed by $1/(1 - R^2) = 1.4065$. For any independent variable, a *VIF* larger than $1/(1 - R^2) = 1.4065$ indicates that it is more closely related to the other independent variables than it is to the dependent variable. It can be seen that *VIF* for four out of the five variables are greater than 1.4065, indicating a strong multicollinearity between these variables, including population, labor force, taxable sales, and the number of vehicle registration. Further analyses showed that there existed correlation among all variables. Figure 4.5 gives the result of one such analysis in which variables *LABOR, TAXABLE*, and *VEHICLES* have been determined to be strongly correlated and eliminated from the model. The resultant model has only two variables, *POP* and *INCOME*, and has the following form:

$$ADT = 9562.60 + 0.0057 \times POP - 0.1077 \times INCOME$$

```
Dependent Variable: ADT
                                 Analysis of Variance

                                     Sum of         Mean
            Source            DF      Squares        Square       F Value        Prob>F
            Model              7 2150949142.2 307278448.88         5.885         0.0001
           Error              99 5169311415.7 52215266.825
           C Total           106 7320260557.9

               Root MSE     7226.01320      R-square        0.2938
               Dep Mean     9474.03738      Adj R-sq        0.2439
               C.V.           76.27174

                                 Parameter Estimates

                      Parameter       Standard     T for H0:                     Variance
   Variable  DF        Estimate          Error   Parameter=0    Prob > |T|       Inflation

   INTERCEP   1           10287   3825.8816617         2.689        0.0084      0.00000000
   POP        1        0.011203     0.01515533         0.739        0.4615    131.13383141
   MUNICI     1        0.015463     0.03886758         0.398        0.6916    122.67217083
   LABOR      1       -0.155892     0.04872705        -3.199        0.0019    443.40518777
   INCOME     1       -0.158794     0.18867112        -0.842        0.4020      1.34944215
   TAXABLE    1     0.000005151     0.00000167         3.091        0.0026    225.06482237
   LANEMILE   1       -2.279193     3.64384736        -0.625        0.5331      9.63497294
   VEHICLES   1        0.059785     0.02240838         2.668        0.0089     33.87965770
```

Figure 4.1 Initial Results of the Multiple Regression Analysis for the State Wide Model

```
          N = 107      Regression Models for Dependent Variable: ADT

   Number in       R-square          C(p)    Variables in Model
    Model
        1       0.15321510      15.71406     VEHICLES
        1       0.12642940      19.46925     POP
   ---------------------------------------------------
        2       0.20949360       9.82415     LABOR VEHICLES
        2       0.20126046      10.97839     LANEMILE VEHICLES
   ---------------------------------------------------------------
        3       0.27152091       3.12831     LABOR TAXABLE VEHICLES
        3       0.23013087       8.93094     POP LABOR TAXABLE
   ---------------------------------------------------------------
        4       0.28512017       3.22177     MUNICI LABOR TAXABLE VEHICLES
        4       0.28194442       3.66699     POP LABOR TAXABLE VEHICLES
   ---------------------------------------------------------------------
        5       0.28898461       4.68000     POP LABOR INCOME TAXABLE VEHICLES
        5       0.28831451       4.77395     MUNICI LABOR INCOME TAXABLE VEHICLES
   ------------------------------------------------------------------------------
        6       0.29270612       6.15827     POP LABOR INCOME TAXABLE LANEMILE VEHICLES
        6       0.29104436       6.39124     POP MUNICI LABOR INCOME TAXABLE VEHICLES
   ------------------------------------------------------------------------------
        7      0.29383505       8.00000     POP MUNICI LABOR INCOME TAXABLE LANEMILE VEHICLES
   ---------------------------------------------------------------------------------------
```

Figure 4.2 R-square Selection for the State Wide Model

```
   Plot of _CP_*_IN_.  Symbol used is 'C'.
   Plot of _P_*_IN_.   Symbol used is '*'.

     '
     '
     '
  25 ˆ
  24 ˆ
  23 ˆ
  22 ˆ
  21 ˆ
  20 ˆ
M 19 ˆ   C
a 18 ˆ
l 17 ˆ
l 16 ˆ   C
o 15 ˆ
w 14 ˆ
s 13 ˆ
  12 ˆ
C 11 ˆ     C
( 10 ˆ     C
p  9 ˆ       C
)  8 ˆ               C
   7 ˆ             *
   6 ˆ           * C
   5 ˆ         * C
   4 ˆ       * C
   3 ˆ     * C C
   2 ˆ   *
   1 ˆ
   0 ˆ
     '
     Šƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒƒƒƒƒƒƒ
       0 1 2 3 4 5 6 7 8 9 1 1 1 1 1 1
                           0 1 2 3 4 5

          Number of regressors in model

              NOTE: 9 obs hidden.
```

Figure 4.3 C_p Test for the State Wide Model

The R^2 is 0.1295, and the adjusted R^2 is 0.1128. The VIF values are slight greater than $1/(1 - R^2) =$ 1.017 indicating there is a low degree of correlation between population and per capita income. This multicollinearity is also suggested by the negative sign of *INCOME* in the equation.

Model: MODEL1
Dependent Variable: ADT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	5	2115442631.5	423088526.29	8.210	0.0001
Error	101	5204817926.4	51532850.756		
C Total	106	7320260557.9			

Root MSE	7178.63850	R-square	0.2890
Dep Mean	9474.03738	Adj R-sq	0.2538
C.V.	75.77169		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|	Variance Inflation
INTERCEP	1	9643.704161	3304.8625584	2.918	0.0043	0.00000000
POP	1	0.014645	0.01063194	1.377	0.1714	65.39160430
LABOR	1	-0.155037	0.04234590	-3.661	0.0004	339.31005524
INCOME	1	-0.181236	0.18123044	-1.000	0.3197	1.26159229
TAXABLE	1	0.000005139	0.00000154	3.342	0.0012	194.25885019
VEHICLES	1	0.058710	0.02211561	2.655	0.0092	33.43715176

Collinearity Diagnostics(intercept adjusted)

Number	Eigenvalue	Condition Index	Var Prop POP	Var Prop LABOR	Var Prop INCOME	Var Prop TAXABLE	Var Prop VEHICLES
1	3.86520	1.00000	0.0010	0.0002	0.0035	0.0003	0.0019
2	0.96564	2.00068	0.0002	0.0000	0.7915	0.0000	0.0006
3	0.15403	5.00938	0.0144	0.0025	0.0948	0.0128	0.0631
4	0.01332	17.03751	0.6351	0.0007	0.0002	0.0111	0.9322
5	0.00182	46.13939	0.3493	0.9966	0.1100	0.9758	0.0022

Figure 4.4 Results of the Multiple Regression Analysis Using Five Variables for the State Wide Model

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	947945594.25	473972797.12	7.736	0.0007
Error	104	6372314963.6	61272259.265		
C Total	106	7320260557.9			

Root MSE	7827.65988	R-square	0.1295
Dep Mean	9474.03738	Adj R-sq	0.1128
C.V.	82.62222		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|	Variance Inflation
INTERCEP	1	9562.596915	3403.1481512	2.810	0.0059	0.00000000
POP	1	0.005704	0.00145030	3.933	0.0002	1.02336734
INCOME	1	-0.107733	0.17798248	-0.605	0.5463	1.02336734

Figure 4.5 Multiple Regression Analysis Using Two Variables for the State Wide Model

The inadequacy of the model is confirmed by testing the model using 11 data points randomly selected from the 118 samples and reserved for this purpose. **Table 4.1** summarizes the results. For each observation, the table indicates the observed ADT, predicted ADT, and percentage error in the estimation. The percentage errors range from 8% to 1096%, with an average difference of 188%.

The result of this modeling effort is not unexpected since the areas being modeled have very different characteristics, as they range from large metropolitan counties to rural counties. While adding more variables such as area type and accessibility to highways to the model may improve the model, the improvement may not be significant. Additionally, county level data do not sufficiently explain variations in ADDT within a county. Therefore, inclusion of counties that have a sophisticated roadway network and complex travel patterns will degrade the model.

Table 4.1 Model Validation for the State Wide Model

Municipal Population	*Labor Force*	*Per Capita Inc.*	*Taxable Sales*	*Lane Miles*	*Vehicles*	*ADT*	*R-ADT*	*% of Error*
74,153	74,121	16,852	1,698,365,000	503	48,965	3,500	8,790	151
11,873	40,841	18,012	1,091,957,000	312	34,344	10,500	9,570	-8
20,958	91,390	29,237	2,795,615,000	821	56,110	29,999	10,493	-65
753,360	1,080,823	19,266	21,427,878,000	2,524	690,359	9,500	18,982	99
62,173	142,882	16,899	2,687,570,000	809	80,011	19,600	6,947	-64
13,677	17,347	14,949	283,646,000	632	14,170	600	7,177	1096
133,731	142,214	18,746	2,284,439,000	563	60,420	14,579	2,617	-82
60,855	88,947	18,202	1,610,518,000	646	52,219	1,650	6,294	281
135,899	194,504	16,858	4,157,759,000	1,379	133,501	13,500	12,033	-10
8,559	8,746	13,955	174,546,000	369	7,926	2,489	7,630	206
136,957	130,984	20,846	3,333,858,000	411	105,100	14,948	13,694	-8

4.1.2 Rural Area Model

It can be seen that the state wide model cannot adequately account for the underlying causes that contribute to ADT. The model may be improved by reducing the variation in the characteristics of the counties considered in the model. It is expected that with the limited data available, a data set that is more homogeneous may produce better results. Based on this expectation, a threshold value of 100,000 for county population and availability of traffic counts were used as the criteria to choose a subset of rural counties from the state wide model. Eight counties were selected (see **Table 4.2**).

With a total of 30 count stations, 27 data points were randomly selected to generate the rural ADT model and the rest were used to test the model. As for the state wide model, the following six variables were considered initially:

(1) Population (*POP*): the total population within a county;
(2) Municipalities Population (*MUNICI*): the total municipalities population within a county;
(3) Labor Force (*LABOR*): the total labor force within a county;
(4) Per Capita Income (*INCOME*): the figure of per capita income of a county;
(5) Taxable Sales (*TAXABLE*): the taxable sales of a county; and
(6) Lane Miles (*LANEMILE*): the total lane miles of county roads in a county.

Table 4.2 Counties Included in the Rural Area Model

County	*Population*
Bradford	24,182
De Soto	25,048
Walton	33,615
Sumter	34,788
Gadsden	43,378
Jackson	43,891
Columbia	48,376
Highlands	74,507

The results of the multiple regression analysis using all six independent variables are summarized in **Figure 4.5**. There is a fair relationship between ADT and the independent variables. The coefficient of determination R^2 is 0.4525. In other words, 45.25% of the sum of error squares in ADT can be associated with the variation of these independent variables. This is a noticeable improvement over the state wide model. The overall F-test is significant ($Prob{>}F = 0.0765$) while some t statistics for individual variables are insignificant (e.g. TAXABLE: $Prob{>}|T| = 0.9608$, etc.). The prob.-values related to some individual t statistics are large, indicating that some of the individual independent variables in the model are not as important as others. The large *VIF* values for *TAXABLE*, *LABOR*, *POP* and *MUNICI* indicate a substantial correlation between these variables.

The result of R^2 seelction, shown in **Figures 4.6**, indicates that the state road lane-miles, county total population, and county labor force were the most significant variables. The result of the C_p test, shown in **Figure 4.7**, indicates that there is a definite corner at $p = 3$, where the C_p values increase rapidly with smaller subset sizes. Hence, a model with four variables appears to be a good choice. The four variables are county population (*POP*), labor force (*LABOR*), state roads lane miles (*LANEMILE*), and number of vehicle registrations (*VEHICLES*).

The new regression model using the above four variables has the following form:

$$ADT = 4853.489444 + 0.122587 \times POP + 0.261858 \times LABOR - 18.930235 \times LANEMILE - 0.0032338 \times VEHICLES$$

As given in **Figure 4.8**, the R^2 of the regression model with 4 variables is 0.4488, and $1/(1 - R^2) =$ 1.842. It can be seen that variables *POP*, *LABOR*, and *LANEMILE* still have *VIF* values greater than 1.842, indicating significant multicollinearity among these variables. More analyses showed that all variables are correlated as in the case of the state wide model.

Model: MODEL1
Dependent Variable: ADT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	7	140954357.9	20136336.843	2.244	0.0765
Error	19	170522136.62	8974849.2956		
C Total	26	311476494.52			

Root MSE	2995.80528	R-square	0.4525
Dep Mean	4431.59259	Adj R-sq	0.2508
C.V.	67.60110		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|	Variance Inflation
INTERCEP	1	-520.050735	17262.912118	-0.030	0.9763	0.00000000
POP	1	0.107277	0.31766789	0.338	0.7393	38.57028564
MUNICI	1	0.089906	1.01473854	0.089	0.9303	34.19457031
LABOR	1	0.151413	0.69186075	0.219	0.8291	36.20283409
INCOME	1	0.410106	1.32043788	0.311	0.7595	3.10755396
TAXABLE	1	0.000001515	0.00003040	0.050	0.9608	41.69901013
LANEMILE	1	-18.368434	5.89999293	-3.113	0.0057	3.19301467
VEHICLES	1	-0.031489	0.06715864	-0.469	0.6445	2.60807982

Figure 4.6 Initial Results of the Multiple Regression Analysis for the Rural Area Model

N = 27 Regression Models for Dependent Variable: ADT

Number in Model	R-square	C(p)	Variables in Model
1	0.18953670	5.12752	LANEMILE
1	0.04083525	10.28828	INCOME
2	0.42186163	-0.93543	POP LANEMILE
2	0.41942416	-0.85083	LABOR LANEMILE
3	0.44169907	0.37610	POP INCOME LANEMILE
3	0.43435654	0.63093	POP LABOR LANEMILE
4	0.44879671	2.12978	POP LABOR LANEMILE VEHICLES
4	0.44758634	2.17178	POP INCOME LANEMILE VEHICLES
5	0.45217040	4.01269	POP LABOR INCOME LANEMILE VEHICLES
5	0.44975646	4.09647	POP MUNICI LABOR LANEMILE VEHICLES
6	0.45246455	6.00248	POP MUNICI LABOR INCOME LANEMILE VEHICLES
6	0.45230991	6.00785	POP LABOR INCOME TAXABLE LANEMILE VEHICLES
7	0.45253610	8.00000	POP MUNICI LABOR INCOME TAXABLE LANEMILE VEHICLES

Figure 4.7 R-square Selection for the Rural Area Model

```
  Plot of _CP_*_IN_.  Symbol used is 'C'.
  Plot of _P_*_IN_.   Symbol used is '*'.

     ‚
     ‚
     ‚
  25 ˆ
  24 ˆ
  23 ˆ
  22 ˆ
  21 ˆ
  20 ˆ
M 19 ˆ   C
a 18 ˆ
l 17 ˆ
l 16 ˆ   C
o 15 ˆ
w 14 ˆ
s 13 ˆ
  12 ˆ
C 11 ˆ     C
( 10 ˆ     C
p  9 ˆ       C
)  8 ˆ               C
   7 ˆ             *
   6 ˆ           * C
   5 ˆ         * C
   4 ˆ       * C
   3 ˆ     * C C
   2 ˆ   *
   1 ˆ
   0 ˆ
     ‚
     Šƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒƒƒƒƒƒƒƒ
       0 1 2 3 4 5 6 7 8 9 1 1 1 1 1 1
                           0 1 2 3 4 5
           Number of regressors in model
```

NOTE: 9 obs hidden.

Figure 4.8 C_p Test for the Rural Area Model

```
Model: MODEL1
Dependent Variable: ADT

                          Analysis of Variance

                                Sum of           Mean
         Source          DF        Squares          Square       F Value       Prob>F

         Model            4 139789627.12    34947406.78         4.478       0.0085
         Error           22  171686867.4    7803948.518
         C Total         26 311476494.52

             Root MSE    2793.55482      R-square       0.4488
             Dep Mean    4431.59259      Adj R-sq       0.3486
             C.V.          63.03727

                          Parameter Estimates

                     Parameter       Standard     T for H0:                   Variance
Variable  DF          Estimate          Error    Parameter=0    Prob > |T|     Inflation

INTERCEP   1       4853.489444   2142.5951644          2.265        0.0337    0.00000000
POP        1          0.122587     0.14324455          0.856        0.4013    9.01934915
LABOR      1          0.261858     0.34402186          0.761        0.4546   10.29413919
LANEMILE   1        -18.930235     4.57896698         -4.134        0.0004    2.21179934
VEHICLES   1         -0.032338     0.04259670         -0.759        0.4558    1.20665132

               Collinearity Diagnostics(intercept adjusted)

                              Condition  Var Prop   Var Prop   Var Prop   Var Prop
      Number  Eigenvalue         Index  POP        LABOR      LANEMILE   VEHICLES

           1     2.75307       1.00000    0.0132     0.0118     0.0404     0.0274
           2     0.86143       1.78772    0.0004     0.0011     0.0741     0.8134
           3     0.33050       2.88619    0.0725     0.0337     0.8286     0.1591
           4     0.05500       7.07508    0.9139     0.9533     0.0569     0.0001
```

Figure 4.9 Results of the Multiple Regression Analysis Using Four Variables for the Rural Area Model

There are three testing data points for examining the model's predictive capability. **Table 4.3** summarized the results. For each observation, the table indicates the observed ADTs, predicted ADTs, and percentage error in the estimation. The percent difference between these two ADT values range from 6.87% to 83.96%, with an average difference of 35.59%. While the size of the test data set is rather small, the test results do indicate smaller errors compared to the state wide model. This, along with the increase in R^2 in the rural area model, suggests that including counties that have similar characteristics in a model will improve model performance. Still, the variable set used is not sufficiently inclusive. To further improve the model, more variables need to be identified and included.

Table 4.3 Model Validation for the Rural Area Model

County	Population	Municipal Population	Labor Force	Per Capita Income	Taxable Sales	Lane Mileage	Vehicles	ADT	R-ADT	% of Error
Walton	33615	6872	11177	14128	316743000	489	5769	2300	2458	6.87
Jackson	43891	13677	17347	14949	283646000	632	14170	2800	2354	-15.93
Highlands	74507	18367	26236	16541	571296000	382	23149	7000	12877	83.96

4.3 Small-Medium Urban Area Model

Urban areas usually have a travel demand and pattern quite different from those in rural areas, therefore need to be considered separately. The size of an urban area and its characteristics may also have a significant impact on travel demand and travel patterns. For the same reason that the rural area model performs better than the state wide model, models that consider an individual group of urban areas with similar characteristics will certainly be more accurate than a single model that includes large, medium, and small urban areas. Therefore, a model that include counties that are neither rural nor major urban counties, like Miami-Dade or Broward counties, was developed.

To model ADT for small to medium urban areas, a criterion of county population between 100,000 and 400,000 is used to determine the counties to be included in the model. As a result, counties with major cities are excluded, such as Dade County (Miami), Broward County (Ft. Lauderdale), Palm Beach County, Orange County (Orlando), Duval County (Jacksonville), and Pinellas County (Pensacola).

4.3.1 Data Collection and processing

The 1995 ADT data are available in 10 counties including Charlotte, Bay, Okaloosa, Leon, Marion, Escambia, Volusia, and Polk counties. With a total of 270 count stations, a sample of 243 is randomly selected to generate the small-medium urban ADT model. Unlike for the state wide model and the rural area model, 14 variables are considered initially:

(1) *DU_SF*: total single family dwelling units in a TAZ.
(2) *POP_SF*: single family population in a TAZ.
(3) *SAUTO*: total single family automobile ownership in a TAZ.
(4) *DU_MF*: total multi-family dwelling units in a TAZ.
(5) *Pop_MF*: multi-family population in the TAZ.
(6) *MAUTO*: total multi-family automobile ownership in a TAZ.
(7) *HOT_OCC*: population in hotel/motels in a TAZ.
(8) *IND_EMP*: industrial employment in a TAZ.
(9) *COM_EMP*: commercial employment in a TAZ.
(10) *SER_EMP*: service employment in a TAZ.
(11) *SCH_ENR*: school enrollment in a TAZ.
(12) *LANES*: number of lanes at the count station location in two direction.
(13) *ATYPE*: area type of the count station location.
(14) *FTYPE*: facility type of the road located the count station.

All of these 14 variables can be obtained directly or indirectly from the ZDATA1 and ZDATA2 files in FSUTMS. The data extraction is performed manually, which is a tedious and labor intensive process. However, these data may be easily obtained when applying the model to individual locations. The socioeconomic data have been aggregated by TAZ. Data processing is described below.

The numbers of occupied single- and multi-family dwelling units (*DU_SF* and *DU_MF*), respectively, are calculated as follows:

$$DU\_SF = SFDU\_Total \times (1 - vacant\_SF/100)$$
$$DU\_MF = MFDU\_Total \times (1 - vacant\_MF/100)$$

where *vacant_SF* and *vacant_MF* are numbers representing vacancy rates of single- and multifamily dwelling units, respectively.

In the ZDATA files, automobile ownership is given as percentages of dwelling units with no vehicle, one vehicle or more than one vehicle. The total vehicle ownership is calculated assuming an average automobile ownership of 2.3 vehicles per single-family dwelling unit and 2.2 vehicles per multi-family dwelling unit for those that have more than two vehicles:

$$SAUTO = 1 \times DU\_SF \times Veh\_1 / 100 + 2.3 \times DU\_SF \times Veh\_2 / 100$$
$$MAUTO = 1 \times DU\_MF \times Veh\_1 / 100 + 2.2 \times DU\_MF \times Veh\_2 / 100$$

Although this study these factors (2.2 and 2.3) are assumed, if the automobile ownership is determined to be a significant factor, better estimates may be obtained by analyzing census data. At the same time, it is unlikely that inaccuracy in the automobile ownership estimates for dwelling units with more than one vehicle will be so significant as to affect the inclusion or exclusion of this factor in the model.

The values for socioeconomic variables (*DU_SF*, *POP_SF*, *SAUTO*, *DU_MF*, *Pop_MF*, *MAUTO*, *HOT_OCC*, *IND_EMP*, *COM_EMP*, *SER_EMP*, *SCH_ENR*) are obtained by summing the values of each variable from the TAZs adjacent to the location where a traffic count is collected. This method is simple but has some potential problems associated with the nonuniform sizes of TAZs and the nonuniform sizes of roadway segments.

The variable *LANES* is the number of lanes in both directions on a roadway. The off-system roads may have 2, 4, or 6 lanes.

There are five area types, or land use types in FSUTMS databases: Central Business District, fringe area, residential area, outlying business district, and rural area. The meanings of these classification are as follows:

- Central Business District (CBD): CBD is an area where the predominant land use is intense business activities. CBD is characterized by large numbers of pedestrians, commercial vehicles, loading and unloading of goods and people, a large demand for parking spaces, and a high degree of turnover in parking.
- Fringe Area: A fringe area is the portion of a municipality immediately outside the CBD. This kind of area exhibits a wide range of business activities (small businesses, light industry,

warehousing, automobile service centers, and intermediate strip development with some concentrated residential areas).

- Residential Area: A residential area is an area within the influence of a municipality in which the predominant land use is residential development (small businesses may be present). It is characterized by few pedestrians and low parking turnover.
- Outlying Business District (OBD): An OBD is an area within the influence of a municipality that is normally separated by some distance from the CBD and its fringe area, but that has the intense activity characteristics of a central area. The principal land use is business, and there may be heavy traffic or through movements, causing vehicles to operate at lower speeds than in fringe areas. Also characterized by large demand for parking and high turnover, and moderate pedestrian traffic. This category does not include off-street shopping on only one side of a street. An area with moderate to heavy strip development on both sides of a street should be coded OBD.
- Rural area: A rural area is a sparsely developed area within the influence of a municipality in which the predominant land use is other than those described in the four preceding categories.

Area types are coded as integers. Because the coding is ordinal, new values are assigned to the area type to reflect their possible association with ADT. The new numerical values, or the weights, are determined based on the result of a single variable regression in which ADT is related to area types. These new values assigned to area types (variable *ATYPE1*) are shown in **Table 4.4**.

Table 4.4 Area Type Coding in Small-Medium Urban Area Model

Area Type	*Original Value (ATYPE)*	*Reassigned Value (ATYPE1)*
Central Business District (CBD)	1	1
Fringe Area	2	5
Residential Area	3	3
Outlying Business District (OBD)	4	4
Rural Area	5	2

Similarly, there are four facility types used in FSUTMS database, which are also coded numerically. New values for different facility type are assigned (to a new variable *FTYPE1*) after a study of the regression relationship between *FTYPE* and ADT. These reassignments of values for different facility types are given in **Table 4.5**.

Table 4.5 Facility Type Coding in Small-Medium Urban Area Model

Facility Type	*Original Value (FTYPE)*	*Reassigned Value (FTYPE1)*
Divided Arterial	20	6
Undivided Arterial	30	2
Collector	40	1
Centroid Connector	50	1

4.3.2 Model Development

The results of the multiple regression analysis using all 14 independent variables are summarized in **Figure 4.9**. There is a strong relationship between ADT and the independent variables. The coefficient of determination R^2 is 0.7433. In other words, up to 74.33% of the sum of error squares in ADT may be associated with the variation in these independent variables. The overall *F-test* is significant (*Prob>F* = 0.0001) while some of t statistics for individual variables are insignificant (e.g. *DU_SF*: *Prob>*$|T|$ = 0.9313 and *DU_MF*: *Prob>*$|T|$ = 0.9416, etc.). The prob.-values related to some individual t statistics are large, indicating that some of the individual independent variables such as *DU_SF, POP_SF, SAUTO, DU_MF, POP_MF, SER_EMP* and *SCH_ENR* in the model are not as important as others.

The results of R^2 selection and C_p test are given in **Figures 4.10** and **4.11**, respectively. From **Figure 4.10**, it may be observed that of the single variable subsets, the number of lanes explains 63.53% of the variation in ADT. The second most significant variable, *ATYPE1* (the area type of a road) accounted for 26.75% of the variation in ADT.

Figure 4.11 reveals that there appears a corner at $p = 5$, where the C_p values increase rapidly with smaller subset sizes. Hence, a model with five variables appears to be a good choice. The five variables are *LANES, FTYPE1, ATYPE1, HOT_OCC,* and *DU_SF.* However, further investigations revealed that there exists multicollinearity between *DU_SF* and *HOT_OCC* and between *LANES* and *FTYPE1*. Therefore, *DU_SF* and *FTYPE1* are eliminated and one more variable, *COM_EMP*, from the next best subset of variables (subset 7 in **Figure 4.10**) is added, resulting in a four-variable subset: *LANES, ATYPE1, HOT_OCC,* and *COM_EMP*.

The result of the regression with four variables is shown in **Figure 4.12**. The R^2 is 0.7251, and the adjusted R^2 is 0.7206. It can be seen that the *VIF* value for all variables is smaller than $1/(1- R^2)$ = 2.1087, indicating that there is no multicollinearity between these variables. The model has the following form:

$$ADT = -13418 + 6770.23 \times LANES + 1580.14 \times ATYPE1 + 2.85 \times COM\_EMP + 1.78 \times HOT\_OCC$$

The testing data set consists of 27 testing data points and is used to examine the model's predictive capability. **Table 4.6** summarizes the results. For each observation, the table indicates the observed ADTs (ADT), predicted ADTs (R-ADT), and percentage error in the estimation. The percent difference between these two ADT values range from 0.68% to 56.44%, with an average difference of 27.22%.

Model: MODEL1
Dependent Variable: ADT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	14	26087158585	1863368470.4	48.030	0.0001
Error	231	8961894855	38796081.623		
C Total	245	35049053440			

Root MSE	6228.65006	R-square	0.7443
Dep Mean	11706.21951	Adj R-sq	0.7288
C.V.	53.20804		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-11572	1710.9243271	-6.763	0.0001
DU_SF	1	-0.387318	4.48503079	-0.086	0.9313
POP_SF	1	-0.816062	1.61883523	-0.504	0.6147
SAUTO	1	0.182037	0.69346454	0.263	0.7932
DU_MF	1	0.231926	3.27637593	0.071	0.9436
POP_MF	1	0.649053	1.76930086	0.367	0.7141
MAUTO	1	-0.761665	0.77941020	-0.977	0.3295
HOT_OCC	1	2.126634	0.38884759	5.469	0.0001
IND_EMP	1	1.685168	1.25402052	1.344	0.1803
COM_EMP	1	2.154274	1.45850219	1.477	0.1410
SER_EMP	1	0.438896	0.90642870	0.484	0.6287
SCH_ENR	1	-0.178522	0.32317076	-0.552	0.5812
LANES	1	7604.067931	503.57073086	15.100	0.0001
ATYPE1	1	1021.329690	538.78305861	1.896	0.0593
FTYPE1	1	-735.318916	371.23639942	-1.981	0.0488

Figure 4.10 Results of the Multiple Regression Analysis for the Small-Medium Urban Area Model with 14 Initial Variables

Number in Model	R-square	C(p)	Variables in Model
1	0.63538898	87.39592	LANES
1	0.26752473	419.73087	ATYPE1
2	0.70965864	22.29942	HOT_OCC LANES
2	0.66263737	64.77925	LANES ATYPE1
3	0.72072835	14.29885	DU_SF HOT_OCC LANES
3	0.72048047	14.52279	HOT_OCC LANES ATYPE1
4	0.72871509	9.08350	DU_SF HOT_OCC LANES ATYPE1
4	0.72835634	9.40760	DU_SF HOT_OCC LANES FTYPE1
5	0.73531937	5.11707	DU_SF HOT_OCC LANES ATYPE1 FTYPE1
5	0.73530514	5.12993	POP_SF HOT_OCC LANES ATYPE1 FTYPE1
6	0.73921712	3.59578	POP_SF HOT_OCC COM_EMP LANES ATYPE1 FTYPE1
6	0.73894572	3.84097	DU_SF HOT_OCC COM_EMP LANES ATYPE1 FTYPE1
7	0.74160009	3.44297	POP_SF HOT_OCC IND_EMP COM_EMP LANES ATYPE1 FTYPE1
7	0.74140946	3.61518	DU_SF HOT_OCC IND_EMP COM_EMP LANES ATYPE1 FTYPE1
8	0.74332413	3.88544	POP_SF MAUTO HOT_OCC IND_EMP COM_EMP LANES ATYPE1 FTYPE1
8	0.74294931	4.22405	DU_SF MAUTO HOT_OCC IND_EMP COM_EMP LANES ATYPE1 FTYPE1
9	0.74368654	5.55803	POP_SF POP_MF MAUTO HOT_OCC IND_EMP COM_EMP LANES ATYPE1 FTYPE1
9	0.74356505	5.66778	POP_SF MAUTO HOT_OCC IND_EMP COM_EMP SCH_ENR LANES ATYPE1 FTYPE1
10	0.74395089	7.31921	POP_SF POP_MF MAUTO HOT_OCC IND_EMP COM_EMP SCH_ENR LANES ATYPE1 FTYPE1
10	0.74384359	7.41614	POP_SF POP_MF MAUTO HOT_OCC IND_EMP COM_EMP SER_EMP LANES ATYPE1 FTYPE1
11	0.74422750	9.06931	POP_SF POP_MF MAUTO HOT_OCC IND_EMP COM_EMP SER_EMP SCH_ENR LANES ATYPE1 FTYPE1
11	0.74405743	9.22296	POP_SF DU_MF MAUTO HOT_OCC IND_EMP COM_EMP SER_EMP SCH_ENR LANES ATYPE1 FTYPE1
12	0.74429333	11.00984	POP_SF SAUTO POP_MF MAUTO HOT_OCC IND_EMP COM_EMP SER_EMP SCH_ENR LANES ATYPE1 FTYPE1
12	0.74422788	11.06897	DU_SF POP_SF POP_MF MAUTO HOT_OCC IND_EMP COM_EMP SER_EMP SCH_ENR LANES ATYPE1 FTYPE1
13	0.74429868	13.00501	DU_SF POP_SF SAUTO POP_MF MAUTO HOT_OCC IND_EMP COM_EMP SER_EMP SCH_ENR LANES ATYPE1 FTYPE1
13	0.74429597	13.00746	POP_SF SAUTO DU_MF POP_MF MAUTO HOT_OCC IND_EMP COM_EMP SER_EMP SCH_ENR LANES ATYPE1 FTYPE1
14	0.74430423	15.00000	DU_SF POP_SF SAUTO DU_MF POP_MF MAUTO HOT_OCC IND_EMP COM_EMP SER_EMP SCH_ENR LANES ATYPE1 FTYPE1

Figure 4.11 R-square Selection for the Small-Medium Urban Area Model

```
  Plot of _CP_*_IN_.  Symbol used is 'C'.
  Plot of _P_*_IN_.   Symbol used is '*'.
  Plot of _CP_*_IN_.  Symbol used is 'C'.
  Plot of _P_*_IN_.   Symbol used is '*'.
     ‚
     ‚
     ‚
  25 ˆ
  24 ˆ
  23 ˆ
  22 ˆ     C
  21 ˆ
  20 ˆ
M 19 ˆ
a 18 ˆ
l 17 ˆ
l 16 ˆ
o 15 ˆ       C                     C
w 14 ˆ       C                   *
s 13 ˆ                         * C
  12 ˆ                       *
C 11 ˆ                     *   C
( 10 ˆ                   *
p  9 ˆ         C       *     C
)  8 ˆ               *
   7 ˆ             *       C
   6 ˆ           *       C
   5 ˆ         * C
   4 ˆ       *     C C C
   3 ˆ     *         C
   2 ˆ   *
   1 ˆ
   0 ˆ
     ‚
     Šƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒƒƒƒƒƒƒƒ
       0 1 2 3 4 5 6 7 8 9 1 1 1 1 1 1
                           0 1 2 3 4 5

           Number of regressors in model

     NOTE: 23 obs hidden.  3 obs out of range.
```

Figure 4.12 C_p Test for the Small-Medium Urban Area Model

```
                         Dependent Variable: ADT
                         Analysis of Variance
                          Analysis of Variance

                              Sum of          Mean
Source              DF       Squares        Square      F Value       Prob>F
Model                4   25474601647  6368650411.7      160.868       0.0001
Error              244  9659793557.1  39589317.857
C Total            248   35134395204

    Root MSE       6292.00428      R-square        0.7251
    Dep Mean      11686.99197      Adj R-sq        0.7206
    C.V.             53.83767

                          Parameter Estimates

                 Parameter       Standard    T for H0:                  Variance
Variable  DF      Estimate          Error  Parameter=0   Prob > |T|    Inflation
INTERCEP   1        -13418   1544.7280683       -8.686       0.0001   0.00000000
HOT_OCC    1      1.784224     0.25257557        7.064       0.0001   1.10844245
COM_EMP    1      2.851354     1.34518093        2.120       0.0350   1.11132078
LANES      1   6770.231853   386.62544687       17.511       0.0001   1.29503563
ATYPE1     1   1580.141839   490.26022163        3.223       0.0014   1.31503884
```

Figure 4.13 Final Multiple Regression Analysis with Four Variables for the Small-Medium Urban Area Model

Table 4.6 Model Validation for the Small-Medium Urban Area Model

HOT-OCC	*COM-EMP*	*LANES*	*ATYPE*	*ATYPE1*	*ADT*	*R-ADT*	*% of ERROR*
0	47	2	30	3	3884	4997	28.66%
0	45	2	30	3	4100	4991	21.73%
0	0	2	30	3	4775	4863	1.84%
0	13	2	50	2	5300	3320	-37.36%
0	34	2	30	3	5800	4960	-14.48%
0	387	2	10	1	6441	2806	-56.44%
0	465	2	30	3	7068	6189	-12.44%
0	450	2	50	2	7500	4566	-39.12%
0	17	2	30	3	8059	4911	-39.06%
0	64	2	30	3	8400	5045	-39.94%
36	649	2	10	1	8885	3617	-59.29%
0	180	2	30	3	9702	5376	-44.59%
0	336	2	30	3	10267	5821	-43.30%
1966	358	2	40	4	10920	10972	0.48%
0	7	4	30	3	15194	18423	21.25%
0	108	4	30	3	17800	18711	5.12%
60	0	4	20	5	15893	21671	36.36%
60	112	4	20	5	16582	21990	32.61%
40	20	2	20	5	5698	8152	43.07%
0	21	4	40	4	21973	20043	-8.78%
0	89	4	30	3	22500	18657	-17.08%
0	24	4	40	4	22733	20052	-11.79%
0	457	4	20	5	23275	22867	-1.75%
0	3	4	20	5	29703	21572	-27.37%
908	17	4	20	5	33738	23232	-31.14%
0	23	6	20	5	43379	35170	-18.92%
764	56	4	20	5	51529	23086	-55.20%

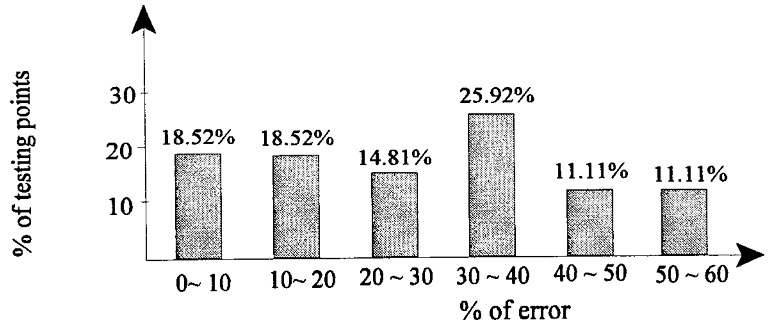

Figure 4.14 Error Distribution in Test Data for Small-Medium Urban Model

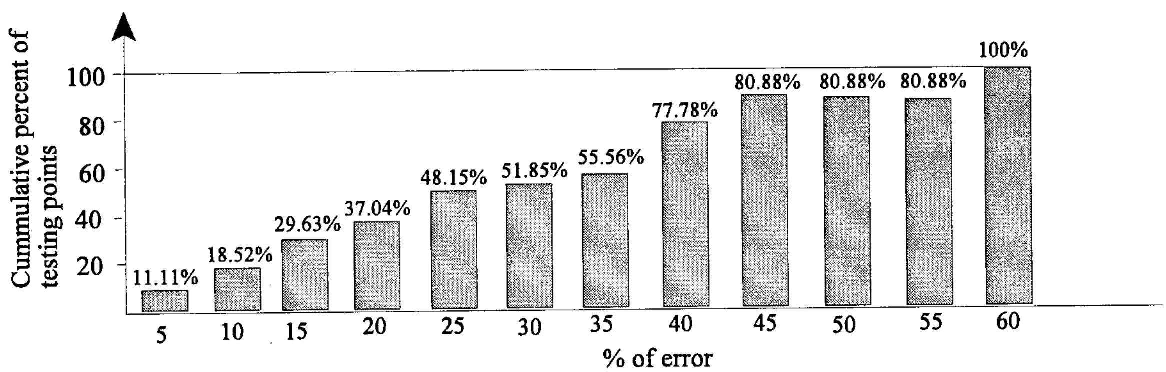

Figure 4.15 Cumulative Percent of Testing Points in Given Error Range for Small-Medium Urban Area Model

4.4 Large Metropolitan Area Model - Broward County Model

The transportation systems in large urban areas consists of many different types of facilities and are located in many different land uses. They also exhibit complex travel patterns. There are several large urban areas in Florida, including Miami-Dade, Broward, Palm Beach, and Tampa, each with a population close to or exceeding one million. In this section an investigation on the development of a regression model for a large urban area, namely Broward County, is reported. Broward County has been chosen because of the excellent availability of data in digital format, which reduces efforts required for data processing and provides an excellent opportunity to perform more sophisticated analysis using geographic information systems (GIS). Although GIS is used in this analysis, it is only used for the purpose of investigating the spatial effects of land uses and accessibility on ADT. The objective is to identify any spatial patterns of land uses or specific network configurations that appear to be associated with certain ADT patterns. If such associations are found, methods will need to be devised to obtain data without the use of GIS for actual application of the model.

4.4.1 Data Collection and Processing

ADTs are obtained from average quarterly traffic counts in 1997 from Broward County. Data from 443 count stations are used for model development. Most of these count stations are on off-system roads. About 60 count stations that are on state minor arterial roads are also used because of the similar characteristics of state and county minor arterial roads. Of the total data, 90% (399 count stations) is used for model development, and 10% of the data (44 count stations) is used for model validation. These two sets of data are selected randomly from the ADT database.

Three kinds of independent variables are included to estimate ADT. The first group is the roadway characteristics, which includes the number of lanes, functional classifications, facility type, area types, etc. The second group of variables is the socioeconomic characteristics of the area surrounding a count station. The socioeconomic data are available at the TAZ level and are aggregated for each count station using GIS. The last group of data reflects road network connectivity and is a measure of the accessibility to a site (a count station location) on one road to other roads in the network. The initial predictors in the three groups are described below in more detail.

Roadway Characteristics Data

This set of data describes the characteristics of the roadway sections. The data include:

(1) Number of lanes (*L*): the number of lanes on a roadway;
(2) Area type (*AREA1*): land use type may be one of the following: rural area, central business, district (CBD), fringe area, residential area, and outlying business district (OBD);
(3) Functional classification (*FCLASS1*): four functional classifications in Broward County are used: state minor arterial, county minor arterial, county collector, city collector, and local and unclassified;
(4) Facility type (*FAC1*):

Area type and functional classification are all of ordinal values. Their values are reassigned based on single variable analysis to identify the relationship between each of them with ADT. The reassigned values for each variable are given in **Tables 4.7**and **4.8.**

Table 4.7 Area Type Recoding for the Broward County Model

Area Type	*Original Value (ATYPE)*	*Reassigned Value (ATYPE1)*
Rural Area	5	1
Central Business District	1	2
Fringe Area	2	2
Residential Area	3	3
Outlying Business District (OBD)	4	4

Table 4.8 Recoded Functional Classifications in the Broward County Model

Facility Type	*Original Value (FTYPE)*	*Reassigned Value (FTYPE1)*
State Minor Arterial	6	2
County Minor Arterial	2	2
County Collector	3	1
City Collector	5	1
Local and Unclassified	0 and 4	0

The functional classification system used by Broward County includes three additional function classes: state principal arterial (expressways), state principal arterial, and state ramps. These three classes are not considered since data from count stations on state principal arterial roads are not included in the model due to the significant differences between such roads and off-system roads. The data for the above predictors are provided by Broward County Metropolitan Planning Organization (MPO) as a database file that is an attribute table of the roadway GIS layer.

Socioeconomic data

These variables reflect the socioeconomic characteristics surrounding a count station. Variables are chosen as independent variables with the expectation that land use patterns and auto ownership will influence ADT to some degree. Eleven variables are considered initially:

(1) Population (*POP*): the total population within a certain distance of a count station;
(2) Single-family population (*SFPOP*): the total single-family population within a certain distance of a count station;
(3) Multi-family population (*MFPOP*): the total single-family population within a certain distance of a count station;
(4) Single-family dwelling units (*SFDUS*): the total occupied single-family housing units within a certain distance of a count station;
(5) Multi-family dwelling units (*MFDUS*): the total occupied multi-family housing units within a certain distance of a count station;
(6) Auto ownership (*AUTO*): the estimated total number of automobiles within a certain distance of a count station;
(7) Industrial Employment (*INDEMP*): the total industrial employment number within a certain distance of a count station;
(8) Commercial Employment (*COMMEMP*): the total commercial employment number within a certain distance of a count station;
(9) Service Employment (*SEREMP*): the total service employment number within a certain distance of a count station;

(10) School enrollment (*SCHOOL*): the total school enrollment number within a certain distance of a count station;
(11) Hotel occupancy (*HTL*): the total hotel occupants within a certain distance of a count station.

The data for the above predictors are obtained from Broward County MPO in a database format. A GIS layer of the TAZ structure is also available. To associate the socioeconomic conditions in TAZs to count stations, data are aggregated using a simple buffer zone method. A program is written in Arc/Info Micro Language (AML) to estimate the totals of population, dwelling units, employment, etc., within a buffer zone of a count station. Several radii have been chosen for the buffer zones, including 0.25, 0.5, 1.0, 1.5, 2.0, and 3.0 mile(s). The purpose is to test in what geographic context the socioeconomic data will have a noticeable impact on ADT. The final data chosen is generated using a 0.25-mile buffer zone. The choice is made after analyzing the correlation between ADT and the above factors. Because the radii do not seem to affect significantly the degree of correlation between ADT and the socioeconomic factors, a small buffer zone is used. A smaller buffer zone means a more localized impact, which promises to simplify the classification of roadways by examining land use in a small area. However, as it turns out later the socioeconomic impact on ADT is minimal. This point will be discussed in the final chapter where the conclusions will be presented.

Accessibility Data

Accessibility measures whether a off-system road has an easy access to state roads such as expressways and whether it is easily accessed by other county and city roads. The word "easy'" here means close proximity. Two variables, *ACCESS1* and *ACCESS2*, are used to indicate the presence of state roads and county roads nearby, respectively:

(1) Accessibility to state roads (*ACCESS1*): this variable will assume a value of 1 when there are state roads nearby, and 0 otherwise;
(2) Accessibility to off-system roads (*ACCESS2*): this variable will be 1 when there are other county roads nearby, and 0 otherwise.

As mentioned before, road network, roadway attributes, and count station locations are all available in GIS format. This allows GIS to be used to create the accessibility data. The accessibility measurements for a count station are obtained using an AML program. The program examines a buffer zone surrounding a count station to determine the presence of other (different) roads that have a functional classification of state principal arterial (expressways), state principal arterial other than expressways, state minor arterial, count minor arterial, county collector, or city collector. Various radii including 0.5, 1.0, 1.5, 2.0, 3.0 miles are investigated. Both 0.5- and 1.0- mile buffers produce similar statistical results. The final model uses the accessibility data obtained with an 1.0-mile buffer.

4.4.2 Model Development

Full Model Regression

The results of the multiple regression analysis using all 16 independent variables are summarized in **Figure 4.15**. There is a strong relationship between *ADT* and the independent variables. The coefficient of determination, R^2, is *0.6422*. In other words, *64.22%* of the sum of error squares in *ADT* may be associated with the variation in these independent variables. The overall *F*-test is very significant (*Prob.* $>$ *F* = *0.0001*) while some of the *t* statistics for individual variables are insignificant (e.g. *HTL: prob.* $> |T| = .6060$, etc). The significance of *F* test indicates that the model has significant overall utility. At the same time, the prob.-values related to some individual *t* statistics are large, indicating that some of the individual independent variables in the model are not as important as others.

Variable Selection

The results of *R*-Square selection (see **Figure 4.16**) indicate that of the single variable subsets, the first best, *L* (number of lanes), explains *49.99%* of the variation in *ADT*. The second most significant variable, *FCLASS1* (functional classification), accounts for 6.30% of variation in *ADT*. The other variables only slightly improve the R^2. The result of the *Cp* regression is shown in **Figure 4.17** which indicates that nine variables should be included in the model. The nine variables are accessibility to county roads (*ACCESS2*), number of lanes (*LAN*), function classification (*FCLASS1*), facility type (*FAC1*), area type (*AREA1*), auto ownership (*AUTO*), service employment (*SEREMP*), school enrollment (SCHOOL), and multi-family dwelling units (*MFDUS*). However, when the *VIF* selection is used to detect multicollinearity, it is found that several of the variables among the nine are correlated. As a result, facility type (*FAC1*), school enrollment (SCHOOL), and multi-family dwelling units (*MFDUS*) are eliminated from the model. A principal component analysis is then performed, which indicates that service employment (*SEREMP*) is insignificant compared to the other variables. Therefore, the final set of variables include five: accessibility to county roads (*ACCESS2*), number of lanes (*L*), function classification (*FCLASS1*), area type (*AREA1*), and auto ownership (*AUTO*).

Final Model

Before the final model runs, outliers in the data set that reflect abnormal conditions are detected. Six outliers are found and eliminated, leaving a data set containing 393 data points. Using this data set, a model with five selected independent variables, *ACCESS2, LANES, AREA1, FCLASS1,* and *AUTO*, is constructed. The output from SAS is shown in **Figure 4.18.** The regression of these five variables is shown in **Table 4.9**. The R^2 is 0.6120 and the adjusted R^2 is 0.6069. The final model has the following form:

$$ADT = -12886 + 4689.86 \times LANES + 5227.57 \times FCLASS1 + 1388.27 \times AREA1 + 0.15 \times AUTO - 1224.06 \times ACCESS2$$

The signs of the coefficients are as expected. For example, the positive sign for *L* indicates that roads with more lanes tend to have a heavier traffic. The positive sign for *FCLASS1* means that a road of a function higher in the network hierarchy will carry more daily traffic. The positive sign for *AREA1* (area type or land use) indicates that a road located in a densely developed area (like a residential area) would carry more traffic. The positive sign for *AUTO* explains the positive relationship between ADT with the total number of automobiles nearby. The negative sign of *ACCESS2* implies that county minor arterials or collectors nearby will compete with each other. In other words, the availability of multiple choices of routes in an area will tend to reduce the traffic on individual roads.

<u>Model Validation</u>

A data set of 34 testing points are used to examine the model's predictive capability. **Table 4.16** summarizes the result. For each observation, the table indicates the observed ADTs, predicted ADTs, and percentage error in the estimation. The absolute percentage difference between these two ADT values ranges from 0.86% to 61.99%, with an average difference of 10.07%. The system wide error is -13.68%, which means the model tends to underestimate total traffic.

Figure 4.31 illustrates the distribution of the errors. The bar chart shows the number of testing points whose errors fall into a particular range. It may be seen that testing points with large errors are few. Most testing points have an error under 40 percent. **Figure 4.32** displays the cumulative number of testing points whose errors are below a certain level. For instance, 55.6% of the testing points have an error smaller than 25% while 83.3% of the testing points have an error less than 33%.

Model: MODEL1
Dependent Variable: ADT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	16	29411093137	1838193321	41.624	0.0001
Error	376	16604716482	44161480.004		
C Total	392	46015809618			

Root MSE	6645.41045	R-square	0.6392
Dep Mean	18955.72519	Adj R-sq	0.6238
C.V.	35.05754		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-14123	2272.9633179	-6.213	0.0001
ACCESS1	1	118.129402	319.16710086	0.370	0.7115
ACCESS2	1	-904.159732	438.46703891	-2.062	0.0399
L	1	4474.047444	284.62777994	15.719	0.0001
AREA1	1	1651.888347	611.47270873	2.701	0.0072
FCLASS1	1	5808.054250	725.18432089	8.009	0.0001
HTL	1	-0.141809	0.17843855	-0.795	0.4273
POP	1	-0.002806	0.08883123	-0.032	0.9748
AUTO	1	0.166660	0.13195292	1.263	0.2074
INDEMP	1	-0.105433	0.19443324	-0.542	0.5880
COMMEMP	1	0.112912	0.27362331	0.413	0.6801
EMP	1	-0.139897	0.09239200	-1.514	0.1308
SCHOOL	1	0.203288	0.13427954	1.514	0.1309
SFPOP	1	3.636252	2.86588473	1.269	0.2053
MFPOP	1	-4.951974	2.02345070	-2.447	0.0149
SFDUS	1	-11.115729	8.28537266	-1.342	0.1805
MFDUS	1	10.747325	3.87594413	2.773	0.0058

Figure 4.16 Initial Results of the Multiple Regression Analysis for the Broward County Model

```
N = 393     Regression Models for Dependent Variable: ADT

      Number in   R-square       C(p) Variables in Model
       Model

          1   0.50539609  126.37221 L
          1   0.28191452  359.23771 FCLASS1
       -------------------------------------
          2   0.58043944   50.17780 L FCLASS1
          2   0.53673656   95.71576 L AUTO
       -------------------------------------
          3   0.59821954   33.65112 L FCLASS1 AUTO
          3   0.59661315   35.32497 L FCLASS1 POP
       --------------------------------------------
          4   0.61458968   18.59360 L FCLASS1 AUTO EMP
          4   0.61028407   23.08000 L FCLASS1 AUTO COMMEMP
       -----------------------------------------------------
          5   0.62280954   12.02860 L AREA1 FCLASS1 AUTO EMP
          5   0.61954001   15.43541 L AREA1 FCLASS1 POP EMP
       ------------------------------------------------------
          6   0.62702235    9.63889 ACCESS2 L AREA1 FCLASS1 AUTO EMP
          6   0.62549064   11.23491 L AREA1 FCLASS1 AUTO EMP MFDUS
       -------------------------------------------------------------
          7   0.62964383    8.90733 L AREA1 FCLASS1 POP EMP MFPOP MFDUS
          7   0.62958283    8.97090 ACCESS2 L AREA1 FCLASS1 AUTO EMP MFDUS
       -----------------------------------------------------------------------
          8   0.63364085    6.74247 ACCESS2 L AREA1 FCLASS1 POP EMP MFPOP MFDUS
          8   0.63265360    7.77118 L AREA1 FCLASS1 AUTO EMP SCHOOL MFPOP MFDUS
       --------------------------------------------------------------------------
          9   0.63646081    5.80411 ACCESS2 L AREA1 FCLASS1 AUTO EMP SCHOOL MFPOP MFDUS
          9   0.63534176    6.97015 ACCESS2 L AREA1 FCLASS1 POP EMP SCHOOL MFPOP MFDUS
       --------------------------------------------------------------------------------
         10   0.63681828    7.43163 ACCESS2 L AREA1 FCLASS1 HTL AUTO EMP SCHOOL MFPOP MFDUS
         10   0.63670179    7.55302 ACCESS2 L AREA1 FCLASS1 AUTO EMP SCHOOL MFPOP SFDUS MFDUS
       ---------------------------------------------------------------------------------------
         11   0.63823284    7.95768 ACCESS2 L AREA1 FCLASS1 AUTO EMP SCHOOL SFPOP MFPOP SFDUS MFDUS
         11   0.63711142    9.12619 ACCESS2 L AREA1 FCLASS1 HTL AUTO EMP SCHOOL MFPOP SFDUS MFDUS
       -------------------------------------------------------------------------------------------
         12   0.63867697    9.49490 ACCESS2 L AREA1 FCLASS1 HTL AUTO EMP SCHOOL SFPOP MFPOP SFDUS
                                    MFDUS
         12   0.63844167    9.74008 ACCESS1 ACCESS2 L AREA1 FCLASS1 AUTO EMP SCHOOL SFPOP MFPOP
                                    SFDUS MFDUS
       ---------------------------------------------------------------------------------------------
         13   0.63880238   11.36422 ACCESS2 L AREA1 FCLASS1 HTL AUTO INDEMP EMP SCHOOL SFPOP MFPOP
                                    SFDUS MFDUS
         13   0.63879840   11.36837 ACCESS1 ACCESS2 L AREA1 FCLASS1 HTL AUTO EMP SCHOOL SFPOP MFPOP
                                    SFDUS MFDUS
       ---------------------------------------------------------------------------------------------
         14   0.63902036   13.13708 ACCESS2 L AREA1 FCLASS1 HTL AUTO INDEMP COMMEMP EMP SCHOOL SFPOP
                                    MFPOP SFDUS MFDUS
         14   0.63898657   13.17229 ACCESS1 ACCESS2 L AREA1 FCLASS1 HTL AUTO INDEMP EMP SCHOOL SFPOP
                                    MFPOP SFDUS MFDUS
       ---------------------------------------------------------------------------------------------
         15   0.63915096   15.00100 ACCESS1 ACCESS2 L AREA1 FCLASS1 HTL AUTO INDEMP COMMEMP EMP
                                    SCHOOL SFPOP MFPOP SFDUS MFDUS
         15   0.63902045   15.13699 ACCESS2 L AREA1 FCLASS1 HTL POP AUTO INDEMP COMMEMP EMP SCHOOL
                                    SFPOP MFPOP SFDUS MFDUS
       ---------------------------------------------------------------------------------------------
         16   0.63915192   17.00000 ACCESS1 ACCESS2 L AREA1 FCLASS1 HTL POP AUTO INDEMP COMMEMP EMP
                                    SCHOOL SFPOP MFPOP SFDUS MFDUS
       ---------------------------------------------------------------------------------------------
```

Figure 4.17 R-square Selection for the Broward County Model

```
     Plot of _CP_*_IN_.  Symbol used is 'C'.
     Plot of _P_*_IN_.   Symbol used is '*'.

     ,         C
     ,         C
     ,
  25 ^
  24 ^
  23 ^
  22 ^           C
  21 ^           C
  20 ^
M 19 ^
a 18 ^                                   C
l 17 ^                                 *
l 16 ^             C                 * C
o 15 ^             C               * C
w 14 ^               C           *
s 13 ^                         *   C
  12 ^               C C   C C C C
C 11 ^                 C C C C C
( 10 ^                   *
p  9 ^                 *
)  8 ^               *
   7 ^             *
   6 ^           *
   5 ^         *
   4 ^       *
   3 ^     *
   2 ^   *
   1 ^
   0 ^
     ,
     Šƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒˆƒƒƒƒƒƒƒƒ
      0 1 2 3 4 5 6 7 8 9 1 1 1 1 1 1
                          0 1 2 3 4 5

            Number of regressors in model

  NOTE: 24 obs hidden.  6 obs out of range.
```

Figure 4.18 C_p Test for the Broward County Model

Model: MODEL1
Dependent Variable: ADT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	5	28159517335	5631903467	122.060	0.0001
Error	387	17856292283	46140290.137		
C Total	392	46015809618			

Root MSE	6792.66444	R-square	0.6120
Dep Mean	18955.72519	Adj R-sq	0.6069
C.V.	35.83437		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|	Variance Inflation
INTERCEP	1	-12886	2152.5779271	-5.986	0.0001	0.00000000
L	1	4689.860252	267.45583171	17.535	0.0001	1.20999812
ACCESS2	1	-1224.057937	421.07458492	-2.907	0.0039	1.11462578
AREA1	1	1388.279943	595.87012528	2.330	0.0203	1.08548103
FCLASS1	1	5227.572162	709.61578411	7.367	0.0001	1.25897429
AUTO	1	0.150255	0.03273980	4.589	0.0001	1.14816277

Figure 4.19 Final Results of the Multiple Regression Analysis for the Broward County Model

Table 4.9 Model Validation for the Broward County Model

ACCESS2	L	FCLASS	Fclass1	AREA	AREA1	AUTO	ADT	R-ADT	% Error
1	2	3	1	3	3	6627	8100	5658	-30.15%
2.5	4	3	1	4	4	2240	8600	13931	61.99%
2.5	2	3	1	4	4	27459	9400	8340	-11.28%
0	4	0	0	3	3	18168	10100	12768	26.42%
2.5	2	3	1	4	4	35416	10700	9536	-10.88%
1	2	3	1	3	3	10534	11700	6245	-46.62%
1	2	5	1	3	3	43222	13400	11156	-16.75%
2.5	2	2	2	4	4	3615	14100	9985	-29.18%
1	2	6	2	3	3	22792	14900	13314	-10.64%
1.5	4	2	2	3	3	2126	15200	18977	24.85%
1.5	4	5	1	3	3	39151	15400	19312	25.40%
2.5	4	3	1	5	1	17094	15600	11998	-23.09%
1.5	4	0	0	3	3	21832	16500	11482	-30.41%
0	4	6	2	3	3	1489	16900	20717	22.59%
2.5	4	2	2	2	2	43264	17000	22546	32.62%
1	4	3	1	3	3	39355	17000	19955	17.38%
2.5	4	3	1	3	3	30032	17300	16718	-3.36%
2.5	4	3	1	3	3	30280	18700	16755	-10.40%
2.5	2	3	1	3	3	37851	18800	8513	-54.72%
2.5	4	3	1	3	3	25928	19800	16101	-18.68%
1	4	2	2	3	3	44333	20500	25931	26.49%
2.5	4	2	2	3	3	43953	23900	24037	0.57%
1	4	2	2	3	3	40001	25500	25280	-0.86%
2.5	4	3	1	4	4	30215	25500	18134	-28.89%
1	4	3	1	4	4	17281	22400	18027	-19.52%
1	4	3	1	3	3	3513	28500	14570	-48.88%
1	4	5	1	3	3	40066	29600	20062	-32.22%
2.5	6	3	1	3	3	25348	29000	25394	-12.43%
1	6	6	2	3	3	26904	30900	32692	5.80%
2.5	4	3	1	2	2	28381	31700	15082	-52.42%
1.5	4	5	1	3	3	32702	31500	18343	-41.77%
1	6	6	2	3	3	23454	31600	32173	1.81%
2.5	6	2	2	3	3	39221	35700	32706	-8.39%
2.5	6	2	2	3	3	31136	39000	31491	-19.25%
1.5	6	2	2	3	3	3384	40900	28546	-30.21%
0	6	6	2	4	4	38909	44900	37108	-17.35%

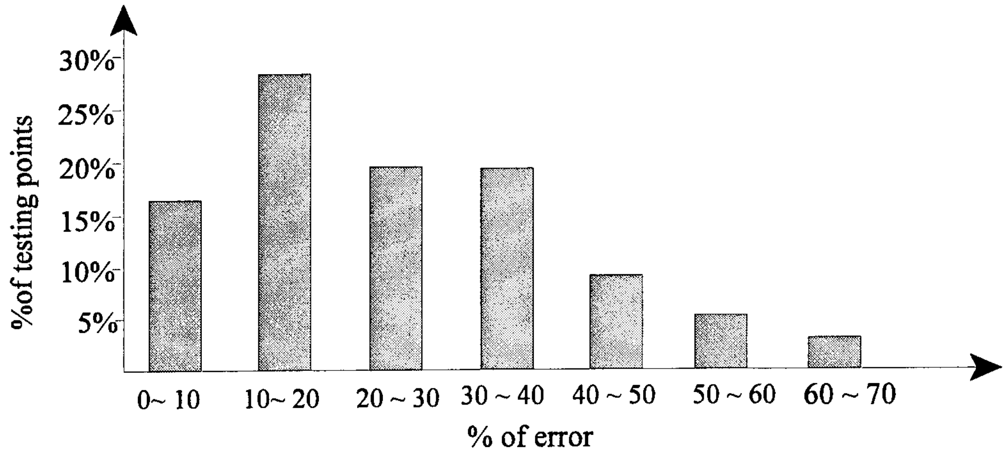

Figure 4.20 Error Distribution in Testing Data Set for the Broward County Model

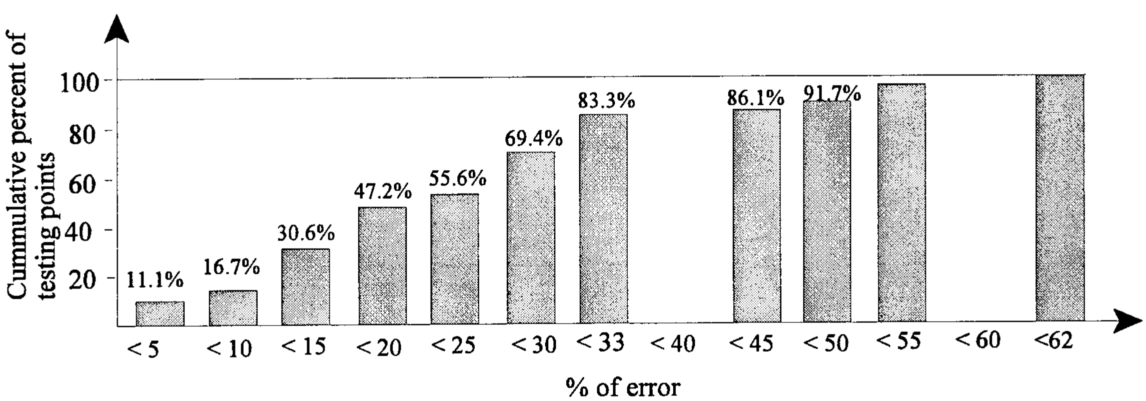

Figure 4.21 Cumulative Percent of Testing Points in a Given Error Range for the Broward County Model

4.4.3 Discussions

Functional classification and area types are included in the Broward County model. While it appears logical that functional classification and number of lanes should be correlated, the correlation is in fact not very strong. A closer examination revealed that a road higher in the hierarchy of the classification system does not always have more lanes. This and other observations including, for example, a roadway with a high ADT has a classification at the low end in the hierarchy, raise some concerns about the functional classification system utilized in Broward County. The residual tests have indicated that the outliers, i.e., abnormal data, have some relationship with the functional classification variable. It seems that these outliers' functional classifications are not consistent with the magnitude of the ADTs associated with them. For example, one of the outliers is from a count station on a segment of Lyons Road, which is a city collector with ADT of 28,262 in 1997. This ADT is much higher than the mean daily traffic of 15,051 on 166 city collectors. Thus, the functional classification for this segment of Lyons Road does not reflect its inner relationship with ADT. A closer examination of Lyons Road reveals that it is a 6-lane roadway, and while this particular segment is a city collector, its other portions are classified as a county minor arterial. This suggests that road classification should not be based on road ownership but rather on their traffic moving functions. Broward County is currently completing the reclassification of its roadways according to the FHWA functional classification system.

The Broward County model demonstrates that characteristics of the roadways themselves are more important to the variation in ADT compared with the socioeconomic factors, which do not appear to be as significant as expected. One of the original objectives of this study is to determine if a buffer method is used, what geographic extent will be to which local land uses will impact traffic on nearby roads. Different radii of 0.25 mile, 0.5 mile, 1.0 mile, 1.5 miles, 2 miles, and 3 miles were experimented. The results show, however, that these variables have a minimum impact on the R^2 of the models. One possible reason is that the buffer method is not able to account for through traffic not originated locally. Another possible reason is that when several count stations are close to each other, there will be overlapping in the socioeconomic data using the "buffer zone" method, which is not considered a good statistical experiment design. When the ADT distribution and land uses patterns were visualized using GIS, it is observed that there is a positive relationship between ADT and land use intensity. Therefore, it cannot be concluded that socioeconomic factors do not contribute to ADT.

Area type is another variable that is worth further study. Area type is intended to reflect land use intensity. However, it is unclear how area types are defined or updated. Inaccurate area types will cause errors in the model while a good classification of area types will likely improve the model. Such an improved system may also be used as one of the criteria to define the categories of roads for ADT estimation.

5. SUMMARY AND CONCLUSIONS

In this research, an extensive data collection has been conducted. The results show that FDOT has limited information on off-system, which is not sufficient to support model development. Data of off-system roads are collected by local agencies including MPOs and public works departments. Availability of digital data is limited, which requires additional resources for data processing so that the data may be used in model development. Also lacking is the land use information. While many urban areas have FSUTMS data, the models are generally old. For practical application, newer data are needed.

For model development, multiple linear regression is chosen as the modeling method based on the literature review, and FDOT's preferences. The choice is based on the desirable simplicity of regression analysis and existing literature, that supports its appropriateness for application to ADT estimates. Four models have been developed using the same regression technique. The four models are defined based primarily on types of land uses, urban population sizes, and unique characteristics exhibited by some of the cities. The models are summarized in **Table 5.1** and their performances in **Table 5.2**, respectively. **Table 5.1** gives for each model the model name, independent variables, number of variables, and sample size used to generate the models. **Table 5.2** compares the model performances b their based on their R^2, adjusted R^2, which is considered a better indication of a model's explanatory capability, root mean square error, the size of the testing data set, and minimum, maximum, and average percent of errors in absolute values.

Table 5.1 Summary of Four Regression Models

Model Name	*Model (dependent variable: ADT)*	*Number of Variables*	*Sample Size*
State-wide	$9643.704161 + 0.014645 \times POP - 0.155037 \times LABOR - 0.181236 \times INCOME + 0.000005139 \times TAXABLE + 0.058710 \times VEHICLES$	5	107
Rural area	$4853.489444 + 0.122587 \times POP + 0.261858 \times LABOR - 18.930235 \times LANEMILE - 0.0032338 \times VEHICLES$	4	27
Small-medium urban area	$ADT = -13418 + 6770.23 \times LANES + 1580.14 \times ATYPE1 + 2.85 \times COM\_EMP + 1.78 \times HOT\_OCC$	4	245
Broward	$-12886 + 4689.86 \times L + 5227.57 \times FCLASS1 + 1388.27 \times AREA1 + 0.15 \times AUTO - 1224.06 \times ACCESS2$	5	393

Table 5.2 Summary of Performance of Six Models

Model Name	R^2	*Adj.* R^2	*Testing Date Points*	*Min. % |Error|*	*Max. % |Error|*	*Avg. % |Error|*
State-wide	0.2890	0.2538	11	8.00	1096.00	188.00
Rural area	0.4488	0.3486	3	6.87	83.96	35.59
Small-medium urban area	0.6937	0.6856	30	2.30	62.10	22.66
Broward	0.6120	0.6069	36	0.86	61.99	23.73

The first two models, the state wide and rural models, have four and five variables, respectively. The data used are at the county level, which is updated annually, although the update lags about three years. The advantage of the data source is, in addition to being continually updated, being readily available. However, both model suffer from a common problem, which is the correlation among all the variables. This suggests that the current variables set is inadequate and more variables must be considered.

These two models do not perform well as indicated by their low R^2 and large errors in the testings. One of the possible reasons for the poor performance of the model may be that some important independent variables were not included in the model. For instance, information on area type and accessibility was not available while it has been learned from other models that these two variables may be important. Such information is, perhaps, more important than just providing information on local environment. It will also allow the models to account for variations in the data set within the same county. In fact, since all variables are aggregated at the county level, the model is not able to consider differences of ADT within the same county. On the other hand, while number of lanes has been an important variable for other models, it is not for the rural model, since most roads are of two lanes, therefore, it does not offer power of explaining variations in ADT.

Another possible reason for the rural area model not to perform well may be the limited sample size. The sample size for the state wide model is 107 from 25 counties, with each county averaging only about 5 samples. The rural model only has 27 data points for model construction, which are from eight counties, averaging a little over three data points from each county.

The large variances in the size and characteristics of the counties also contribute to the low R^2. The population sizes range from 24,182 to 2,031,336. Some of the counties are rural while others are urbanized including major metropolitan areas, such as Miami-Dade County.

In the small-medium urban area model, the sample contains 270 ADTs from ten counties, including Charlotte, Bya, Okaloosa, Leon, Marion, Escambia, Volusia, Polk, Pinellas, and Palm Beach Counties. Besides the ADT, other data are obtained from FSUTMS ZDATA files and are aggregated at the TAZ level. None of the variables share the same value as those in the state wide model and the rural area model. The model performance is significantly improved compared to the first two models. The percent different between the observed ADT and the predicted ADT values vary from 1.96% to 62.10%, with an average difference of 22.66%.

The Broward County Model is unique in several ways. It is a large urban area with complex urban forms, transportation systems, and travel patterns. No literature has been found on estimating ADT for such a large urban area using statistics or other techniques. It is therefore interesting to investigate whether a model different from the existing long term travel forecasting models can be built, and what factors will account for ADT variations in a large urban area.

The model has five independent variables, number of lanes, function classification, area type, automobile ownership, and access to county roads. With the exception of automobile ownership, the values of other variables are easily obtained. However, it also reflects the importance of the necessity of ensuring the accuracy of the function classification system and the area type classification system. In particular, the functional classification system utilized in Broward County needs some special attention. The residual tests have indicated that the outliers, i.e. abnormal data, have some relationship to the functional classification variable. It seems that these outliers' functional classifications are not consistent with the magnitude of the ADTs associated with them. For example, one of the outliers is from a count station on a segment of Lyons Road, which is a city collector with ADT of 28,262 in 1997. This ADT is much higher than the mean daily traffic of 15,051 on 166 city collectors. Thus, the functional classification for this segment of Lyons Road does not reflect its inner relationship with ADT. A closer examination of Lyons Road reveals that it is a 6-lane roadway, and while this particular segment is a city collector, its other portions are classified as a county minor arterial. This suggests that road classification should not be based on road ownership but rather on their traffic moving functions. Broward County is currently evaluating its roadway functional classification system and may reclassify its roadways in the future.

The Broward County model demonstrates that characteristics of the roadways themselves are more important to the variation in ADT compared with the socioeconomic factors, which do not appear to be as significant as expected. One of the original objectives of this study is to determine that if a buffer method is used, to what geographic extent local land uses will impact traffic on nearby roads. Different radii of 0.25 mile, 0.5 mile, 1.0 mile, 1.5 miles, 2 miles, and 3 miles were experimented. The results show, however, that these variables have minimum impact on the R^2 of the models. One possible reason is that the buffer method is not able to account for through traffic not originated or destined locally. Another possible reason may be that when several count stations are close to each other, there will be overlapping in the socioeconomic data using the "buffer zone" method, which is not considered a good statistical experiment design. It is also possible that the large retirement population in Broward County may have affect the model. These and other possible causes remain to be studied.

While the last four model have achieved an adjusted R^2 ranging from 0.61 to 0.69, meaning that 61% to 69% of the variability in ADT is explained by the independent variables, improvement on accuracy of the models is still desired. Additional research is also needed to study how the coefficients of such statistical models will change with time. Currently, the research is on-going at the Lehman Center for Transportation Research, Florida International University. Possible improvements of the model are being investigated, such as inclusion of nonlinear relationship between ADT and some of the predictors. The land uses will also be further studied to better understand the impact of socioeconomic characteristics of an area on traffic and if necessary, methods for collecting and processing such data economically. Future study will also attempt to apply the developed methodology to other Florida urban areas.

The models generally have a rather large negative intercept. While the intercepts themselves do not have physical meanings, their large negative values tend to make the model underestimate ADT more often. This is perhaps a result of not including all relevant variables that have a significant impact on ADT.

Finally, the regression models developed do not have a temporal dimension. This means that the model coefficients will stay constant even when the conditions in the system being modeled have changed. The consequence may be that the model error will grow larger and larger as time elapses. To control the model error, the model will need to be recalibrated periodically.

6. RECOMMENDATIONS

This research has identified sets of variables that may be used as predictors of AADT. These sets of variables, along with any that may be identified in future studies, may be used to classify roads into categories of similar roads, which may then be used to estimate AADT for roads that do not have traffic counts using traffic count data for roads in the same category. However, before this classification is attempted, more study is needed to further investigate the issues that arise from this research. The following research is recommended to be considered for future effort to improve AADT prediction methods:

(1) *Rural Model Improvement.* Further investigate the development of a rural county model that incorporate more variables, such as land uses and accessibility that reflect local conditions. Since rural counties may not have FSUTMS data, the main data will be at the county level. Some local data may also be gathered from the relevant county agencies.

(2) *Functional classification.* The functional classification of a road indicates its intended function in a hierarchy of a roadway system, which has important implication on ADT. It also helps to distinguish two roads that have the same number of lanes but carry different traffic volumes. Some problems have been identified concerning Broward County functional classification system. It is desirable to use a standard functional classification system throughout the state. Such system may be the one adopted by the Federal Highway Administration (FHWA).

(3) *Accessibility.* In the Broward County model, accessibility of a location is evaluated by determining the existence of nearby county roads. The evaluation is carried out using the buffer zone method, which will identify all county road segments that fall into a radius from the location being considered. In practice, street connectivity is important. However, considering street connectivity will involve substantial effort in studying the street patterns. This has not been accomplished in the current research, because of the effort required to define network routes in the GIS base map.

Alternatively, other measurements of accessibility may be developed. For instance, results from a preliminary study indicates that when a count is taken near an intersection, the types of the roads, that intersect, may be an important factor. For instance, traffic volume on a minor arterial at an intersection, with the cross street being a collector, may be higher than if the cross street is a local road. Whether a road is connected to expressways also appear to have a bearing on ADT. These and other phenomena should be carefully studied.

For the rural model, accessibility information is not included due to the lack of information and the resources required to obtain such information. Future work may include the development of accessibility data for rural roads, and improve the model by including such information in the model.

(4) *Area type.* Area type has been an important factor in the models. However, the determination of area types may not be straightforward. This is especially important for urban areas, since a urban area tends to have much greater variations in land use intensities, and land use patterns may be complex. In fact, an expert on traffic and travel models from Chicago area has suggested to avoid using area types at all due to the ambiguity in the definition. It may be necessary to study the methodologies used to define the area types for the FSUTMS models and determine if such methodologies are sound. If the current area types are not to be used for urban areas, then other methods are needed to describe the land uses.

For the current version of the rural model, land use information is missing. The model may be improved by adding area type information, which may, for instance, include categories such as county seat, other urban, and rural.

(5) *Socioeconomic impact.* It agrees with our intuition that land use intensity should have some impact on ADT. Examples include, arterial roads with strip commercial developments. From studying the employment distribution in Broward County, it appears that corridors that have heavy commercial activities also tend to have higher ADTs. However, the simple methods of aggregating socioeconomic data developed in this research do not adequately capture such correlations. Local effects of land uses in a small area may only be noticeable on local roads and may be insignificant on a collector or an arterial, since the latter serve a much larger area. One possible approach to study the impact of economic activities on ADT of a road segment is to define a service area for the corridor of which the road segment is a part. This larger service area will account for some of the through traffic that is not originated or destined locally.

(6) *Model testing.* Test the Broward model on other large urban areas including the Miami-Dade County. Also develop a model based on Miami-Dade County data, since much of the needed data are available in GIS format including detailed land use information, which is not available in Broward County. The main benefits will be to compare the models, in order to determine a common set of variables that may be used for establishing categories for estimating ADT, and determine to what degree differences in geographic characteristics will affect the model structure.

(7) *AADT Lookup table.* Develop tables that may be conveniently used to determine AADT for a roadway based on its characteristics. The tables may be constructed using the variables identified in the regression models.

(8) *Data collection procedures.* Develop practical data collection procedures to allow data needed for model application be gathered easily and inexpensively.

(9) *Model update.* Regression models are developed based on existing data, which define the model coefficients. Once developed, the coefficients of a model are fixed unless new data are available to recalibrate the model. Whether the model will continue to perform well or

how fast the model will deteriorate with time remains unknown. As a result, additional research will be needed to answer these questions. While theoretically it is possible to use the same set of variables and apply data from different years, and then compare the resultant models, lack of historical data, especially socioeconomic data, may make this a challenge. A long term approach would be to continually check the model with new traffic data and in the course of several years observe the trend of model performance.

(10) *New methodologies*. Other potential techniques, such as neural networks may be investigated. As mentioned in the literature review, neural networks are good at capturing nonlinear relationship between AADT and variables. They are also insensitive to correlations between independent variables, which are undesirable for regression models and must be eliminated. Neural networks are relatively easy to develop, and may be continually updated with new data. Self learning and updating are important features of neural networks and ensure that a model will remain current without the need for periodical recalibration.

(11) *Seasonal factors*. One problem encountered in the project is the lack of AADT data. As a result, ADT data are used in place of AADT. Broward County, for instance, has begun to convert ADT to AADT starting for 1998 data. The seasonal factors applied are developed based on traffic counts from permanent count stations on the state roads. The applicability of these seasonal factors to off-system roads needs to be studied since many off-system roads possess rather different characteristics when compared to state roads. Therefore, seasonal factors need to be studied that can be applied to different types of off-system roads. Good seasonal factors will eliminate biases in the data set resulting from the different times when data are collected.

REFERENCES

Albright, D. (1991), "An Imperative For, and Current Progress Toward, National Traffic Monitoring Standards", *ITE Journal*, June, pp. 22-26.

Albright, D. (1987), "A Quick Cluster Control Method: Permanent Control Station Cluster Analysis in Average Daily Traffic Calculations.", *Transportation Research Record* 1134, Transportation Research Board, National Research Council, Washington, D.C., pp. 57-64.

Albright, D. (1993), "Standards, Innovation, and the Future of Traffic Monitoring", *ITE Journal*, Jan, pp.31-36.

Aldrin, M.(1995). "A Statistical Approach to the Modeling of Daily Car Traffic", *Traffic Engineering and Control*, Vol 36, Issue 9, pp. 489-493.

Bester, C. J. and Joubert, J.D. (1986). "Classifying a Rural Road Network for Traffic Counting", *Transportation Research Record* 1090, Transportation Research Board, National Research Council, Washington, D.C., pp. 22-26.

Cheng, C. (1992), *Optimum Sampling for Traffic Volume Estimation*, PhD Dissertation, University of Minnesota, Carlson School of Management.

Crevo, C.C. and Virkud, U. (1994), "Practical Approach to Deriving Peak-hour Estimates from 24-hour Travel Demand Models", *Transportation Research Record* 1443, Transportation Research Board, National Research Council, Washington, D.C., pp. 30-37.

Dean, M. I., Jefferys, T., Justice, H.A. and Orcutt, F. L. (1993), "Standard File Format for System Detector Counts", *ITE Journal*, November 1993.

Davis, G. A. and Guan, Y. (1996), "Bayesian Assignment of Coverage Count Locations to Factor Groups and Estimation of Mean Daily Traffic", *Transportation Research Record 1542*, Transportation Research Board, National Research Council, Washington, D.C., pp. 30-37.

Deacon, J. A., Pigman, J. G. and Mohenzadeh, A. (1987), *Traffic Volume Estimates and Growth Trands*, UKTRP-87-32, Kentucky Transportation Research Program, University of Kentucky.

Erhunmwunsee, P. (1991), "Estimating Average Annual Daily Traffic Flow from Short Period Counts", *ITE Journal*, November, pp. 23-30.

Faghri, A. and Hua, J. (1995), "Roadway Seasonal Classification Using Neural Networks", *Journal of Computing in Civil Engineering*, American Society of Civil Engineers, Vol. 9, No. 3, pp. 209-215.

FDOC (1995), *1995 Florida County Profile*, Division of Economic Development, Bureau of Economic Analysis, Florida Department of Commerce, Tallahassee, FL 32399-2000.

FDOT (1997), *State Highway System Report 1: All Roads*, Florida Department of Transportation, December 31, 1997.

Flaherty, J. (1993), "Cluster Analysis of Arizona Automatic Traffic Recorder Data", *Transportation Research Record* 1410, pp. 93-99.

Fricker, J.D. and Saha, S.K. (1987), *Traffic Volume Forecasting Methods for Rural State Highways – Final Report*, FHWA/IN/JHRP-86/20, joint Highway Research Project, School of Civil Engineering, Purdue University.

Garber, N.J. and Bayet-Mokhtari, F. (1986), "A Computerized Highway Link Classification System for Traffic Volume Counts", *Transportation Research Record* 1090, Transportation Research Board, National Research Council, Washington, D.C., pp. 1-7.

Kopanezou, H. and Trivellas, T. (1989), "A Time Series Model for Daily Traffic Volume Forecasting", *Proceedings of the Fourth International Conference on Civil and Structural Engineering Computing*, Civil-comp Press, pp. 295-300.

ITE Technical Council Committee 6f-42 (1993), "The Age and Status of Transportation Planning Database", *ITE Journal*, March 1993, pp. 42-44.

Mohamad, D., Sinha, K.C., Kuczek, T. and Scholer, S. F. (1998), "An Annual Average Daily Traffic Prediction Model for County Roads", presented at the Transportation Research Board 77th Annual Meeting, January 11-15, 1998, Washington D. C.

Ochoa, D. L. and Ramsey, G. M..(1986), "Estimating Sampling Error for Cluster Sample Travel Surveys by Replicated Subsampling", *Transportation Research Record* 1090, Transportation Research Board, National Research Council, Washington, D.C., pp. 36-42.

Ritchie, S.G. (1986), "A Statistical Approach to Statewide Traffic Counting", *Transportation Research Record* 1090, Transportation Research Board, National Research Council, Washington, D.C., pp. 14-21.

Ritchie, S.G. and Hallenbeck, M.E. (1986), "Evaluation of a Statewide Highway Data Collection Program", *Transportation Research Record* 1090, Transportation Research Board, National Research Council, Washington, D.C., pp. 27-35.

Sharma, S.C. (1996), "Design Hourly Volume from Road User's Perspective", *Journal of Transportation Engineering*, American Society of Civil Engineers, Vol. 112, No. 4, July, pp. 435-441

Sharma, S.C. , Kiburn, P, and Wu, Y. (1996), "The Precision of Average Daily Traffic Volume Estimates from Seasonal Traffic Counts: Alberta Example". *Canadian Journal of Civil Engineering*, Vol. 23, No. 1, February, pp. 302-304.

Sharma, S.C. and Leng, Y. (1994), "Seasonal Traffic Counts for a Precise Estimation of AADT", *ITE Journal*, September, pp.21-28.

Sharma, S.C., Lingras, P., Xu, F., and Liu, G.X. (1999), "Neural Networks as an Alternative to The Traditional Factor Approach of Aadt Estimation From Traffic Counts", presented at the Transportation Research Board 78th Annual Meeting, Washington, D.C., January 10-14.

Showers, R. H. and Courage, K. G. (1992), "The Permanent Traffic Count Analysis Package", *the 4th International Conference on Microcomputers in Transportation,* Baltimore, M.D, July 22-24, pp. 415-423.

APPENDIX A. SURVEY FORM

OFF-SYSTEM (NON-STATE) ANNUAL AVERAGE DAILY TRAFFIC QUESTIONNAIRE

FDOT District: ____________
District Division/Department: ____________
Person responding to the survey: ____________
Phone: ____________ Fax: ____________
E-Mail: ____________
Mailing Address: ____________

- ***Place a check mark next to the question(s) you would like to provide additional explanations by telephone.***
- ***Attach any forms used by the district relevant to the questions, and use additional paper if needed.***

1. What is the total mileage in your District of:

   State roads: ____________
   Off-system roads: ____________

2. What is the total mileage of the different road categories existing within your district? (check all that apply)

Rural principal	______	Urban freeway	______
Rural minor arterial	______	Urban expressway	______
Rural major collector	______	Urban principal arterial	______
Rural minor collector	______	Urban minor arterial	______
		Urban collector	______

   Other (please specify) ____________

Please FAX this Survey Form back to FIU/LCTR at (305) 348-4057 or by MAIL to:
Florida International University
Department of Civil & Environmental Engineering
Center for Engineering and Applied Sciences
University Park Campus, EAS 3685
Miami, Florida 33199

3. What percentage of the off-system road network is surveyed by:
   FDOT: ____________________
   MPO: ____________________
   City: ____________________
   Public Works: ____________________
   Others (specify): ____________________

4. If your district does collect data for off-system roads, please describe any existing coordination and data sharing between different agencies.

____
____
____
____
____
____
____
____
____

5. Please provide information about the person to contact from; DOT, MPO, public works, and other agencies or firms handling off-system roads data.

Name: ____________________
Agency/Company: ____________________
Phone: ____________________ Fax: ____________________
E-Mail: ____________________
Mailing Address: ____________________

Name: ____________________
Agency/Company: ____________________
Phone: ____________________ Fax: ____________________
E-Mail: ____________________
Mailing Address: ____________________

Name: ____________________
Agency/Company: ____________________
Phone: ____________________ Fax: ____________________
E-Mail: ____________________
Mailing Address: ____________________

6. How are the number and the location of *off-system* segments (traffic breaks) selected for data collection?

7. How are the number and the location of *state road* segments (traffic breaks) selected for data collection?

8. How frequently are the *off-system* segments (traffic breaks) surveyed?

9. How frequently are the *state road* segments (traffic breaks) surveyed?

10. In which months are data usually collected for *off-system* roads?

11. In which months are data usually collected for *state roads*?

12. During what time periods are data usually collected for *off-system* roads?

____

____

13. During what time periods are data usually collected for *state roads*?

____

____

14. What type of information is collected, compiled, analyzed, and reported for *off-system* roads? You may choose to send in the form(s) used for off-system data collection.

____

____

____

____

15. What type of information is collected, compiled, analyzed, and reported for *state roads*? You may choose to send in the form(s) used for off-system data collection.

____

____

____

____

16. Please check the data collection method used for *off-system* roads:

Permanent stations	________
Portable equipment	________
Manual counts	________
Video recording	________
Other (specify)	________

17. Please check the data collection method used for *state roads*:

Permanent stations	________
Portable equipment	________
Manual counts	________
Video recording	________
Other (specify)	________

18. If your district office makes estimation of AADT for *off-system* roads, what **data** and **method** do you use?

Data

Vehicle counts	______
Vehicle classification	______
Vehicle speeds	______
Intersection delays	______
Travel time studies	______
Vehicle weights	______
Statistic records	______
RCI	______
Other (specify)	______

Method (please explain):

19. What **data** and **method** do you use for estimating AADT of *state roads*?

Data

Vehicle counts	______
Vehicle classification	______
Vehicle speeds	______
Intersection delays	______
Travel time studies	______
Vehicle weights	______
Statistic records	______
RCI	______
Other (specify)	______

Method (please explain):

20. Which physical, design, and functional criteria are considered for off-system roads classification. Please explain.

For surveyed roads:

For un-surveyed roads:

21. For the AADT estimates/forecast data, what accuracy level is required?

Off-system roads		**State roads**	
5%	______	5%	______
10%	______	10%	______
15%	______	15%	______
20%	______	20%	______
Other	______	Other	______

22. What are the major problems encountered when collecting data for off-system roads?

____
____
____
____
____
____
____
____
____
____

23. How many people in the district office are involved in off-system data collection, analysis, and reports? Please estimate the full-time employees engaged in the following tasks:

Data collection ____________
Data analysis ____________
Reporting data ____________

24. Estimate the percentage (%) of effort in the following tasks:
Data collection ____________
Data analysis ____________
Reporting data ____________

25. Please provide the following documents:
- Copies of data collection manuals or guidelines used for off-system roads
- GIS coverage in Arc/Info export format and related databases for the following: TAZ, Roads, Land use
- Z data 1 and 2 (digital files)
- Historical traffic counts for the past 20 years for state roads and off-system roads (in digital & printout format)

APPENDIX B. CONTACT PERSONS AT FDOT DISTRICTS

District 1: Sandra L. Thompson, Supervisor Planning Department
Planning - Transportation Statistics
801 N. Broward Street
Barton, FL 33830
Tel: 941-519-2352

District 2 : Joye Brown
Planning Department
P.O.Box 1089, MS 2014
Lake City, FL 32056
Tel: 904-752-3300

District 3: Paul Day
Planning Department
P.O. Box 607
Chipley, FL 32428
Tel: S.C. 767-1539

District 4: William L. Cross, Asst. District Planning Manager
Systems Planning Office
3400 West Commercial Blvd., 3rd Floor
Ft. Lauderdale, FL 33309-3421
Tel: 954-777-4601

District 5: John Kahl
Planning Statistics Office
719 S. Woodlands Blvd.
Deland, FL 32720
Tel: 904-943-5374

District 6: Rolando Jimenez
Planning Department
602 S.Miami Avenue
Miami, FL 33130
Tel: 305-377-5897

District 7: William Gardner
Planning Office
11201 N. McKinley Dr.
Tampa, FL 33612
Tel: 813-975-4834

APPENDIX C. CONTACT PERSONS AT FLORIDA COUNTIES

DISTRICT	COUNTY	NAME	PHONE	ADDRESS
1	CHARLOTTE	ROBERT JOHNSON	941-639-4676	
	COLLIER	GAVIN JONES	941-643-8300	2800 NORTH HORSESHOE DRIVE NAPLES, FL 34104
	DESOTO	JAMES LEIPORT	941-993-4816	200 LAST OAK ST. SUITE 203 ARACADIA, FL
	GLADES	JIM THREAWITHS	941-946-0771	
	HENDRY	LAMAR CARROL	941-675-5222	
	LEE	STEVE JENSEN	941-694-7600	
	POLK	DAVIS HYSLOP	941-534-6486	
	SARASOTA/MANATEE	BILL SPARROWHAWK	941-359-5772	7632 301 blvd. Sarasota, FL. 34243
2	ALACHUA	MARLEY SANDERSON	352-955-2200	2009 NW 67 PL. SUITE A. GAINESVILLE, FL 32653
	CLAY	WANDA FORREST	904-269-6375	
	DUVAL	CLAVIN L. BURNEY	904-630-1903	128 E. FORYTH ST. SUITE 700. JACKSONVILLE, FL 32202
	ST. JOHNS	JEFF ALEXANDER	904-363-6350	9143 PHILIPS HIGHWAY, SUITE350 JACKSONVILLE, FL 32256
3	BAY	GARY CRAMER	1800-226-8914	PO BOX 486 PENSACOLA, FL 321593
	ESCAMBIA	GARY CRAMER	1800-226-8914	PO BOX 486 PENSACOLA, FL 321593
	LEON	MARLON BROWN	850-891-8614	CITY HALL TALLAHASSEE, FL 32301
	LIBERTY	MIKE DONAVAN	850-488-6211	
	OKALOOSA	WILEY PAGE	1800-226-8914	PO BOX 486 PENSACOLA, FL 321593
	WALTON	WILEY PAGE	1800-226-8914	PO BOX 486 PENSACOLA, FL 321593
4	BROWARD	JACK BURIE	954-357-6649	115 S. ANDREWS AVE. FT. LAUDERDALE, FL 33301
	INDIAN RIVER	JACOB RIGER	407-567-8000	1840 25TH ST. VERO BEACH, FL 32960
	MARTIN	DON DONALDSON	561-288-5495	
	ORANGE/OSCEOLA/SEMINOLE	SUMONE BABB	407-481-5672	315 EAST ROBINSON ST. SUITE 355 ORLANDO, FL 32801
	PALM BEACH	RANDY WHITFIELD	561–684-4170	PO BOX 21229 WEST PALM BEACH, FL 33416
	ST. LUCIE	DAVID GINNS	561-462-1576	2300VIRGINIA AVE. FT. PIERCE, FL 34982
5	BREVARD	JIM LIESENFELT	407-635-7815	
	FLAGLER	KENNETH KOCH	904-437-7484	1200 E. MOODY BLVD. #2 BUNNEL, FL 32110
	MARION	GREG SLAY	352-629-8529	PO BOX 1270, OCALA, FL 34478
	SUMTER	DALE PARRETT	352-793-0240	319 E. ANDERSON AVE. BUSHNELL, FL 33513
	VOLUSIA	DARLA ZAKALUZNY	904-322-5160	1190 PELICAN BAY DR. DAYTONA BEACH, FL 32110
6	DADE	MICHEAL MOORE	305-375-4507	111 NW 1ST ST., 9TH FLOOR MIAMI, FL 33128

DISTRICT	*COUNTY*	*NAME*	*PHONE*	*ADDRESS*
	MONROE	MARLON BROWN	305-891-8641	5757 BLUE LAGOON DR. SUITE 170 MIAMI, FL 33126
7	CITRUS	BETY CESTER	352-527-5239	
	HERNANDO	DENIS DIX	352-754-4057	20 N. MAIN ST. ROOM 262 BROOKSVILLE, FL 34601
	HILLSOROUGH	LUCIE AYER	813-272-5940	601 E. KENNEDY BLVD., 18TH FLOOR TAMPA, FL 33602
	PASCO	DOUG UDEN	813-847-8193	7530 LITTLE RD. NEW PORT RICHEY, FL 34654
	PINELLAS	RAMON SOLIS	813-464-4751	14 S. FORT HARRISON AVE. CLEARWATER, FL 34616

APPENDIX D. 1995 COUNTY LEVEL DATA IN 67 COUNTIES IN FLORIDA

MODEL	County	Population	County Road Mileage	Vehicle	Municipalities	Labor Force	Per Capita Income	Taxable Sales
S	Alachua	196,106	1,010	57,199	106,061	119,367	$18,424	$1,916,008,000
	Baker	20,153	220	5,050	4,576	5,846	$14,158	$66,054,000
S	Bay	142,690	503	48,965	74153	74,121	$16,852	$1,698,365,000
S1/R2	Bradford	24,182	176	7,061	6,432	7,168	$13,049	$126,065,000
S	Brevard	450,646	1,200	121,548	213,185	207,726	$18,915	$3,865,377,000
	Broward	1,412,165	2,377	214,513	519,230	669,665	$23,840	$17,185,339,000
	Calhoun	11,858	193	2,469	2,922	3,359	$11,790	$49,384,000
S	Charlotte	129,381	312	34,344	11,873	40,841	$18,012	$1,091,957,000
	Citrus	107,333	266	31,112	10,619	32,222	$15,295	$691,033,000
S	Clay	124,431	338	35,246	16,203	37,413	$18,134	$920,298,000
S	Collier	181,381	821	56,110	20,958	91,390	$29,237	$2,795,615,000
S/R	Columbia	48,376	608	16,907	10,375	18,413	$14,552	$423,906,000
S	Dade	2,031,336	2,524	690,359	753,360	1,080,823	$19,266	$21,427,878,000
S/R	De Soto	25,048	165	9,302	6,575	9,353	$15,043	$140,917,000
	Dixie	12,159	151	2,494	2,254	3,445	$10,334	$36,874,000
I	Duval	701,673	1,817	213,291	710,592	465,410	$19,820	$8,498,490,000
S	Escambia	273,804	809	80,011	62173	142,882	$16,899	$2,687,570,000
	Flagler	40,643	293	10,483	6,434	9,126	$14,845	$207,541,000
	Franklin	10,301	183	2,030	4,061	4,020	$14,458	$54,499,000
S/R	Gadsden	43,378	413	80,011	16,603	14,506	$13,712	$171,439,000
	Gilchrist	12,332	122	2,483	1,846	2,563	$12,622	$26,760,000
	Glades	7,665	228	1,001	1,552	1,659	$15,097	$15,130,000
	Gulf	13,390	133	2,380	5,955	4,692	$14,482	$630,202,000
	Hamilton	11,773	240	2,078	3,588	4,291	$11,876	$81,739,000
	Hardee	19,952	199	7,730	6,724	8,025	$15,490	$109,114,000
	Hendry	28,114	156	10,146	9,129	12,158	$17,174	$176,364,000
	Hernando	120,054	305	26,722	7,730	31,446	$15,251	$626,033,000
S/R	Highlands	74,507	382	23,149	18,367	26,236	$16,541	$571,296,000
	Hillsborough	884,608	1,744	263,784	326,262	539,118	$19,129	$11,467,950,000
	Holmes	17,520	245	4,049	3,966	4,236	$12,356	$50,788,000
	Indian River	97,144	377	28,980	35,259	42,650	$27,220	$881,921,000
S/R	Jackson	43,891	632	14,170	13,677	17,347	$14,949	$283,646,000
	Jefferson	12,950	327	2,806	2,746	3,882	$14,575	$34,039,000
	Lafayette	6,043	124	1,027	920	1,299	$13,098	$10,280,000
	Lake	180,160	688	56,014	56,169	60,061	$17,325	$1,327,379,000
S	Lee	375,381	578	114,707	135,597	170,472	$20,907	$4,590,853,000
S	Leon	213,917	563	60,420	133,731	142,214	$18,746	$2,284,439,000
	Levy	29,738	463	9,027	6,931	8,098	$13,062	$143,215,000
	Liberty	6,400	142	1,230	968	1,848	$14,199	$12,380,000
	Madison	17,231	364	3,197	4,712	5,618	$13,665	$54,541,000
S	Manatee	229,864	670	209,818	65,365	104,692	$21,009	$2,003,742,000
	Marion	226,678	826	74,983	48,911	84,428	$15,972	$2,006,548,000
	Martin	111,069	584	34,991	15,809	50,814	$30,256	$1,382,644,000
	Monroe	81,850	303	21,969	27,639	45,959	$23,582	$1,398,653,000
	Nassau	50,767	323	11,038	12,895	17,527	$19,771	$388,188,000
S	Okaloosa	163,707	646	52,219	60,855	88,947	$18,202	$1,610,518,000
	Okeechobee	30,222	232	10,254	4,997	9,945	$14,227	$227,625,000
T	Orange	749,631	1,543	252,891	241,907	550,142	$19,607	$15,132,506,000
S	Osceola	130,771	638	45,129	50,647	48,607	$15,379	$1,604,611,000
	Palm Beach	972,093	1,915	209,887	262,796	479,616	$32,230	$11,877,879,000
	Pasco	304,938	514	102,910	32,212	81,320	$16,176	$1,955,158,000
	Pinellas	870,884	1,042	231,763	487343	449,829	$22,798	$880,183,700
S	Polk	436,701	1,379	133,501	135,899	194,504	$16,858	$4,157,759,000

MODEL	County	Population	County Road Mileage	Vehicle	Municipalities	Labor Force	Per Capita Income	Taxable Sales
	Putnam	69,481	339	17,180	15,189	20,707	$13,972	$408,609,000
	Santa Rosa	104,110	509	26,503	14,111	30,201	$16,556	$438,982,000
S	Sarasota	295,942	588	88,383	87,170	154,868	$28,761	$3,408,275,000
S	Seminole	330,012	411	105,100	136,957	130,984	$20,846	$3,333,858,000
S	St. Johns	102,174	548	26,391	16,463	38,438	$24,797	$945,782,000
	St. Lucie	172,483	588	50,226	105,795	60,374	$15,773	$1,296,018,000
S/R	Sumter	34,788	369	7,926	8,559	8,746	$13,955	$174,546,000
	Suwannee	30,103	325	8,906	7,253	10,147	$14,345	$168,025,000
	Taylor	17,445	321	4,907	7,213	7,368	$13,690	$155,082,000
	Union	12,433	115	1,857	2,528	4,483	$10,398	$29,523,000
	Volusia	408,261	1,189	115,753	168,554	156,518	$16,991	$3,652,962,000
	Wakulla	17,111	183	3,276	676	4,489	$14,816	$45,245,000
S/R	Walton	33,615	489	5,769	6,827	11,177	$14,128	$316,743,000
	Washington	18,623	268	3,305				

Note: [1] Data used in state wide model
[2] Data used in rural model