Estimating Annual Average Daily Traffic for Non-State Roads in Louisiana

A Thesis

Presented to the

Graduate Faculty of the

University of Louisiana at Lafayette

In Partial Fulfillment of the

Requirements for the Degree

Master of Science

Charles W. LeBoeuf

Fall 2014

Estimating Annual Average Daily Traffic for Non-State Roads in Louisiana

Charles W. LeBoeuf

APPROVED:

Xiaoduan Sun, Chair Professor of Civil Engineering	Kenneth McManis Professor of Civil Engineering
Mohammad Khattak Professor of Civil Engineering	Mary Farmer-Kaiser Interim Dean of the Graduate School

Acknowledgements

First, I would like to give thanks to my thesis advisor and committee chair, Dr. Xiaoduan Sun. Her guidance has greatly helped me throughout the entire thesis process, and her knowledge in transportation engineering has also greatly helped. In addition, I thank her for providing me with the opportunity to be able to work on a project that can enhance my transportation knowledge.

I would also like to thank the other members of my thesis committee: Dr. Kenneth L. McManis and Dr. Mohammad J. Khattak. Their recommendations for improving my thesis are greatly appreciated, and I thank them for accepting the offer to be on my committee.

In addition, I would like to thank my research peer Mr. Subasish Das for assisting in the direction of the project and providing me with material that would be of great help in what to expect for my project.

Also, I would like to thank my parents for being supportive of me throughout my entire educational career and for making me the person I am today.

I am appreciative of the opportunity to be able to attend this University and be a student in the Department of Civil Engineering, both of which are interested in the academic development of their students.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Background

Annual average daily traffic, also known as AADT, or sometimes referred to as *traffic count data*, is defined as the total volume of vehicle traffic on a highway or road during a year divided by 365 days, and it is a measure used for transportation planning and transportation engineering purposes, which include: roadway geometric and pavement design (pavement thickness and width, horizontal curve radii, etc.), transportation forecasting (using current-year AADT to compare to an expected AADT in a future year), travel model validation, roadway safety design and analysis (estimating the number of crashes on a segment of roadway, design of highway safety devices, etc.), and air quality compliance, among others. Some example types of applications that require AADT as a data input include *SafetyAnalyst*, Highway Safety Manual (HSM) model functions, and several types of pavement design software (e.g. MichPAVE, KENSLAB, KENLAYER)

AADT data can be collected by either using automatic traffic recorders (ATRs) permanently located along a particular highway or road that give continuous traffic count data or portable temporary traffic recorders, which can be used for short-term traffic counts, generally for two to three days (1). Examples of each type of traffic counter are shown in **Figure 1**:

Figure 1: Permanent (Left) and Portable Temporary (Right) Traffic Counting Devices Images Courtesy of: Flickr Profile Rob Klug (Left) and City of Knoxville, Tennessee Traffic Engineering Division (Right)

While most State Departments of Transportation (DOTs) have collected and predicted AADT for the aforementioned transportation planning and engineering purposes (2), most non-state (municipal and county) agencies responsible for transportation planning and engineering do not have complete or even extensive AADT data, primarily due to budgetary constraints (high costs) for resources to collect AADT on non-state roads (3). The high cost of collecting AADT for every local road can be attributed to the large proportion of a roadway network that is non-state maintained, especially at the county level. Because of these limitations in obtaining adequate AADT data for local roads, methods for obtaining AADT data must be developed. These methods could include: regression models where several independent variables could be associated with AADT (e.g. demographic data), clustering and regression trees analyzing AADT for roadways with similar characteristics, neural networks, travel demand modeling at the tax parcel level, universal kriging, and vector regression with data-dependent parameters. Each method has particular advantages as well as disadvantages in estimating AADT for local roads, and some of these methods account for any potential changes in AADT (e.g. seasonally, weekly). Each research project and the results from the studies are explained further in more detailed literature reviews in *Chapter 2: Literature Review*.

Within the State of Louisiana, the Louisiana Department of Transportation and Development (DOTD) provides traffic count data for the 16,000-mile state-highway network on a relatively frequent basis (generally every three years), and most of the major cities in the state (e.g. New Orleans and Baton Rouge), as well as their respective Metropolitan Planning Organization (MPO), generally provide traffic count data for the locally maintained roads within a particular MPO study area. However, since most parishes and municipalities do not maintain traffic count data for roads within their jurisdiction due to the aforementioned costs associated with AADT collection, the DOTD has collected traffic count data on these roads in previous years. While more count stations are located on non-state roads than state-maintained highways, AADT data for local roads is not as adequate as on the state highways, generally due to these reasons: the limited availability of this data, the infrequent updating of the data, if at all, and in some locales, traffic count data on a particular road having not been counted in over twenty years. This results in traffic count data that is inadequate for use in current transportation planning and engineering. Since approximately three-fourths of all road mileage in Louisiana is maintained by local governments (parishes and municipalities), estimating AADT on non-state roads is important.

1.2 Objectives

As mentioned, estimating AADT for local roads is important in transportation planning and engineering; therefore, proper methodology in determining a way to estimate AADT for non-state roads in Louisiana is necessary since AADT data for these roads is not collected as frequently as for state highways. An emphasis is made in developing a methodology for rural roads as these roads constitute the majority of the statewide roadway network, but AADT estimation in small urban areas (smaller municipalities not within an MPO) is to be considered since the smaller municipalities do not have resources to collect AADT like the majority of the Parishes in the State. The identification of all variables that could be related to AADT estimation, including demographic and roadway information, as well as the exploration of the relationship between these variables and AADT, is important in developing a methodology for estimating AADT on local roads. Once the variables that are related to AADT have been identified, a reliable and practical methodology is to be developed for the estimation of AADT. Different types of model equations (linear, Poisson, etc.) are to be developed to finalize the selection of model type and can be based on several conditions. A validation of the model through the comparison of the actual collected traffic counts with the estimated AADT can show how reliable and practical the developed model is in estimating AADT. Due to the differences in the demographic, roadway, and AADT data throughout Louisiana, more than one final model to estimate AADT for local roads in Louisiana is to be developed to account for these differences throughout Louisiana.

Chapter 2: Literature Review

2.1 Background

Previous studies to estimate AADT focused on two types of approaches: mathematical model development and machine learning algorithms. Commonly used in engineering disciplines, mathematical models are representations, in mathematical terms, of the behavior of real devices and objects (4). Mathematical models can be developed through the use of linear regression, parcel-level trip generation, and spatial grids with the latter two being more related to this study. Although mathematical models are relatively simple to run, these models can only run at an aggregated level. Machine learning algorithms, which automatically learn programs from the available data, can figure how to perform important tasks by generalizing from example (5). Some types of machine learning algorithms include: clustering, support vector machines (SVR), fuzzy algorithms, and kriging methods. In addition, all of these studies mentioned the importance of AADT in transportation applications.

2.2 Mathematical Model Development

2.2.1 Regression

F. Zhao and S. Chung (3) used GIS and regression to estimate the AADT in Broward County, Florida, a suburban county located along the East Coast of Florida, north of Miami and east of the Everglades. Their report was based on previous studies, including one where a regression model for estimating AADT on non-state-owned roads in Broward County, Florida that included predictors such as: functional classification, the number of lanes, area type, auto ownership, presence of nearby non-state roads, and service employment. Initial predictors included: roadway characteristics, socioeconomic characteristics, accessibility to

Expressways, and regional accessibility to employment. Roadway characteristics, which were obtained from the Broward County MPO, included: the number of lanes on a roadway (as of 1998), the area type (the land use type e.g. residential, commercial), and the functional classification. Socioeconomic characteristics include: employment size along a corridor near a group of count stations, employment, population, and total dwelling units around a count station, and the employment and population around a count station aggregated using buffer sizes based on functional classification. Accessibility to expressways incorporate the minimum distance to an expressway (Interstates and Florida's Turnpike Enterprise Toll Roads (6)) access point from a count station, minimum travel time in minutes from a count station to an expressway access point based on minimum travel speeds for each type of road, the number of expressway access points within a four-mile radius from a count station, and a binary variable to account for roads connecting to an expressway. Regional accessibility to employment includes these variables: network distance to regional mean centers of employment and population, regional accessibility to population and employment centers, and regional accessibility to population and employment defined as the product of the regional accessibility measures. Based on the preliminary analyses, six independent variables were used to generate four regression models, and all four models showed a strong relationship between the independent variables and AADT as well as having no multicollinearity among the independent variables. The final results from this study showed that the more independent variables used, the better a model would perform, and the choice of a model would likely be based on the data processing cost, though this data is generally available and easy to process. However, this study did realize that the current models may not be adequate in meeting the need of engineering design or the calibration of travel demand models, but the performances of the models have shown improvements. Future work for this study includes the examination of spatial patterns of errors resulting from the different models and the reasons the errors occur.

2.2.2 Parcel-Level Trip Generation

T. Wang, A. Gan, and P. Alluri (7) developed a method to estimate AADT on local roads in Florida using a travel demand modeling method with a major component involving a parcel-level trip generation estimating the trips generated by each parcel. The interest of deploying *SafetyAnalyst* on all roads in Florida by the Florida Department of Transportation (FDOT) is a primary reason for this study since AADT is a required input for the program and FDOT does not have this data for local roads. This study, which attempted to improve on regression models that were developed by Q. Xia et al., examined several existing methods for estimating AADT in Kentucky, Alabama, Minnesota, and Indiana. It also included Broward County's regression models to estimate AADT. Other methods for estimating AADT were based on high-resolution satellite images and aerial photographs as well as machine learning algorithms. A study using travel demand modeling, although rare, was used in New Brunswick, Canada, and was included as a basis for this study. Parcel-level demand analysis incorporated four steps based on standard trip generation and trip assignment: network modeling defining the boundaries of the study area, parcel-level trip generation estimating the number of vehicle trips generated by each parcel, parcel-level trip distribution determining where each generated trip by each parcel will do, and parcel-level trip assignment predicting the routes the travelers will take to reach the traffic count sites on major roads which results in the estimated AADTs on local roads in the study area. Model development was performed by two development tools: ESRI's ArcGIS, which preprocesses the input data for the model, and Citilabs' Cube, which builds the highway network from the roadway shape file. To evaluate the model, a sample of count sites was selected to compare the actual to estimated AADT values. The parcel level trip generation greatly improved the estimated AADT versus the overestimated AADT resulting from the regression methodology. The reason for the overestimated AADT was due to the regression method's inability to recognize that the layout of local roads was meant for only minimal, if any, through traffic.

2.2.3 Alternative Methods to Traditional Sampling

W. Seaver, A. Chatterjee, and M. Seaver (8) developed a better understanding of traffic volumes on local roads in Georgia through the use of alternative methods for estimating AADT to the traditional sampling approach that is currently used for FHWA's Highway Performance Monitoring System (HPMS). In traditional sampling approach, the entire roadway network is broken into discrete segments having uniform characteristics of physical features and anticipated traffic volumes, and the road segments are stratified according to functional class and volume groups. Due to the difficulties in obtaining complete inventory data for local roads for sampling purposes as well as the substantial use of resources in preparing for traditional sampling, innovative alternative methods for estimating traffic counts were considered to be more practical. These innovative methods include: spatial grids, the development of mathematical models to estimate local road volumes, and the combination of different approaches. Spatial grids can be set up in each county and a selection of a sample of grids could be done through the use of a randomized procedure. A sample of local road segments within the selected grids then can be identified by using the same randomized procedure again. Mathematical models, which have been a common practice for transportation planning and can be used in estimating existing traffic volumes for reducing costs associated with a traffic volume count program, can be fairly simple and are based on “trend analysis.” In addition, since traffic volumes on rural local roads do not typically undergo drastic changes over time and are usually quite low, the amount of data required for model development and subsequent adjustment is not expected to be very large. The combination of different approaches can be practical as well as cost-effective due to the varying characteristics in the roadway conditions and location. Four road types (Non-Atlanta urban areas, small urban areas, and all rural roads-paved or otherwise) and their characteristics were analyzed within 80 of the 159 counties in Georgia. The initial models, with a total of 45 variables considered, were poor in predictability, but a stratification of counties based on their location within or outside a Metropolitan area was used to anticipate differences in the amount of rural traffic. The developed models can be used to estimate AADT in counties in Georgia that were not included in this study, thus reducing the need for resources to collect AADT on rural roads within the state.

2.3 Machine Learning Approach

2.3.1 Universal Kriging

B. Selby and K. Kockelman (1) studied the use of Universal Kriging for Spatial prediction of AADT in unmeasured locations in Texas. A state department of transportation typically has a few hundred permanently located automatic traffic recorders (ATRs) in conjunction with tens of thousands of portable count stations for short-term count samples that can be spaced far apart due to limited resources. In addition, the FHWA requires counts on high volume roads to be collected every three years, while counts on other roads can be sampled every six years. An FHWA recommendation is that AADT estimates should be within ten percent of the observed AADT values. Universal kriging is a geostatistical technique used to harness known local conditions influencing count and road network spatial information about measured locations; this technique involves spatial interpolation as well as making use of local information (lane count, population, etc.) and drawing on residuals in prediction from nearby sites. Universal kriging is just one of three types of kriging; the other types of kriging are simple kriging and ordinary kriging. In simple kriging, the value of interest at a location is predicted directly from nearby values based on semivariogram, which depicts the spatial autocorrelation of the measured sample points, and a known global mean. In ordinary kriging, a slightly more complicated method is used, requiring the process to estimate an unknown mean as well as the semivariogram. However, since the global-meal assumption cannot be used for this project, universal kriging was used. Previous studies for both future year (using current and past traffic data to estimate counts at the same location at future dates) and current year (estimating counts at a location whose traffic flow have not been measured and uses data from nearby locations during the same time period) prediction methods were studied for this project and include: Box-Jenkins, neural network, nonparametric regression, Gaussian maximum likelihood, "support vector regression with data-dependent parameters," geographically weighted regression (weighted least squares), and restricted maximum likelihood (REML). In addition to universal kriging, Box-Cox transformation, a likelihood-maximizing power transform giving skewed data a more normal distribution, for all traffic counts was used. Weighted least squares (WLS) was chosen over REML due to the ease of implementation, comparable performance from J.K. Eom et el.'s (2006) work, and the fact that WLS does not require the assumption of the error term's distribution. The estimation of the model parameters was done by using a randomly selected collection of the data points from each regional sample analyzed. Although the model uses Box-Cox transformed AADT values, to work directly with AADT estimates, the reverse transformation was used. The models developed used data from the year 2005 in Texas, which includes both large metropolitan areas (Houston and Dallas-Fort Worth) as well as sparsely populated lands (primarily West Texas), and the sample counts, obtained from the Texas Department of Transportation (TxDOT) came from all types of roads in the state. The variables used include the speed limit, number of lanes, and functional class of the roadway segment obtained from TxDOT. A marginal preference for the exponential semivariogram was the result for the test on the subsets of data, and no strongly favored model was shown for the other regional data sets. The median percentage errors show an inconsistent bias hovering near zero on both sides, and the average absolute errors are very high for many of the subsets. Euclidean distance-based kriging fared about as well as network-based metrics, thus suggesting that the latter's complexity is not warranted in these applications.

2.3.2 Vector Regression

A study in Tennessee by M. Castero-Neto, Y. Jeong, M. Jeong, and L. Han (2) used support vector regression with data-dependent parameters to predict AADT within the state of Tennessee. The objective of the research was to evaluate the performance of a modified version of the support vector machine for regression (SVR) technique for forecasting AADT one year into the future without use of any external, or predictor, variables. The attention of SVR has been increasing due to its remarkable characteristics, good generalization performance, absence of local minima, and sparse representation of solution; however, the computation of adequate SVR parameters is crucial to the quality of SVR models developed. The quality and performance of SVR models depend on the settings of these three parameters: the type of kernel, the value of C, and the value of ε for the ε-insensitive loss function. For any particular type of kernel, the quality and performance of the SVR models is affected by the values of C and ε. In a previous study by D. Mattera and S. Haykin (1999), the proposed value of C being equal to the range of the output values (AADT) resulted in a non-robust approach to the outliers as later determined by V. Cherkassky and Y. Ma (2004), whose approach was based on training data that did not resort to resampling. An advantage of this approach was the possible robustness to possible outliers. For short-term forecasts, exponential smoothing (ES) methods (in this study, Holt-ES) can be very effective, but due to the forecast pattern being linear, these methods may not perform well for multiple-step ahead forecasts. For this study, Tennessee Department of Transportation (TDOT) data was used, and their database contains AADT data collected annually since 1985 for more than 10,000 count stations strategically located throughout all counties within the state. The AADT data was aggregated based on the county and the type of road (rural and urban). A total of 25 counties in Tennessee were selected, resulting in 50 time series, in which three forecast methods applied: SVR-DP, Holt-ES, and OLS-regression. The mean absolute percent error (MAPE) and root-mean-square deviation (RMSD) represented the performance of the models over the five predictions (AADT from 2000 to 2004). Results from this study, which are detailed in Table 4, explain that the average MAPE and RMSE for the SVR-DP technique was the lowest of the three techniques, while the OLS technique had the highest model performance test values.

2.3.3 Fuzzy Algorithm

Because AADT can be affected by seasonal changes (e.g. more traffic in the summer than in the winter), M. Gastaldi, R. Rossi, G. Gecchele, and L. Lucia (9) proposed an approach to estimating AADT for seasonal conditions. The Federal Highway Administration (FHWA) provides recommendations concerning traffic monitoring programs to transportation agencies based on portable and permanent count stations. The FHWA's procedure could be affected by three sources of error: day-to-day variations in traffic volumes as traffic volumes fluctuate over time, grouping road segments into significant groups since ATR sites could belong to more than one road group, and assigning the road segments, along which Short Period Traffic Counts (SPTCs) and Permanent Traffic Counts (PTCs) were obtained to the right road group because large errors between estimated and observed AADT can be the result of incorrect assigning of a road section to a road group. Even with the errors that can result from FHWA's current procedure, the proposed approach preserved the framework of the FHWA procedure and allowed analysts to deal with situations where road segments may appear to belong to more than one group and provide the degree of belonging to each group. The four step approach includes: the grouping of ATR sites with Fuzzy algorithm, assigning the road segment where the STCs are available, calculating the measures of uncertainty associated with assigning to road groups, and estimating AADT as a weighted average of STC volumes, adjusted by seasonal adjustment factors of the assigned road groups. This study used data obtained in the year 2005 at 50 ATR sites on the rural road network in the Province of Venice, Italy. To implement the model, three tasks were conducted: establishing road groups with the Fuzzy C-Means algorithm, developing artificial neural networks, and the calculation of AADT. Eight groups (five recreational and three commuter) were developed for the STCs, with AADT estimated for three timeframes: one complete week, the weekdays (Monday through Friday), and each day of the weekend (Saturday and Sunday). To analyze the accuracy of the resulting AADT estimates, MAPE and Standard Deviation of Absolute

Percent Error (SDAPE) were used. Some conclusions that can be made from this study include: AADT estimates based on weekdays are typically more accurate than those obtained on weekends, which can be explained through the noting that weekday traffic patterns are generally more stable than on the weekends, AADT estimated based on average daily volumes make a balance among estimates obtained from single day volumes and thus maximizing the information available from the STC, recreational roads have larger MAPE and SDAPE than commuter roads due to tourism along the Venice coast, and the Summer (July-August) and Winter (November-December) months have larger error values than the rest of year, particularly in the Spring (March-June). Future work for this report could include extending the work to consider the influence of the socioeconomic and land use characteristics of the environment of the particular road section in question.

2.3.4 Forecasting Errors through Clustering

M. Dixon of The National Institute for Advanced Transportation Technology (NIATT) at the University of Idaho (10) studied the effects of errors in AADT forecasting for highways in rural Idaho. Since funding for transportation projects is always an issue due to limited funding available for all transportation projects, making critical decisions in an informed manner is always important. In addition, forecasted AADT volumes are required for use in selecting transportation projects, but inaccurate traffic volume forecasts are responsible for additional costs associated with over and under design. The current practice for forecasting AADT in Idaho is based on annual growth rates representing the average percent increase in AADT volume per year, but the accuracy of the forecasts have come into question by transportation professionals. The methods for forecasting traffic counts evaluated in previous studies in this report include: time series forecasting methods, regression, neural networks, and clustering. The first three methods are models to forecast AADT, and clustering is a method combining groups of traffic count stations based on data similarities. The variables, which were assumed to be independent of each other due to the low correlations among each other, used for clustering were: functional class of the roadway, county population and its annual growth rate, and AADT. To easily explain the results of clustering, a classification and regression tree (CART) method is readily implemented, and this method allows the addition of many variables to create subsets of data having similar characteristics while reducing the variability in the dependent variable (annual AADT growth rate). In addition, a sensitivity analysis was conducted to verify the suitability of the CART method, as this method has not been used previously to forecast AADT, and the process in the sensitivity analysis was repeated eight times to establish a confidence interval of the mean error. The final calibrated regression tree incorporated traffic count data at automatic count stations from 1980 and 1990, and rural portable count locations were used to complete a validation test of the final regression tree. As the importance of this project was to assess the impacts of AADT forecasting errors in applications for transportation planning and design, two applications were chosen: the overlay thickness, which requires the equivalent single axle load (ESAL), and level of service (LOS), the effects of the errors were compared for both applications. Although the existing Idaho Transportation Department (ITD) method of estimating AADT was more accurate, the CART method was considered to be more promising due to its ease, the small amount of data in the calibration data sets, and the potential to update the ITD growth factors more frequently.

Chapter 3: Methodology

3.1 Data Collection

3.1.1 Parish Selection

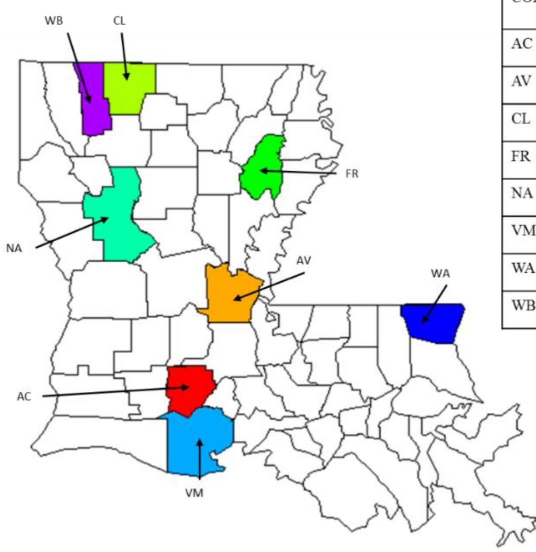
The selection of the parishes for the development of the model was based on the varying characteristics of the parishes in Louisiana (demographic, number of counts, etc.). The following considerations were made in the selection of the parishes for model development:

- Type of parish (Urban, Suburban, or Rural)
  - Population within the parish
  - Small urban area may be located within parish (e.g. Crowley)
- The location of the parish within Louisiana (both North and South Louisiana)
- Interstate accessibility (four parishes)
- The number of count stations within the parish (the more counts the more likely the parish would be selected).

No major urban or suburban parishes were used in the development of the model, as traffic counts are collected on a regular basis in the State's urbanized areas from both the DOTD and Metropolitan Planning Organizations (MPO). The eight parishes considered in the model development are:

- **Interstate Parishes**
  - Acadia
  - Avoyelles
  - Natchitoches
  - Webster
- **Non-Interstate Parishes**
  - Claiborne
  - Franklin
  - Vermilion
  - Washington

Of the four parishes selected that have Interstate access within the parish, Interstate 10 traverses Acadia Parish, Interstate 20 traverses Webster Parish, and Interstate 49 crosses Avoyelles and Natchitoches Parishes. Washington Parish is the only parish selected that does not have direct access to either Interstates or United States Highways. **Figure 2** shows the eight selected parishes as well as the location within Louisiana of each parish, and **Table 1** details the average population and the number of count stations of a particular parish.



CODE	PARISH	POPULATION (2010)	NUMBER OF COUNT STATIONS
AC	Acadia	61,773	1227
AV	Avoyelles	42,073	923
CL	Claiborne	17,195	681
FR	Franklin	20,767	810
NA	Natchitoches	39,566	895
VM	Vermilion	57,999	987
WA	Washington	47,168	1177
WB	Webster	41,207	970

Figure 2: Location, Population, and Number of Count Stations in Louisiana

Average Population (2010)	40,970
Average Number of Count Stations	959
DOTD Districts Selected and Number of Parishes within District	Dist. 03: 2 (Acadia, Vermilion) Dist. 04: 2 (Claiborne, Webster) Dist. 08: 2 (Avoyelles, Natchitoches) Dist. 58: 1 (Franklin) Dist. 62: 1 (Washington)

Table 1: Information on Selected Parishes

As shown in **Figure 2**, no parishes were selected in the more populated and urban/suburban Southeastern Louisiana except for Washington Parish. The population (2010 Census) in several of these parishes exceeds 100,000 (East Baton Rouge, Jefferson, Orleans, St. Tammany, Livingston, Ascension, and Terrebonne, in addition to Lafourche with a population just below 100,000), and the number of count stations in most of the parishes with smaller populations is considerably smaller than in most of the remainder of the State. The population documented in the 2010 Census and the number of count stations in all sixty-four parishes in Louisiana is detailed in Appendix A.

3.1.2 Roadway Data

3.1.2.1 AADT

The DOTD collects traffic counts at 5,067 permanent or portable count locations on the state-maintained highways and has provided 43,755 counts on non-state roads throughout Louisiana. On most state highways, counts have been collected within the last three years; however, no traffic counts have been collected on non-state roads since 2011, and few counts have been collected after 2006. Local AADT data is provided in a non-state dataset by the DOTD, and Appendix B describes the attributes of this table. **Figure 3** details the locations of non-state counts within Acadia Parish.

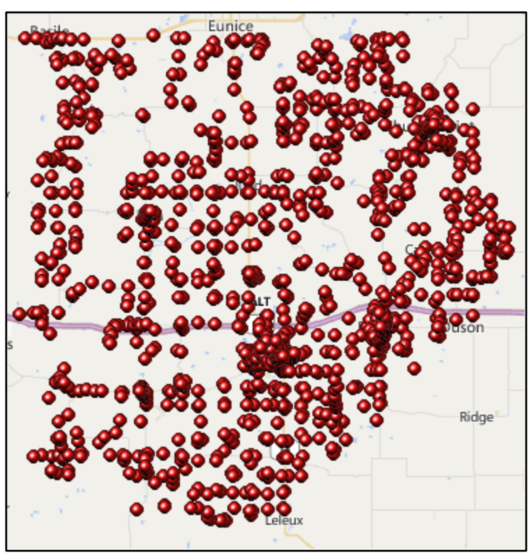


Figure 3: Locations of Counts within Acadia Parish

3.1.2.2 Roadway Network

The Louisiana highway and road network consists of 61,093 miles, including state-maintained highways, parish roads, and city streets, and the statewide network is shown in **Figure 4**.

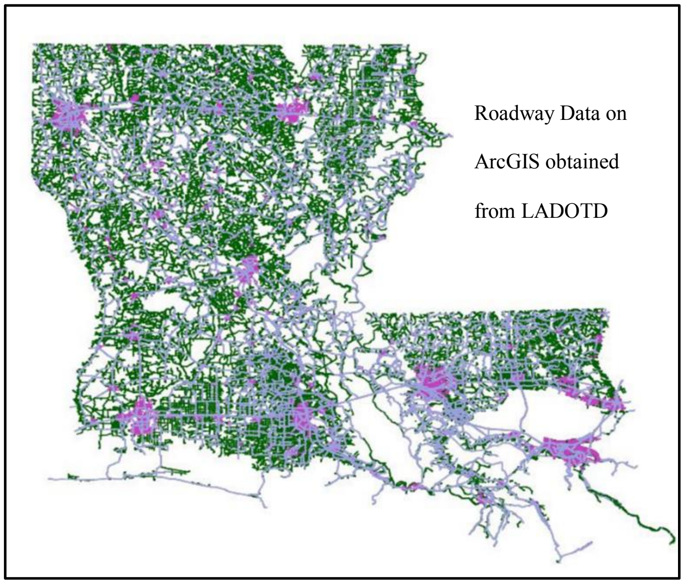


Figure 4: Louisiana Roadway Network

In **Figure 4**, state-maintained highways are depicted in blue, parish-maintained roads are depicted in green, and municipal owned streets are depicted in purple. Local roads account for approximately seventy-three percent of all roadway mileage in Louisiana, totaling over 44,000 miles. Appendix C details the attributes for the statewide roadway network dataset. A notable fact about Louisiana's roadway network is that Louisiana ranks tenth in the nation in the proportion of highways and roads that are state-maintained, primarily due to several miles of state highways that only serve a local purpose; because of this, the DOTD has initiated a voluntary highway transfer program allowing parishes and municipalities to take ownership of a particular road with the State funding any construction-related costs on that particular road before transfer to the parish or municipality (11).

3.1.2.3 Interstates and Major Highways

Two of the variables to be considered in the model determination are related to the shortest distance to Interstate and major state highways, and both of these types of highways have to be established before deriving a shortest distance from the count location to a particular highway. **Figure 5** shows the Interstates within Louisiana.

Figure 5: Interstates within Louisiana
Image Courtesy of Louisiana DOTD

Although six major Interstates and six Loop/Spur Interstates exist within Louisiana, only three Interstates were considered in the study, including:

- Interstate 10, the east-west Interstate through South Louisiana from the Sabine River at the Texas State Line to the Pearl River, part of the State Line with Mississippi, passing through Lake Charles, Lafayette, Baton Rouge and New Orleans along its 274 mile trek through Louisiana

- Interstate 20, the east-west Interstate through North Louisiana between Texas and Mississippi, passing through Shreveport/Bossier City and Monroe on its 190 mile journey in the State
- Interstate 49, the State's major north-south Interstate, with its southern terminus in Lafayette and passing through Alexandria and Shreveport before entering Arkansas.[1]

Although Interstates 12, 55, and 59 are also important in Louisiana, these Interstates are located in urban/suburban New Orleans and Baton Rouge, as well as most of the Loop and Spur Interstates with the exception of Interstates 210 (Lake Charles) and 220 (Shreveport/Bossier City). Future Interstates within Louisiana include the North and South extensions of Interstate 49 and an all-new Interstate 69 in Northwestern Louisiana (12, 13, and 14).

The second variable (distance to major highways) is required as not all areas in Louisiana are directly accessible by Interstates, as shown in **Figure 5**. Notable areas in Louisiana without Interstate access are Northeastern Louisiana (only east-west Interstate 20 crosses this part of Louisiana) and South-Central Louisiana (although Interstate 49 is planned to be extended through this area of the State). However, all parishes do have major highway access, and most areas within the State without nearby Interstate access have access to four-lane major highways, including:

- United States Highway (Hwy.) 90 between Lafayette and New Orleans (future Interstate 49 South corridor), with portions of the highway now controlled-access,

[1] Segments of Interstate 49 are under construction between Interstate 220 in Shreveport and the Arkansas State Line. The segment between Interstate 20 and Interstate 220, the "Inner City Connector", is in the planning stage as of October 2014. (14 and 15).

including around New Iberia, western St. Mary Parish, and from Morgan City to Raceland bypassing Houma-Thibodaux

- United States Hwy. 165 between Iowa and the Arkansas State Line; passing through Alexandria and Monroe
- United States Hwy. 167 from Alexandria to the Arkansas State Line; passing through Ruston
- United States Hwy. 171 from Lake Charles to Shreveport; passing through DeRidder, Leesville and Fort Polk
- United States Hwy. 425 from the Mississippi State Line near Natchez, Mississippi to the Arkansas State Line.

Most of these highways are north-south since Interstate 49 is the only Interstate that traverses north-south through most of Louisiana. Some of the other important United States and State highways considered in the study include:

- United States Hwy. 71, the main north-south United States highway through Western Louisiana, passing through Alexandria and Shreveport
- United States Hwy. 79 from Minden to the Arkansas State Line, passing through Homer
- United States Hwy. 80, north Louisiana's primary east-west United States highway
- Louisiana Hwy. 1, the State's longest highway of any classification (Interstate, United States, State), from Grand Isle to the Texas State Line in far Northwestern Louisiana
- Louisiana Hwy. 2, North Louisiana's primary east-west State highway
- Louisiana Hwy. 9, from Natchitoches Parish to Homer in Claiborne Parish

- Louisiana Hwy. 14, from Lake Charles to New Iberia passing through eastern Calcasieu, Jefferson Davis, Cameron, Vermilion, and Iberia Parishes.

The determination of a major highway for the *shortest distance to major highwa*y analysis within a particular parish is based on the following types of highways:

- United States Highways (e.g. United States Hwy. 90), since these highways are the highest-order highways in Parishes without Interstate access
- Trans-parish (i.e. parish line to parish line) State Highways that serve as a relatively direct connection to the neighboring parishes, which is considered especially in parishes without either Interstate or United States highway access
- State Highways with a terminus (end point) within the study parish and serving as the main route to the Parish Seat of a neighboring parish.

Most of the State Highways considered for analysis are one or two-digit highways (e.g. Louisiana Hwy. 1 or Louisiana Hwy. 14), since the majority of these highways carry more traffic than other state highways. Nonetheless, since many of Louisiana's state-maintained highways serve only a local importance themselves, these highways were not considered as major highways when determining the shortest distance from a count station to a major highway access point due to their lower-level importance in the state highway network. Appendix E shows the types of major highways that were considered.

<u>3.1.2.4 Deriving Shortest Distance</u>

Using ArcGIS, the shortest distance between a count location and access point to an Interstate or major highway can be determined. The ArcGIS feature used was the *Closest Facility Analysis*, which requires the following:

- **Incidents**, which for this analysis, is the location of the count stations
- **Facilities**, which for this analysis, is either where an on-ramp merges with the Interstate mainlanes or at an intersection of a local road with a major highway.

In addition, to prevent any portion of the route between the count station and the Interstate from accessing the Interstate where no actual access exists (i.e. where a street crossed the Interstate at a grade-separated non-interchange), *point barriers* were implemented along an Interstate where a local road crosses the Interstate or where the route would have a possibility of traveling the wrong way on a divided major highway. This was done to prevent wrong-way traveling and accessing an Interstate improperly. Because the route from the count stations to the nearest major state highway will not include any major highways (divided or undivided) or Interstates, point barriers were not needed for the *shortest distance to major highway* analysis. The step-by-step process in calculating the shortest distance from the count station to an Interstate or major highway is as follows:

1. Load all count stations as Incidents
2. Load all access points to the Interstate as Facilities
3. Run the program by clicking SOLVE, which gives the shortest paths to an intersection with major highway or Interstate on-ramp as "routes."

The *point barriers* must be loaded for the *shortest distance to Interstate* analysis due to the aforementioned possibilities that the "route" will access the Interstate at a location where access to the Interstate is not allowed (route suddenly accessing grade-separated intersection) or the "route" travels the wrong-way to an Interstate access point. An example of the shortest route between a count station and Interstate or major highway access is shown in **Figure 6 and Figure 7**.

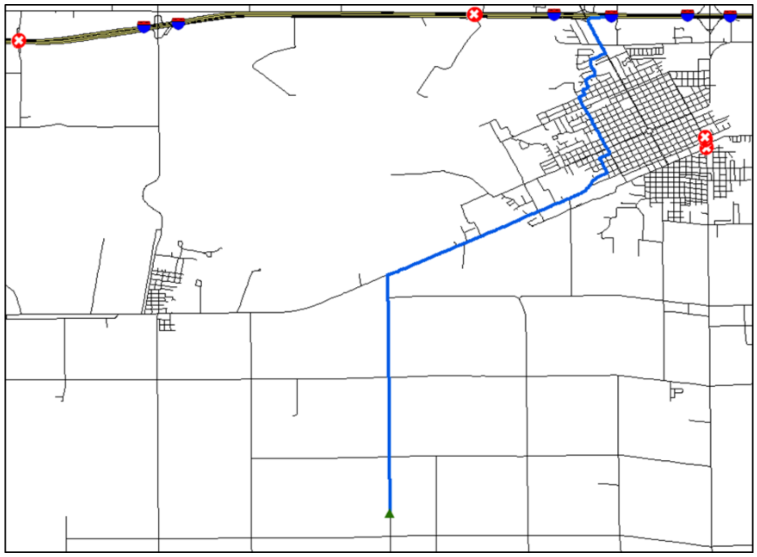


Figure 6: Example Shortest Distance to Interstate Analysis in Acadia Parish

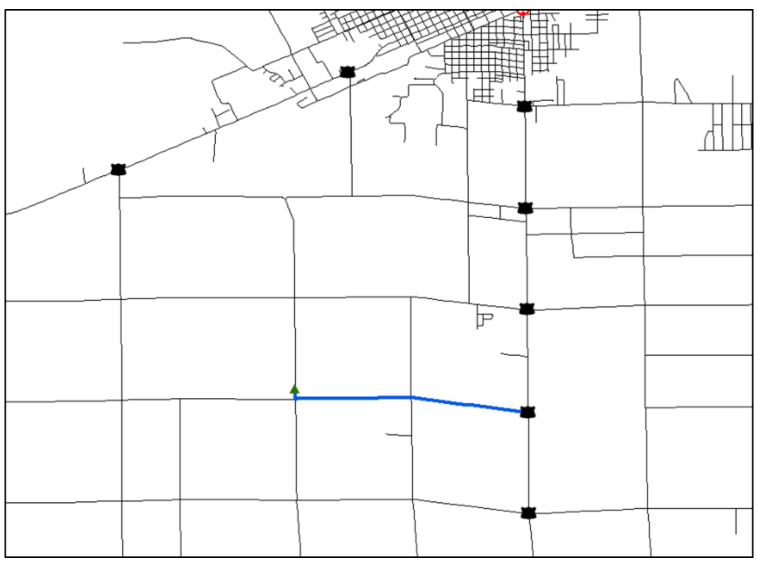


Figure 7: Example Shortest Distance to Major Highway Analysis in Acadia Parish

In both of the above figures, the routes from the count station (green triangles) to an Interstate/major highway access point (red circles) are depicted in blue along the roadway network. The *point barrier* feature is shown as with a white "X" inside a red circle, as shown in **Figure 6**. However, some routes between the count station and Interstate/major highway access may not have been determined due to these reasons: the location of the count station not locking to the roadway network, or the count station is located along a segment of disconnected network, typically near parish lines. Typically, manual measurement to the shortest path by calculating the distance from the problematic count station location to a count station where a "route" to major highway or Interstate access was determined.

3.1.3 Census Data

3.1.3.1 Geographic Units

All Census demographic and geographic data, updated as of the most recent Census in 2010, was obtained from two sources: the demographic information from the American FactFinder, and the block shapefiles from *TIGER.* Census data is subdivided into three units within each parish: Tracts, Block Groups, and Blocks (16, 17, and 18). Each of these units is further detailed below:

- **Tract**: The highest-level geographic unit, relatively permanent statistical subdivisions of a parish, generally defined to contain 1,200 to 8,000 people, identified with an integer number of up to four digits, and special codes exist for special land-use tracts with little or no population (9800s) or to cover large bodies of water (9900s)
- **Block Group:** The intermediate-level geographic unit, the division of tracts and clusters of blocks, generally defined to contain 600 to 3,000 people, identified as first digit of the block code (for example if a particular tract has blocks 2000, 2001, 2002, etc., then those blocks belong to ***block group 2*** of that particular tract)
- **Block:** The lowest-level geographic unit, the division of block groups, generally small statistical areas bounded by visible features such as roads, streets, small bodies

of water, or railroad tracts, all blocks are numbered between 0000 and 9999, and blocks beginning with zero are water-only blocks (i.e. 0XXX).

A map comparing two of the Census geographic subdivisions for Acadia Parish is shown in **Figure 8**.

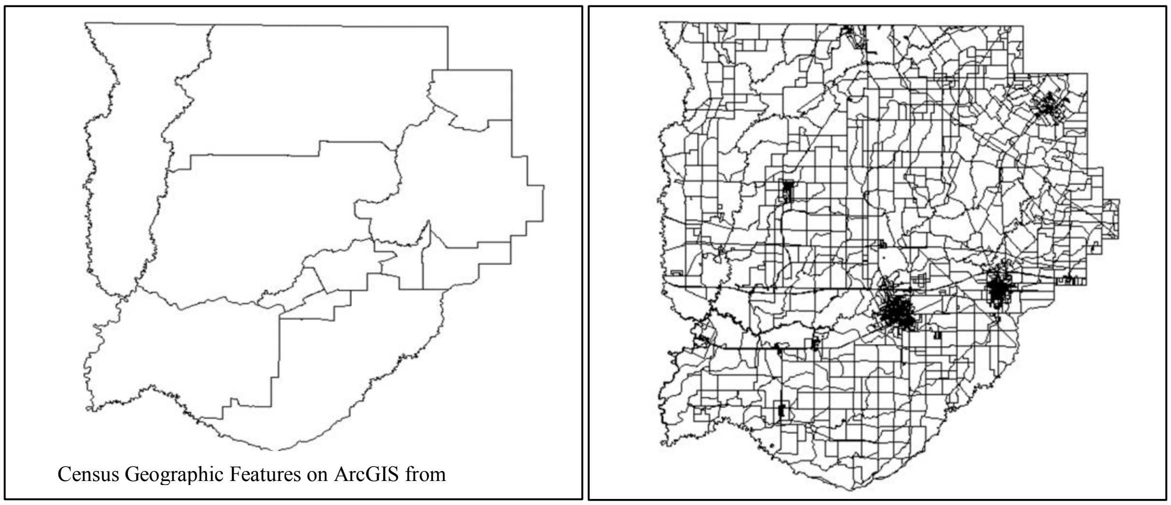


Figure 8: Census Tracts (left) and Census Blocks (right) for Acadia Parish

The amount of available demographic and economic data from the Census website is based on the geographic subdivision of each parish, with more readily available data at the Tract level than the lower two levels. While data at the Census Block level is more accurate than at the Tract level, the amount of available demographic and socioeconomic data at the Block level is considerably lower. The shapefiles obtained from *TIGER* show the shape of each Census subdivision and its geographic attributes. Appendix F gives an explanation to these attributes.

<u>3.1.3.2 Demographic and Socioeconomic Data</u>

Demographic and socioeconomic data for each Census geography in 2010 obtained from Census' *American FactFinder* include, but are not limited to, the following: population, households, employment, and travel time to work. The Census demographic data sets for the population and number of households in the study parish include both rural and urban subtotals within the Census geographic subdivision, which was used in the data processing to separate into rural and urban data sets.

3.2 Data Processing

3.2.1 Merging Census Geographic and Demographic Datasets

The Census geographic and demographic data is initially provided in separate data sets, and before any models can be developed, these sets must be merged. This three-step process was used to merge the two data sets:

1. Add both data sets to ArcGIS
2. Use ArcGIS's *Join* feature to merge the tables, with the required attributes from both tables: the GEOid from the Census geographic shapefiles and the GEOID2 from the Census demographic data
3. The merged table has both demographic and geographic data

This merged dataset can be used to determine the demographic and socioeconomic information of the Census geographic subdivision. However, if the data is not available for a particular geographic unit, then Longitudinal Employer-Household Dynamics is to be used to collect the data not available.

3.2.2 Merging Census and Roadway Data

Once the Census data sets have been merged, then this merged Census data set is combined with the AADT data set to create a data set of all interested variables that can be used in model development. **Figure 9** shows the process in merging data, with Acadia Parish as an example (The remaining counts are shown in parenthesis).

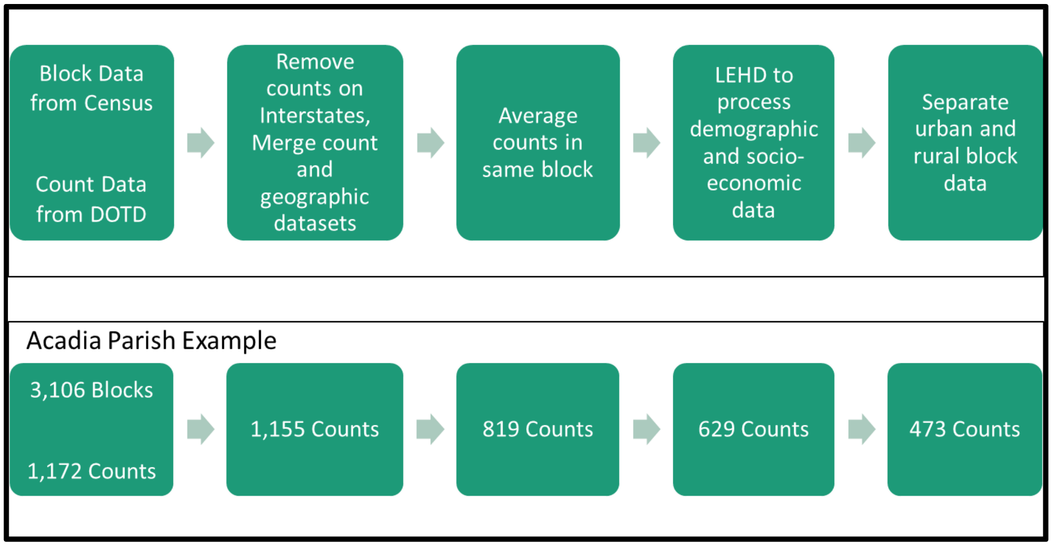


Figure 9: Data Merging Process for Acadia Parish

Initially, the geographic dataset includes information on all Blocks within a particular parish (3,106 Blocks in the example), while the AADT dataset includes all non-state and some state counts in a particular parish (1,172 counts). Because some counts in the non-state AADT database are on major state highways or Interstates, those counts are removed (1,155 counts remaining); following the removal of those counts, LEHD is used to locate which Block the count is located in. Because more than one count may be located within a particular Block, the third step averages the observed AADT and the distance to an Interstate or major highway of the count stations within a particular Census block (reducing to 819 counts). The fourth step merges the socioeconomic datasets that were collected from the LEHD datasets, which are available from the Census (19); this data is merged with the LEHD Residential Area Characteristics (RAC) economic (629 counts remaining). The final step in creating a data set with the roadway, demographic, and socioeconomic information is to remove any count stations that are located in urban areas (dataset containing 473 counts).

Afterwards, the rural dataset is processed further, and counts containing these three attributes are retained in the final dataset:

- Counts greater than one-tenth of a mile from a major highway
- Population of Block greater than five
- Observed AADT less than 2,000

Table 2 shows the changes from this data processing in an initial model development for Acadia Parish as an example.

Attribute	MAPE	MAPE Change	Number Remaining	Percent Remaining
Original	551.6%	-	473	-
Distance greater than 1/10 mile from major highway	449.0%	18.6%	346	73.1%
Block Population Greater than 5	440.2%	2.0%	315	91.0%
ADT less than 2,000	379.1%	13.9%	309	98.1%

Table 2: Data Processing Example for Acadia Parish

In this example, the Mean Absolute Percentage Error (MAPE) improved between the initial and final rural datasets, and about two-thirds of the original rural dataset was retained in the data processing. The formula for calculating MAPE is detailed more in **Chapter 4: Results**.

3.3 Model Development

3.3.1 Model Variables

After all the required data had been collected and processed, the variables that would be used in the model development were determined, including the demographic, socioeconomic, and the roadway network data which are further detailed below:

- *TOTAL_POPULATION*: the total population of the Census block that the count station is located in
- *TOTAL_JOBS*: the number of jobs within the Census block
- *POPULATION_TO_JOB _COMPARISON*: a factor to compare if the Census block was more residential (larger number of households) or more commercial (larger number of jobs); if the population was greater than the number of jobs, then a value of 2 was used, otherwise 1
- *DISTANCE_INTERSTATE*; the shortest distance (in miles) between the count and Interstate access point (on-ramp merge with mainlanes)
- *DISTANCE_MAJOR HIGHWAYS*; the shortest distance (in miles) between the count and intersection with major highway

Because of the variations in the attributes for each parish (e.g. demographic, Interstate access), more than one final model was developed, with at least one for all parishes with Interstate access within the Parish and at least one for parishes not having Interstate access within the parish.

3.3.2 Creating an Initial Model

JMP ("Jump") software, a Graphical User Interface (GUI) software developed by the Statistical Analysis System (SAS) Institute, was used for model development and selection. Poisson modeling would be the final model selected. The results generated from JMP were used to determine the type of model to use and is shown in **Figure 10.**

Generalized Linear Model Fit

Response: Adt_Avg
Distribution: Poisson
Link: Log
Estimation Method: Maximum Likelihood
Observations (or Sum Wgts) = 309

Whole Model Test

Model	-LogLikelihood	L-R ChiSquare	DF	Prob>ChiSq
Difference	3147.27037	6294.541	5	<.0001*
Full	46044.2596			
Reduced	49191.53			

Goodness Of Fit Statistic	ChiSquare	DF	Prob>ChiSq
Pearson	115342.3	303	<.0001*
Deviance	89940.43	303	<.0001*

AICc
92100.797

Effect Tests

Source	DF	L-R ChiSquare	Prob>ChiSq
Distance_Interstate	1	1298.105	<.0001*
Distance_State_Hwys	1	1163.5081	<.0001*
Total_Population	1	135.81069	<.0001*
Total_Jobs	1	442.20621	<.0001*
Greater_Less_Median	1	1.2757168	0.2587

Parameter Estimates

Term	Estimate	Std Error	L-R ChiSquare	Prob>ChiSq	Lower CL	Upper CL
Intercept	5.5877835	0.0143518	112280.43	<.0001*	5.5596363	5.6158942
Distance_Interstate	-0.022865	0.0006429	1298.105	<.0001*	-0.024126	-0.021606
Distance_State_Hwys	0.1004742	0.0029045	1163.5081	<.0001*	0.094777	0.1061626
Total_Population	0.0014217	0.0001215	135.81069	<.0001*	0.0011833	0.0016598
Total_Jobs	0.0065174	0.000308	442.20621	<.0001*	0.0059133	0.0071206
Greater_Less_Median	0.0090746	0.0080349	1.2757168	0.2587	-0.006672	0.0248246

Figure 10: Model Results from JMP

A major consideration in model selection is the use of probability tests (Chi-Square-, p-, and/or t-tests) to determine if a model and its parameters are statistically significant. If the probability of being greater than a particular probability test is greater than 0.05 (95% confidence interval or two-standard deviations), the model or parameter estimate is considered not to be statistically significant. Initial models studied include: OLS regression, exponential, Poisson, and negative binomial. Since AADT data is discrete, only the Poisson and negative binomial models would be explored further. The Poisson model would be selected as that model was statistically significant (using the aforementioned probability tests). The basic equations for Poisson models are:

$$\ln(AADT) = (\, c + \sum k_i x_i)$$

$$AADT = e^{(\, c + \sum k x_i)}$$

Where:

x_i is the i^{th} independent variable of the model

c is the intercept of model function

k_i is the coefficient of independent variable

The first equation is a generalized linear model as the equation is in terms of the natural-log of the AADT. The first equation must be converted to AADT by taking the exponential of the linear portion of the first equation, resulting in the second equation.

3.3.3 Results of Initial Poisson Models

Using the JMP software, ten initial models were developed: eight Parish-specific models and two models combining the data from the four Interstate and four non-Interstate parishes, which are shown in **Table 3 and Table 4** as well as **Figure 11 to Figure 20.**

MODEL TYPE	FUNCTION
Acadia	*AADT =* exp [*5.587784 – 0.022865 * (DISTANCE_TO_INTERSTATE) + 0.100474 * (DISTANCE_TO_MAJOR_HIGHWAYS) + 0.001422 * (TOTAL_POPULATION) + 0.006517 * (TOTAL_JOBS) + 0.009075 * (POP_TO_JOB_COMPARISON)*]
Avoyelles	*AADT* = exp [*5.274059 + 0.006048 * (DISTANCE_TO_INTERSTATE) – 0.015370 * (DISTANCE_TO_MAJOR_HIGHWAYS) + 0.005035 * (TOTAL_POPULATION) – 0.004790 * (TOTAL_JOBS) + 0.158717 * (POP_TO_JOB_COMPARISON)*]
Natchitoches	*AADT* = exp [*4.985668 + 0.003497 * (DISTANCE_TO_INTERSTATE) – 0.022803 * (DISTANCE_TO_MAJOR_HIGHWAYS) + 0.003313 * (TOTAL_POPULATION) + 0.004970 * (TOTAL_JOBS) + 0.186539 * (POP_TO_JOB_COMPARISON)*]
Webster	*AADT* = exp [*6.885442 + 0.001682 * (DISTANCE_TO_INTERSTATE) – 0.099051 * (DISTANCE_TO_MAJOR_HIGHWAYS) – 0.002901 * (TOTAL_POPULATION) + 0.004136 * (TOTAL_JOBS) – 0.658609* (POP_TO_JOB_COMPARISON)*]
Combination	*AADT* = exp [*5.975311 + 0.006731 * (DISTANCE_TO_INTERSTATE) – 0.053398 * (DISTANCE_TO_MAJOR_HIGHWAYS) + 0.033997 * (TOTAL_POPULATION) – 0.037866 * (TOTAL_JOBS) – 0.375535* (POP_TO_JOB_COMPARISON)*]

Table 3: Poisson Model Functions for Parishes with Interstate Access

MODEL TYPE	FUNCTION
Claiborne	*AADT* = exp [*5.153063 – 0.124589 * (DISTANCE_TO_MAJOR_HIGHWAYS) – 0.000398 * (TOTAL_POPULATION) + 0.013775 * (TOTAL_JOBS) + 0.027771 * (POP_TO_JOB_COMPARISON)*]
Franklin	*AADT* = exp [*6.262507 – 0.063446 * (DISTANCE_TO_MAJOR_HIGHWAYS) + 0.006997 * (TOTAL_POPULATION) – 0.014393 * (TOTAL_JOBS) – 0.390779 * (POP_TO_JOB_COMPARISON)*]
Vermilion	*AADT* = exp [*5.739465 – 0.017628 * (DISTANCE_TO_MAJOR_HIGHWAYS) + 0.002380 * (TOTAL_POPULATION) + 0.005512 * (TOTAL_JOBS) – 0.008011 * (POP_TO_JOB_COMPARISON)*]
Washington	*AADT* = exp [*5.724025 – 0.007453 * (DISTANCE_TO_MAJOR_HIGHWAYS) + 0.000835 * (TOTAL_POPULATION) + 0.004292 * (TOTAL_JOBS) – 0.153256 * (POP_TO_JOB_COMPARISON)*]
Combination	*AADT* = exp [*5.544503 – 0.025703 * (DISTANCE_TO_MAJOR_HIGHWAYS) – 0.000458 * (TOTAL_POPULATION) + 0.012488 * (TOTAL_JOBS) – 0.026565 * (POP_TO_JOB_COMPARISON)*]

Table 4: Poisson Model Functions for Parishes without Direct Interstate Access

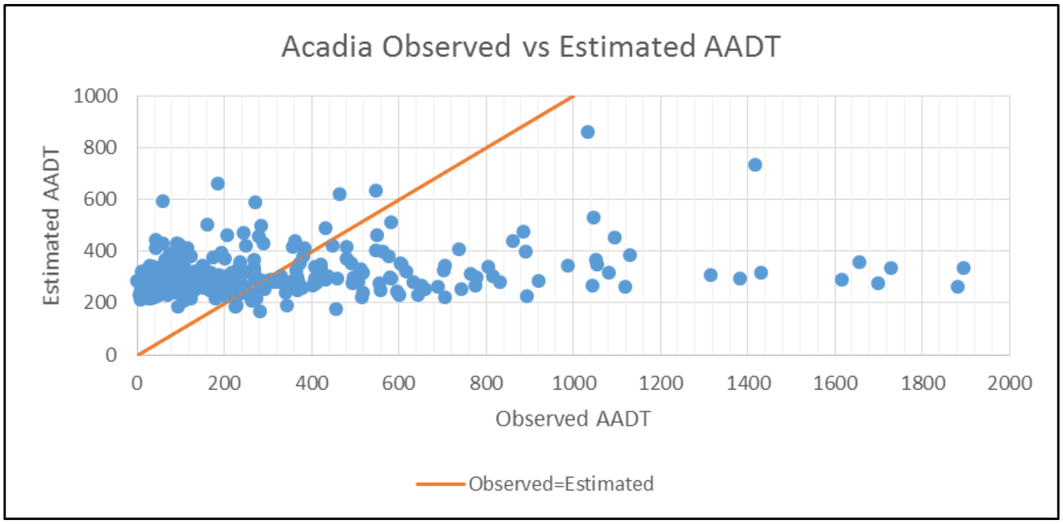


Figure 11: Acadia Poisson Observed versus Estimated AADT

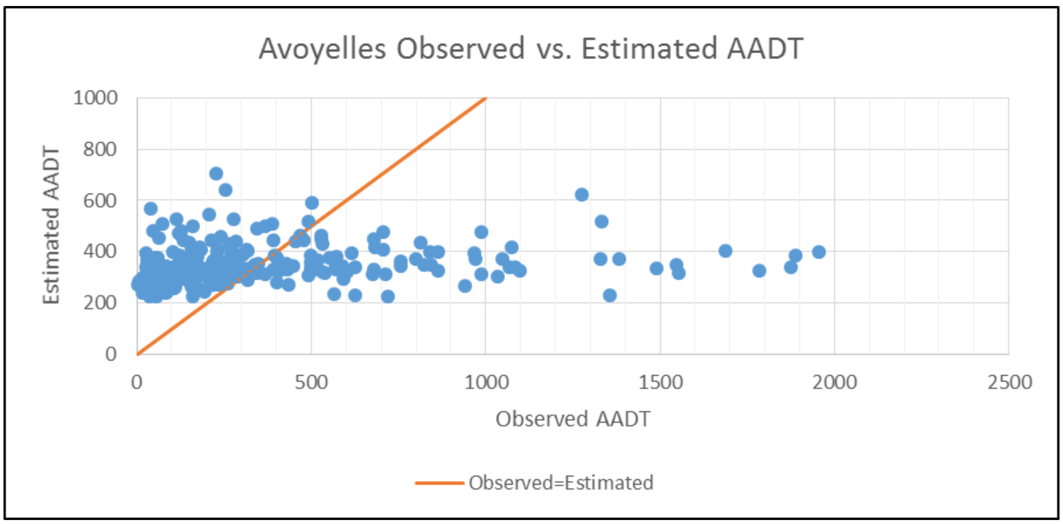


Figure 12: Avoyelles Poisson Observed versus Estimated AADT

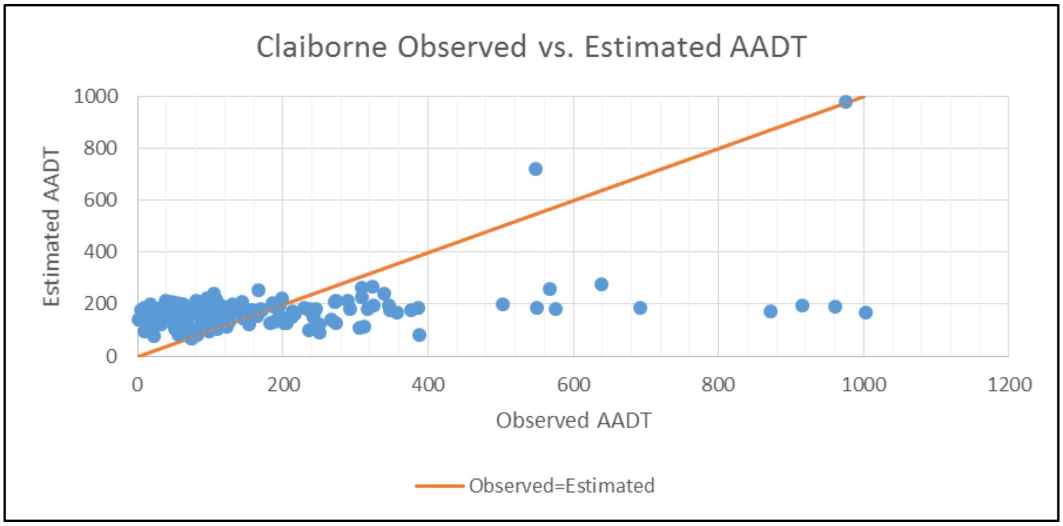


Figure 13: Claiborne Poisson Observed versus Estimated AADT

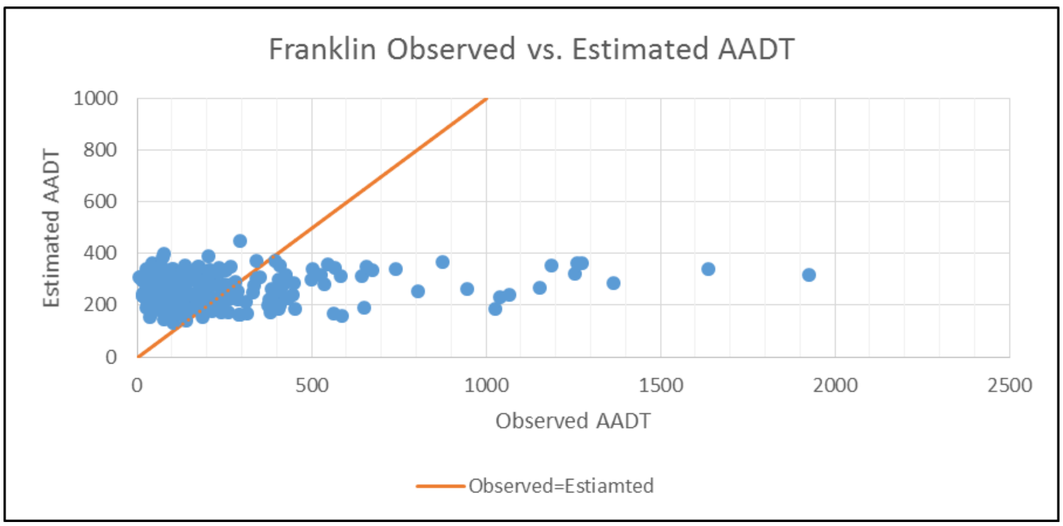


Figure 14: Franklin Poisson Observed versus Estimated AADT

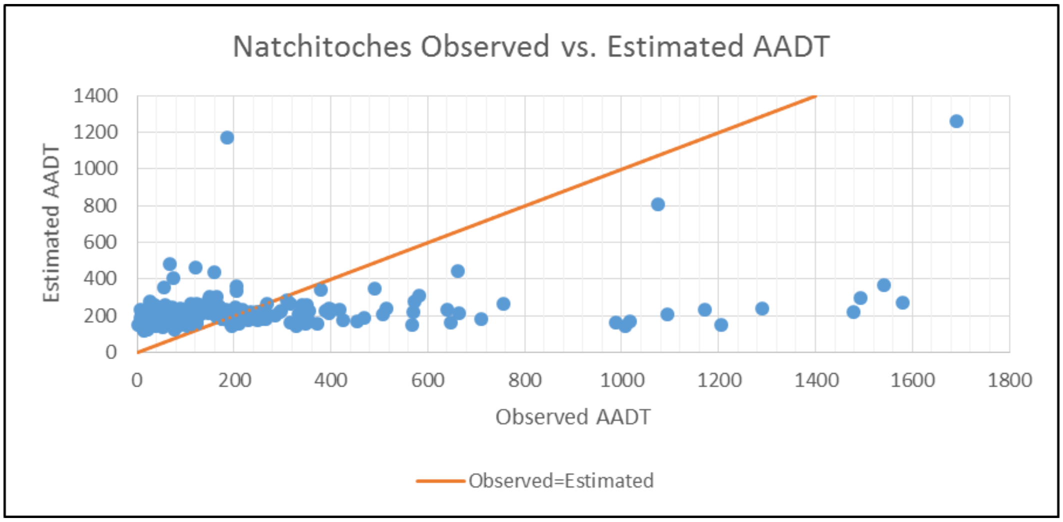


Figure 15: Natchitoches Poisson Observed versus Estimated AADT

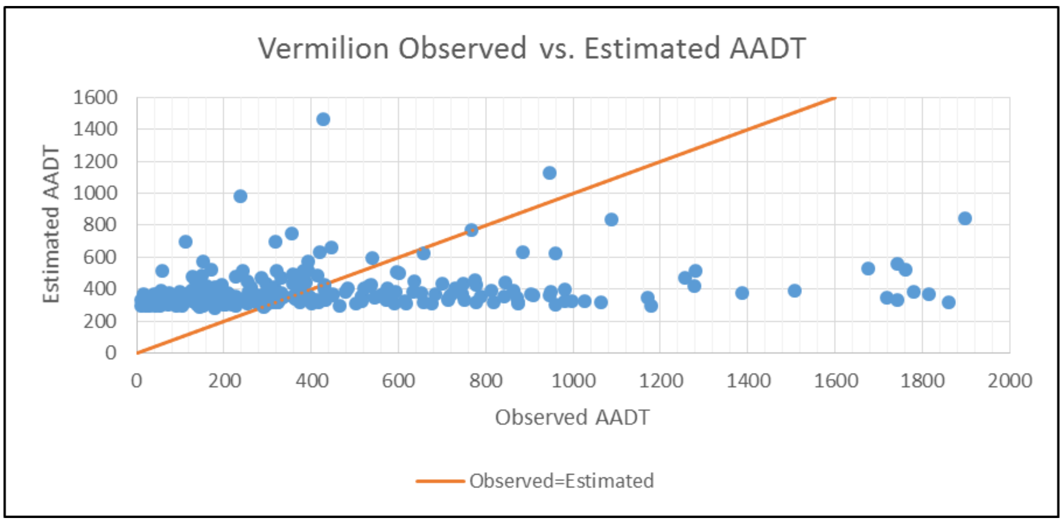


Figure 16: Vermilion Poisson Observed versus Estimated AADT

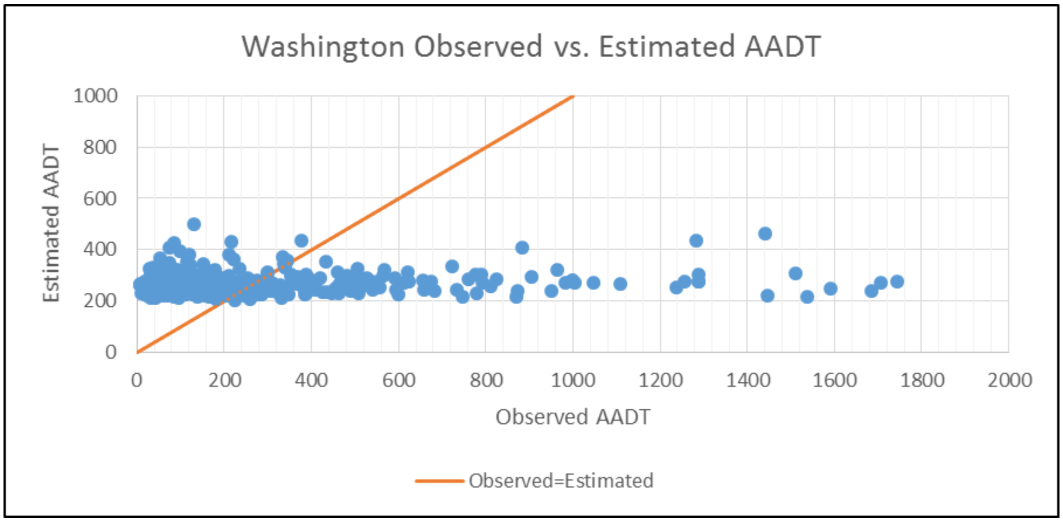


Figure 17: Washington Poisson Observed versus Estimated AADT

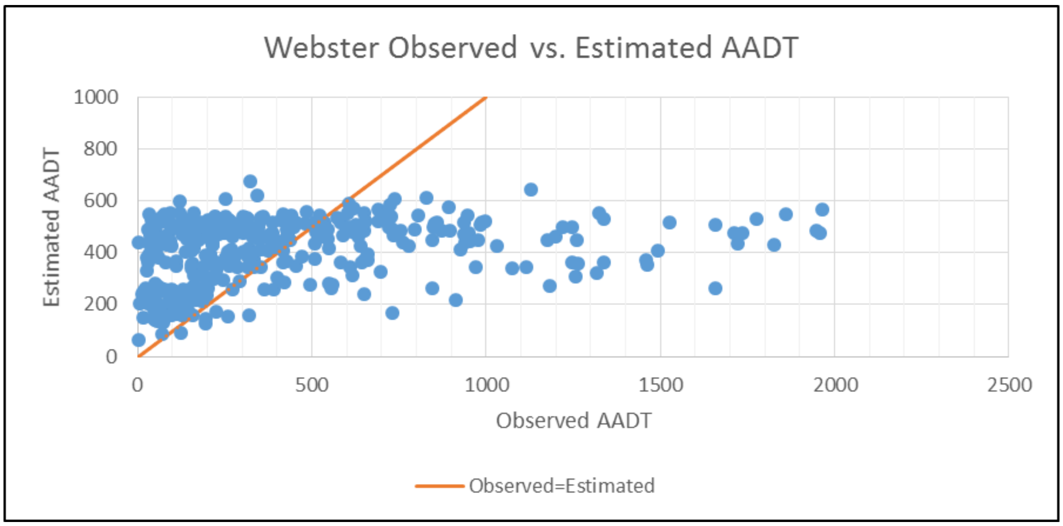


Figure 18: Webster Poisson Observed versus Estimated AADT

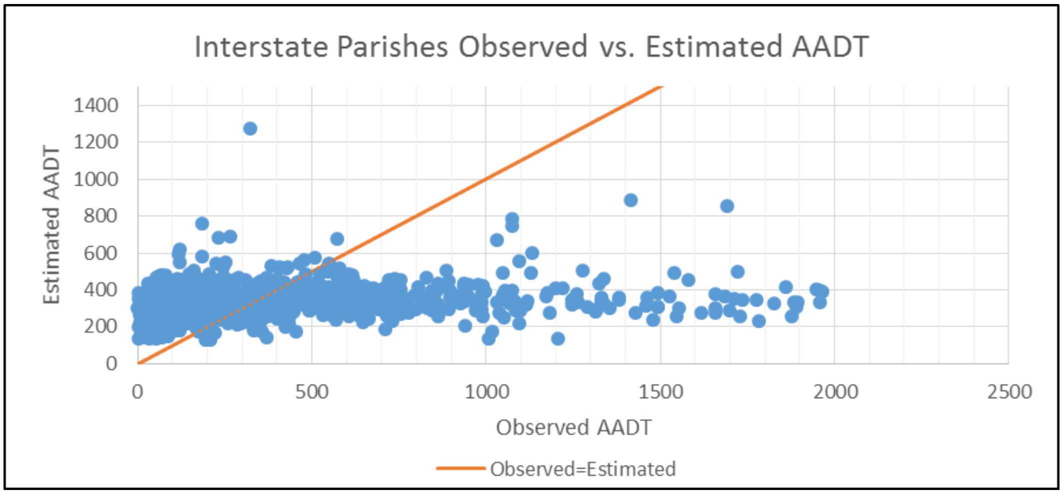


Figure 19: Interstate Parishes Poisson Observed versus Estimated AADT

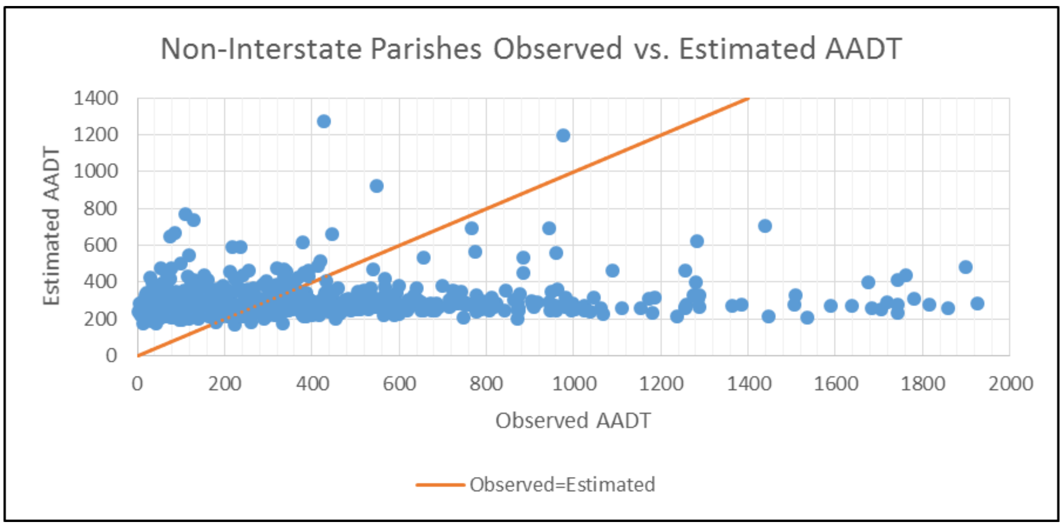


Figure 20: Non-Interstate Parishes Poisson Observed versus Estimated AADT

The 45-degree line *Observed=Estimated* was included in each of the above figures to compare the estimated AADT values to the observed AADT values, and the closer the observations were to this line, the better fit the model was. All ten initial models (eight parish-specific and two combination) overestimated the AADT for lower observed AADT values (typically below 400) and underestimated the AADT for the higher observed AADT values (usually above 400), as shown where the *Observed=Estimated* line crosses the datasets. In addition, while the range of observed AADT values was between 0 and 2,000 (with the exception of Claiborne Parish), the range of the estimated values was more narrow (usually between 300 and 600), with the exception of the Natchitoches and Vermilion Parish-specific models. The analysis of the estimated AADT range disregards the combination models since the maximum and minimum estimated AADT values for the combination models are similar to the overall maximum and minimum estimated AADT values from the parish-specific models. Because of the estimation errors detailed, no further analysis was performed for the Poisson model.

3.3.4 Support Vector Regression

Because the fit of the Poisson models was still poor even though among regression models, the Poisson model had the best fit, machine learning would have to be used, and support vector regression (SVR or support vector machine-SVM) was further studied due to its strong theoretical foundation, good generalization performance, the absence of local minima, and sparse representation of solutions. In addition, SVR can enhance prediction accuracy and provides an efficient way to compute SVR parameters. The quality and performance of the SVR models depend on the setting of three parameters: kernel type, value of the penalty for excess deviation during training (*C*), and error-term value for the ε-insensitive loss function (ε) (2). An open-source software programming language, R, is used to estimate AADT with these parameters:

- SVM-Type, which in this study is eps-regression
- SVM-Kernel, radial in this study
- Cost, a value of 100 in the study
- Gamma, a value of 1
- Epsilon, a value of 0.1

In addition, the number of support vectors is determined before running the SVR analysis. Once all parameters are determined, R is used to run the SVR analysis. Also, the values can be graphically summarized to better analyze the results since the initial estimated values are not shown in the script window in R. Appendix G shows how to use R for model development as well as sensitivity analysis.

3.3.5 Sensitivity Analysis

To determine if the final model selection will be parish-specific (meaning a total of eight models) or combination (two models), a sensitivity analysis was conducted. In this sensitivity analysis, the minimum and maximum values of each independent variable of the combined parishes (four Interstate and four non-Interstate) were taken, and each independent variable is divided into twenty equal increments between and including the minimum and maximum values with the exception of the population to job comparison variable as the value of that variable is binary, either *1* or *2*. Determining what type of model to use is based on the relationship among the parish-specific models; if the relationship among the individual models is similar, then a combination model can be used, but individual parish models must be used if no clear relationship among the individual models is shown to exist. The results of the sensitivity analysis is detailed in *4.1 Sensitivity Analysis.*

Chapter 4: Results and Discussion

4.1 Sensitivity Analysis

Using SVR, a sensitivity analysis was conducted for two datasets, one for the parishes with Interstate access, and one for parishes without Interstate access. The sensitivity analysis was used to determine if individual models for each parish or one model for all parishes with and without Interstate access would be used. The results are detailed in **Figure 21 and Figure 22.**

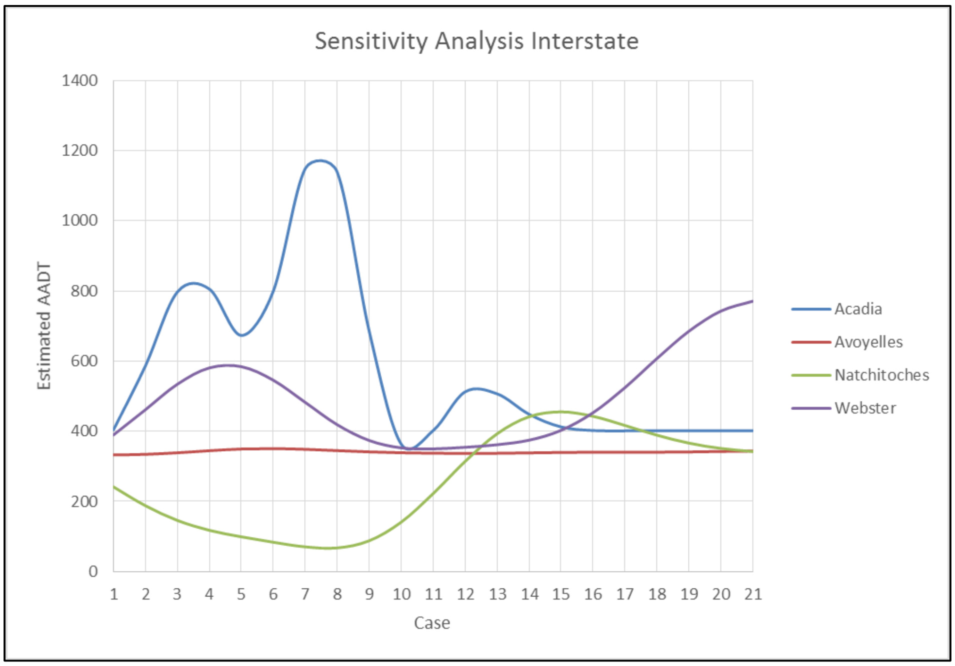


Figure 21: Sensitivity Analysis Interstate Parishes

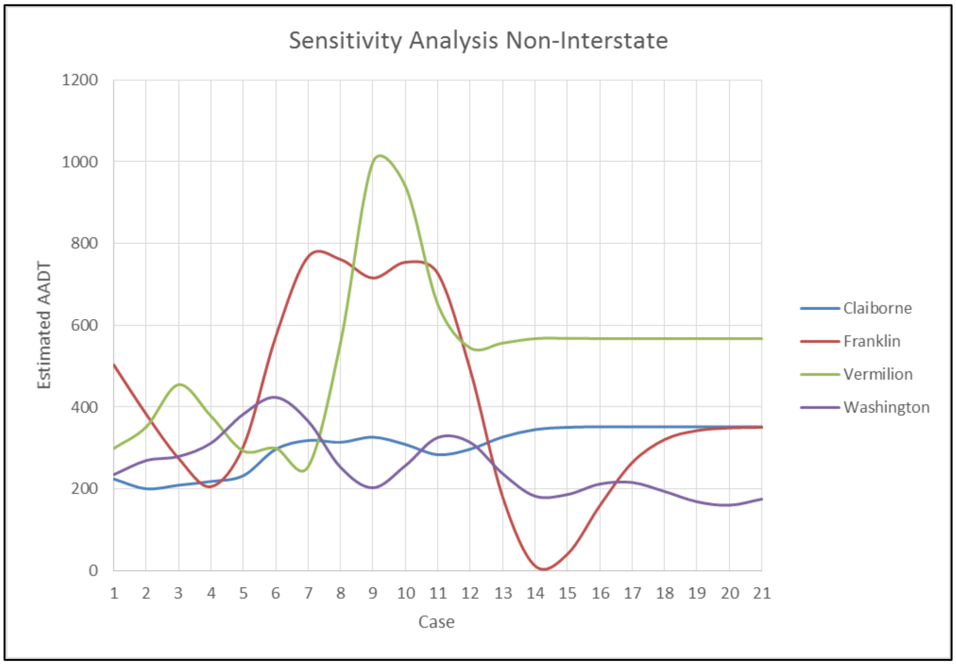


Figure 22: Sensitivity Analysis Non-Interstate Parishes

The results of the sensitivity analyses above determined that individual parish models would have to be used since no clear relationship exists among the individual models. A likely reason for this unclear relationship is that the maximum values for the entire group of parishes may not be applicable to certain parishes. One notable observation from the results of the sensitivity analysis is the "plateau" that occurs in some models as the case number increases; this likely results when the values of all the independent variables exceed the maximum values of a particular parish before reaching the overall maximum values for the combination datasets.

4.2 SVR Parish Models

After determining that individual parish models would need to be developed, SVR would be used to create the eight parish-specific models. **Figure 23 to Figure 30** show the results using SVR to estimate AADT, and two bands are shown to compare how the estimated

AADT is related to the observed AADT (estimated AADT within 100 or 200 of the observed AADT).

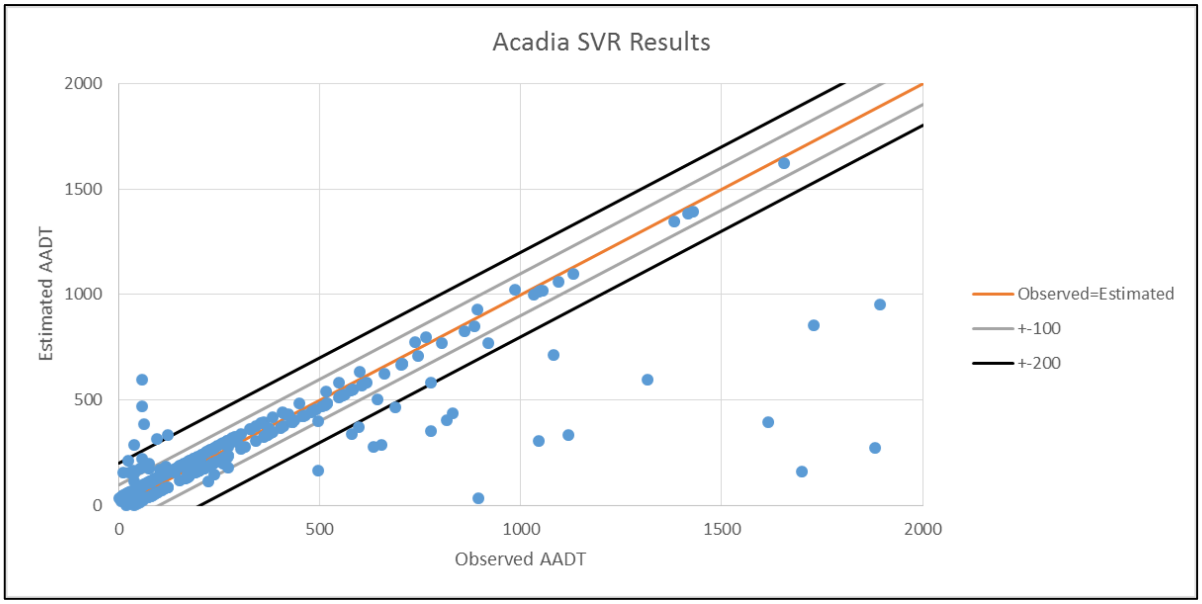


Figure 23: Acadia SVR Observed vs. Estimated

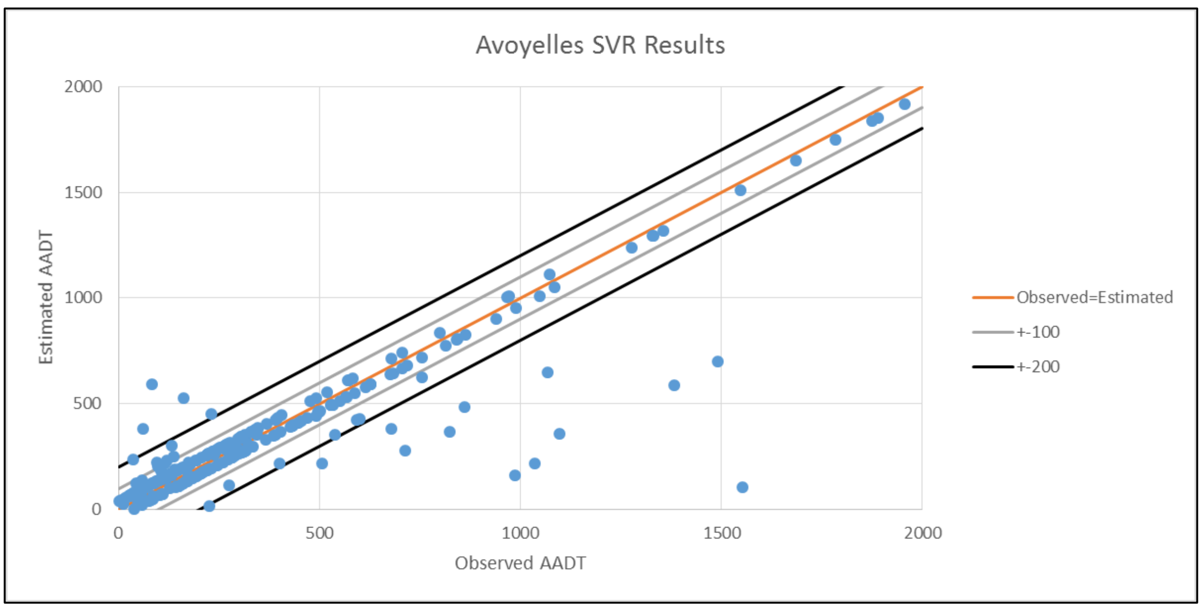


Figure 24: Avoyelles SVR Observed vs. Estimated

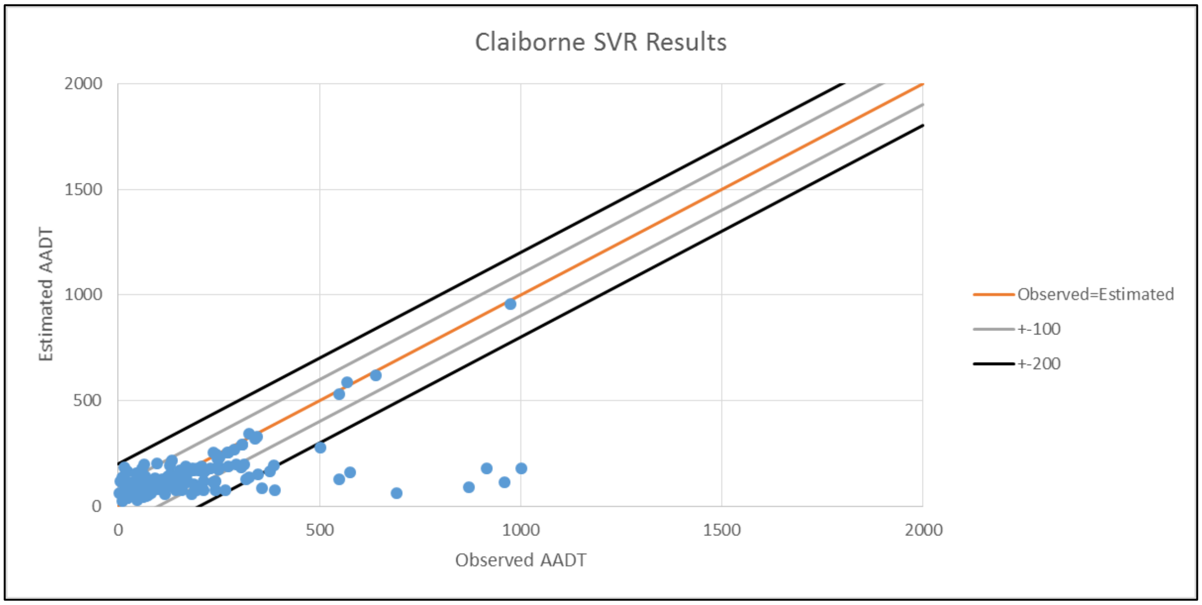


Figure 25: Claiborne SVR Observed vs. Estimated

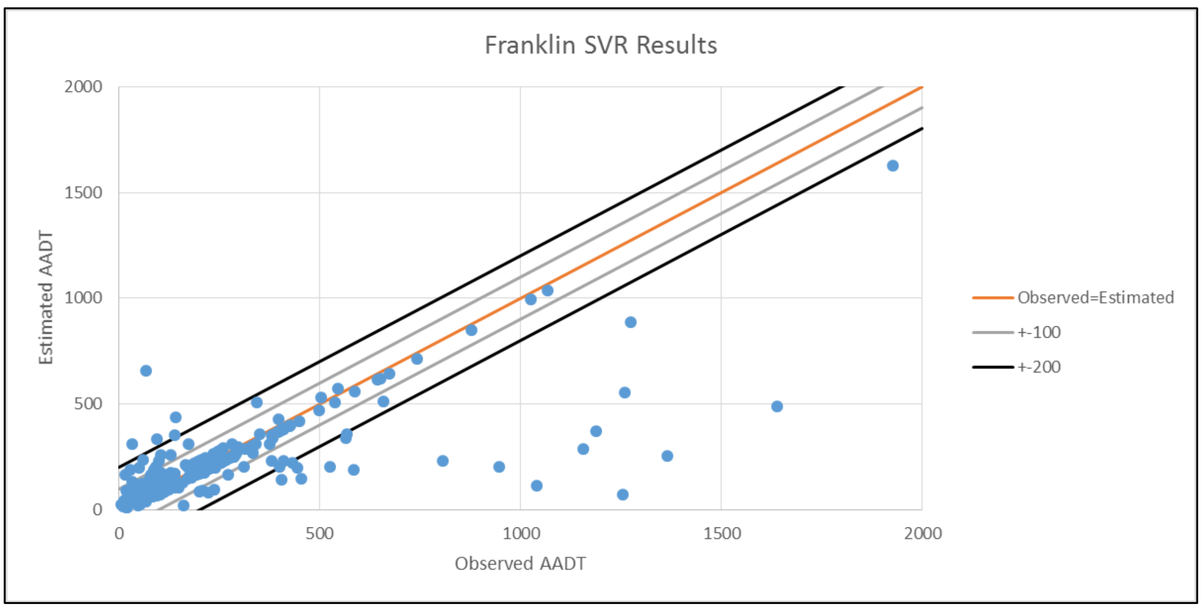


Figure 26: Franklin SVR Observed vs. Estimated

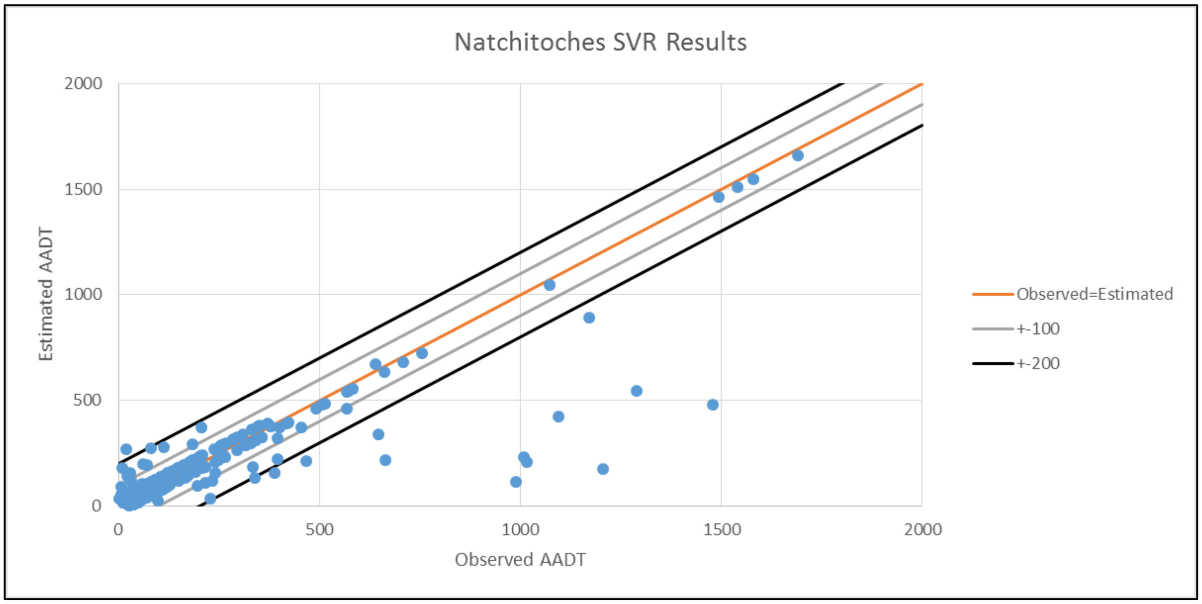


Figure 27: Natchitoches SVR Observed vs. Estimated

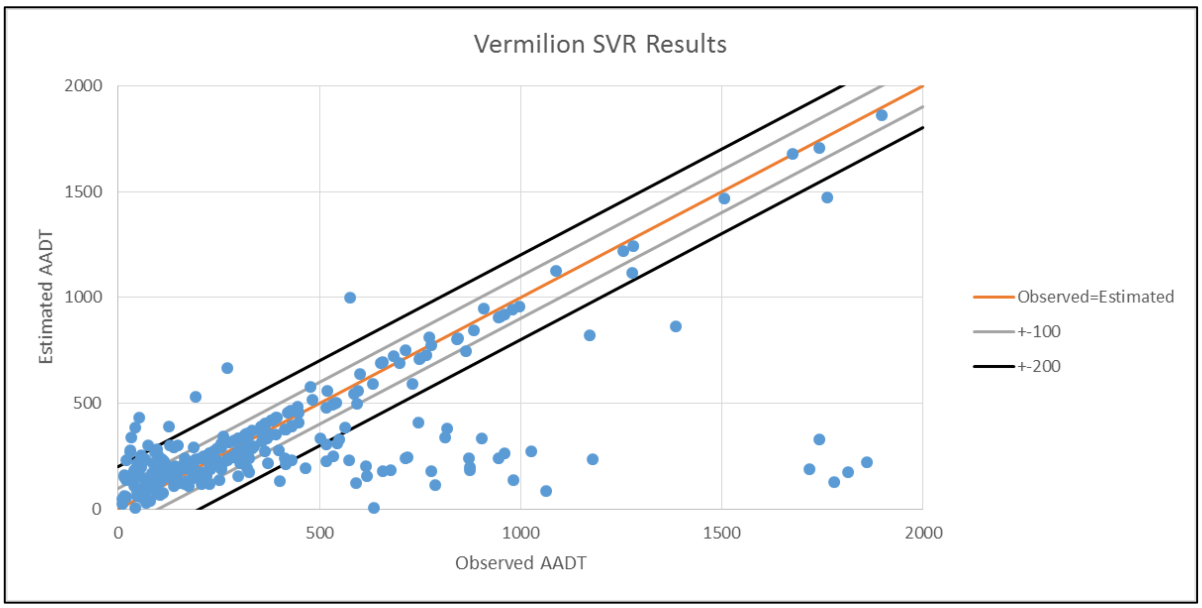


Figure 28: Vermilion SVR Observed vs. Estimated

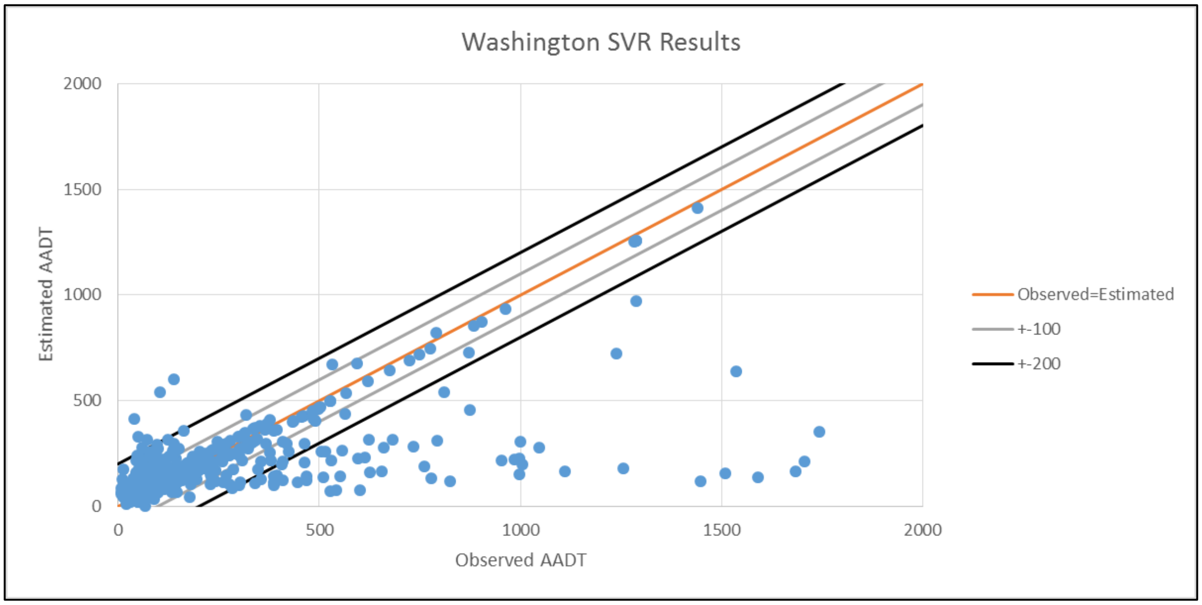


Figure 29: Washington SVR Observed vs. Estimated

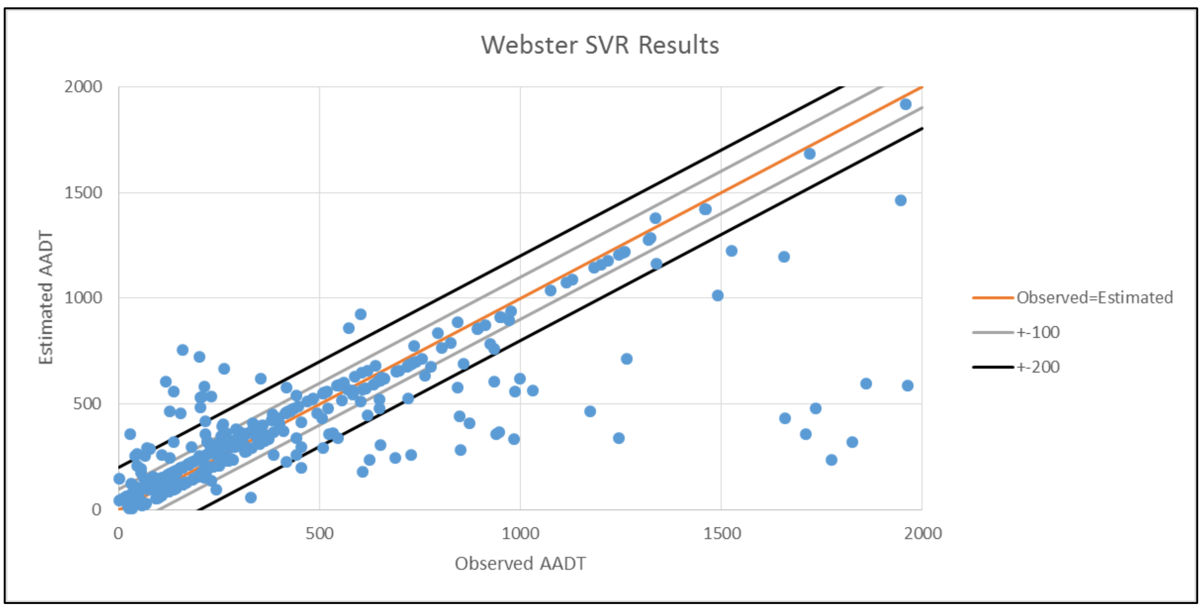


Figure 30: Webster SVR Observed vs. Estimated

Compared to the initial models, the SVR greatly improved the estimation of the AADT. Far more observations followed the same direction as the *Observed=Estimated* line. Nonetheless, for all eight models, SVR did underestimate the AADT at higher observed AADT values, but the underestimation of the AADT in SVR was not nearly as problematic as in the Poisson models. In other words, the SVR "captured" most of the overestimation of AADT, but some underestimation of AADT still existed. Also, the SVR models did overestimate some AADT values, typically occurring at lower observed AADT values. Nonetheless, the SVR did better estimate the AADT for some of the higher observed AADT values. The next section explains in more detail how close the estimated AADT was to the observed AADT using the two bands shown (within 100 or 200 of the observed).

4.3 Comparison of Poisson and SVR Models

Because SVR does not determine probability estimates (standard deviation), two bands were used (within 100 and 200 of the observed AADT) to determine how SVR estimated the AADT in comparison to the observed AADT. The percentage of observations within these two bands is shown in **Table 5.**

			±100		±200	
Model		Sample size	Count	Percent	Count	Percent
Interstate	Acadia	307	270	87%	284	92%
	Avoyelles	272	242	89%	255	94%
	Natchitoches	243	213	88%	229	94%
	Webster	393	306	78%	338	86%
Non-Interstate	Claiborne	179	140	78%	168	94%
	Franklin	233	187	80%	209	90%
	Vermilion	306	212	69%	254	83%
	Washington	422	304	72%	362	86%

Table 5: Fit within 100 or 200 of Observed Values for SVR

For six of the eight models, more than three quarters (75%) of all estimations were within 100 of the observed, and the percentage of estimations within 100 of the observed for the remaining two models were near or above 70%. The average percentage of estimations within 100 of the observed for parishes with direct Interstate access was 85.5% while the average for parishes without direct Interstate access was somewhat lower at 75%, owning to the fact that the two models where the percentage of estimations within 100 of the observed AADT was less than 75% were both for parishes without direct Interstate access. The number of estimations within 200 of the observed for all eight models exceeded 80%, and five of the models exceeded 90%. In addition to determining how many estimations were within 100 or 200 of the observed AADT, the Mean Absolute Percentage Error (MAPE) for the Poisson and SVR models was compared to show the improvement in fit between these models. The formula for MAPE is shown below:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{E_i - O_i}{O_i}\right|$$

Where:

E_i is the estimated value at observation i

O_i is the observed value at observation i

n is the total number observations of the model

Table 6 shows the comparison between the Poisson and SVR MAPE calculations as well as the percentage improvement using SVR.

Model		Poisson MAPE	SVR MAPE	Percent Improvement
Interstate	Acadia	379.2%	75.9%	80.0%
	Avoyelles	252.5%	48.9%	80.6%
	Natchitoches	256.1%	74.2%	71.0%
	Webster	230.3%	66.4%	71.2%
Non-Interstate	Claiborne	215.8%	113.9%	47.2%
	Franklin	201.4%	56.5%	71.9%
	Vermilion	197.9%	75.9%	61.6%
	Washington	207.3%	76.0%	63.3%

Table 6: MAPE Results

The MAPE for the Poisson models further iterates the poor fit shown when using Poisson to estimate AADT. The MAPE for all eight models significantly exceeded 100%, and all but one model exceeded 200%. In addition, while the MAPE for all eight models was large, the MAPE for the Acadia model was nearly 50% greater than the model with the second highest MAPE - Avoyelles. Using SVR improved the MAPE by as much as 80%, and an improvement of at least 60% was shown for all but one model (Claiborne). Although the MAPE improved using SVR versus Poisson modeling, the MAPE was still relatively high for all eight models. Only one model (Avoyelles) had an SVR MAPE of less than 50% while the MAPE for the Claiborne SVR model still exceeded 100%. The main reason for the relatively high MAPE is not caused by the aforementioned outliers (underestimating AADT at higher observed AADT values) but occurs when the estimated AADT is significantly greater than the observed AADT. The MAPE in some of these observations exceeds 1,000% while the MAPE for the observations where the estimated AADT was less than observed was significantly lower than when the estimated AADT was greater than observed. In addition, the absolute percentage error (percentage error for each individual observation) for most of the individual observations was less than the MAPE of the entire model due to the relatively few observations where the MAPE exceeded 1,000%. The much greater MAPE when observed values are small and estimated values are significantly larger than the observed is an inherent problem of using MAPE since the denominator of the MAPE equation is the observed value. Nonetheless, using MAPE to compare between the Poisson and SVR models is a good fit parameter since SVR does not calculate probability statistics including standard deviation and variance.

Chapter 5: Conclusion

To consider the variability in characteristics, including demographic, economic, and roadway attributes throughout Louisiana, eight parishes were selected for model development. An emphasis was placed on rural parishes having a high number of count stations in a particular parish, as these parishes do not have the resources to collect AADT counts on a regular basis. The type of model selected was Poisson as this type of model is commonly used for the discrete data of AADT. Creating selection models proved this as the Poisson model was the only one that was statistically significant through the use of a Chi-square test, and the coefficients of the model were also determined to be statistically significant using the same statistical test. Initially, ten models were developed for rural non-state roads in Louisiana, with eight of those being Parish-specific and two models combining the data from four Parishes each. One of those combination models included a variable considering the distance between a count station and an Interstate access location (on-ramp) for those Parishes having direct Interstate access within the study Parish, while the second combination model was for the Parishes that did not have Interstate access. Using the JMP program, Poisson models were developed, and a major and common error was shown for all developed models. The models tended to overestimate the AADT at lower observed AADT values and underestimate the AADT at higher observed values (i.e. *estimated* > *observed* when observed AADT was less than a particular value and *estimated* < *observed* when observed AADT was greater than a particular value). That particular observed AADT value for all models developed was similar, typically between 300 and 500. Another method would need to be used in an attempt to “capture” this error.

Even with Poisson being the best regression model, the results were still less than satisfactory. Therefore, another method of estimation would need to be used, and Support Vector Regression (SVR) was used in an attempt to create a better fit model. Before developing the models, a sensitivity analysis was conducted to determine which type of model would be used, each Parish having a model that was specific to that Parish or a combination model that could be used on all four of the similar Parishes (Interstate or Non-Interstate). This analysis determined that the prediction models would have to be specific to each Parish as using each model on the combination values for the independent variables resulted in vastly different estimated AADT values. Using SVR, the relationship between the observed and estimated AADT greatly improved. Two goodness-of-fit parameters were used to understand the relationship between the observed and estimated AADT: determining how many estimations were within a particular value of the observed AADT (100 and 200 were used in this project), and using Mean Absolute Percentage Error (MAPE). Using both of these goodness-of-fit parameters showed a major improvement in using SVR, though some drawbacks were shown in using MAPE that are inherent to that goodness-of-fit parameter (much higher MAPE for smaller observed AADT).

While using SVR significantly improves AADT estimation and the relationship between the observed and estimated AADT in contrast to the Poisson models, some drawbacks do exist when using SVR. The most noticeable drawback is that probability estimates, including standard deviation and variance, cannot be estimated using SVR. Another drawback is that SVR does not determine mathematical equations, likely the result of SVR being a machine learning algorithm. Also, as a result of the sensitivity analysis, SVR must be used for individual parish datasets, and the application of SVR must be conducted as a “test” batch.

Some considerations that can be looked further into include the results of the sensitivity analysis and models for the smaller urban areas in the study parishes. A new type of sensitivity analysis could be used to determine if a combination parish model could be used, but the use of several different sensitivity analyses may not change the results of the initial model. While the focus of this study was the rural areas in Louisiana, the smaller municipalities (or urban areas), like the rural parishes, likely do not collect traffic counts on their roads. Therefore, a separate use of SVR in estimating AADT in the smaller urban areas may provide a better result than combining all non-state counts within a particular parish. A notable consideration in estimating AADT in small urban areas is that likely only one model would be needed for the parish's entire municipalities, not individual small urban models like the individual parish models.

References

1. Selby, B. and Kockelman, K. (2011). *Spatial Prediction of AADT in Unmeasured Locations by Universal Kriging.* Research Report. The University of Texas, Austin, 2011.

2. Castero-Neto, M., Jeong, Y., Jeong, M. K., and Han, L. D. *AADT Prediction Using Support Vector Regression with Data-Dependent Parameters.* Research Report. University of Tennessee, Knoxville, 2008.

3. Zhao, F. and Chung, S. *Estimation of Annual Average Daily Traffic in a Florida County Using GIS and Regression*. Research Report. Florida International University, Miami, 2001.

4. Dabbaghian, V. *Chapter 1: What is Mathematical Modeling?*. Simon Fraser University, Burnaby, British Columbia, Canada. 2014.

5. Domingos, P. *A Few Useful Things to Know about Machine Learning.* University of Washington, Seattle. 2012.

6. System Description. Florida's Turnpike Enterprise, Ocoee. http://www.floridasturnpike.com/about_system.cfm. Accessed August 27, 2014.

7. Wang, T., Gan, A., and Alluri, P. *Estimating Annual Average Daily Traffic for Local Roads for Highway Safety Analysis*. Research Report. Florida International University, Miami, 2013.

8. Seaver, W. L., Chatterjee, A., and Seaver, M. L. *Estimation of Traffic Volume on Rural Local Roads*. Research Report. University of Tennessee, Knoxville and University of Georgia, Athens, 2000.

9. Gastaldi, M., Rossi, R., Gecchele, G., and Lucia, L.D. *Annual Average Daily Traffic Estimation from Seasonal Traffic Counts*. Research Report. Padova, Italy: University of Padova, Padova, Italy, 2012.

10. Dixon, Michael. *The Effects of Errors in Annual Average Daily Traffic Forecasting: Study of Highways in Rural Idaho*. Research Report. University of Idaho, Moscow, 2004.

11. Road Transfer Program. Louisiana Department of Transportation and Development, Baton Rouge. http://wwwsp.dotd.la.gov/Inside_LaDOTD/Divisions/Multimodal/Road_Transfer/Pages/default.aspx. Accessed August 19, 2014.

12. DOTD I-69, SIU 15 Project Site. Louisiana Department of Transportation and Development, Baton Rouge. http://www.i69dotd.com/. Accessed October 16, 2014.

13. Geaux South 49. Louisiana Department of Transportation and Development, Baton Rouge. http://www.geauxsouth49.com/. Accessed October 16, 2014.

14. I-49 North. Louisiana Department of Transportation and Development, Baton Rouge. http://wwwapps.dotd.la.gov/administration/public_info/projects/home.aspx?key=21. Accessed September 24, 2014.

15. I-49 Inner-City Connector- Shreveport. Northwest Louisiana Council of Governments, Shreveport. http://www.i49shreveport.com/Site/. Accessed September 24, 2014.

16. Geographic Terms and Concepts-Census Tract. United States Census Bureau, Suitland, Maryland. https://www.census.gov/geo/reference/gtc/gtc_ct.html. Accessed August 20, 2014.

17. Geographic Terms and Concepts-Block Group. United States Census Bureau, Suitland, Maryland. https://www.census.gov/geo/reference/gtc/gtc_bg.html. Accessed August 20, 2014.

18. Geographic Terms and Concepts-Block. United States Census Bureau, Suitland, Maryland. https://www.census.gov/geo/reference/gtc/gtc_block.html. Accessed August 20, 2014.

19. Longitudinal Employer-Household Dynamics. United States Census Bureau, Suitland, Maryland. http://lehd.ces.census.gov/data/. Accessed September 17, 2014.

Appendix A: Selection of Parishes in Louisiana Based on the Number of Count Stations and Population (Census 2010)

District	Parish Name	# Count Stations	Population (Census 2010)	District	Parish Name	# Count Stations	Population (Census 2010)
02	Jefferson	877	31,594	07	Beauregard	859	35,654
02	Lafourche	1017	96,318	07	Calcasieu	1257	192,768
02	Orleans	139	343,829	07	Cameron	364	6,839
02	Plaquemines	319	23,042	07	Jefferson Davis	328	432,552
02	St. Bernard	195	35,897	08	Avoyelles	923	42,073
02	St. Charles	364	52,780	08	Grant	724	22,309
02	Terrebonne	563	111,860	08	Natchitoches	895	39,566
03	Acadia	1227	61,773	08	Rapides	1241	131,613
03	Evangeline	1025	33,984	08	Sabine	755	24,233
03	Iberia	719	73,240	08	Vernon	862	52,334
03	Lafayette	1107	221,578	08	Winn	747	15,313
03	St. Landry	1104	83,384	58	Caldwell	512	10,132
03	St. Martin	735	52,160	58	Catahoula	467	10,407
03	St. Mary	671	54,650	58	Concordia	294	20,822
03	Vermillion	987	57,999	58	Franklin	810	20,767
04	Beinville	803	14,353	58	Lasalle	600	14,890
04	Bossier	699	116,979	58	Tensas	275	5,252
04	Caddo	1083	254,969	61	Ascension	791	107,215
04	Claiborne	681	17,195	61	Assumption	165	23,421
04	Desoto	708	26,656	61	East Baton Rouge	815	440,171
04	Red River	321	9,091	61	East Feliciana	409	20,267
04	Webster	970	41,207	61	Iberville	428	33,387
05	East Carroll	343	7,759	61	Point Coupee	304	22,802
05	Jackson	498	16,274	61	St. James	293	22,102
05	Lincoln	760	46,735	61	West Baton Rouge	317	23,788
05	Madison	287	12,093	61	West Feliciana	61	15,625
05	Morehouse	698	27,979	62	Livingston	939	128,026
05	Ouachita	894	153,720	62	St. Helena	433	11,203
05	Richland	726	20,725	62	St. John The Baptist	284	45,924
05	Union	1072	22,721	62	St. Tammany	1433	233,740
05	West Carroll	509	11,604	62	Tangipahoa	1199	121,097
07	Allen	690	25,764	62	Washington	1177	47,168

Appendix B: Attributes of the Local AADT Dataset from the Louisiana Department of Transportation and Development

The attributes for each non-state (local) count station (location and observed AADT in a particular year) in Louisiana is detailed below:

- **STATION:** The five or six digit ID for the count station, with the first digit or two digits being the parish code
- **DISTRICT:** The LA DOTD District the count station is located in (Appendix D gives further explanation on the Districts)
- **PARISH CODE:** The parish the count station is located in. The code starts with the value of *1*, being Acadia Parish, the first parish in alphabetical order and increases by 1 for each successive parish in alphabetical order
- **STREET NAME:** The name of the street the count station is located on
- **LRS ID:** The state-issued ID for a particular roadway segment, which is in the format PPP-X-NNNNNN-TTT-S-F-LL and is described below:
    - PPP- Parish FIPS
    - X- Prefix Code (N,S,E, or W)
    - T- Type Code (Ave., Blvd., St., etc.)
    - S- Suffix (N, S, E, or W)
    - F- Feature Type Code (Main direction, Frontage Road, or Ramp)
    - L- Sequential Occurrence.
- **LRS LOGMILE:** Logmile on the roadway segment where the count station is located

- **YEAR 1, YEAR 2, … YEAR 6:** The year when the AADT was recorded; not all stations have data for six different years. Year 1 is the most recent year the AADT was recorded, and no stations have more than six different years of recorded data
- **ADT 1, ADT 2, …, ADT 6:** The recorded AADT for a particular year (For example, if year 1 was in 2007, then the ADT 1 that is given is what was recorded in 2007)
- **LATITUDE:** The coordinate detailing the Y-Axis component of the location of the count station
- **LONGITUDE:** The coordinate detailing the X-Axis component of the location of the count station.

Since some count stations have more than one year of count data available for local roads, only the data recorded for the most recent year (YEAR 1) is to be used as the dependent variable in model determination. The Latitude and Longitude of the count station is to be used in ESRI's ArcGIS program to locate where the count station is on the State roadway network.

Appendix C: Statewide Roadway Network Attributes

The attributes of the Louisiana Statewide roadway network, which includes a roadway segment's descriptive characteristics (name of road, length of roadway segment, etc.) is detailed below:

- **NAME**, the street name (e.g. Main)
- **STREET_CATEGORY**, the type of street (Ave., St., Blvd., Rd., etc.)
- **SUFFIX**, if the roadway has a directional identifier (e.g. North)
- **FULL_NAME**, the full name of the roadway (e.g. North Main Street)
- **DOTD_DISTRICT**, the DOTD District the road segment is located in
- **PARISH_FIPS**, the Census Parish code
- **CONTROL_SECTION**, the State-identified code for a roadway
- **LRS_ID**, based on the Control Section and additional information to distinguish between roadway segments (the state LRS_ID is in the XXX-XX-F-LLL format where XXX-XX is the control section, F is the feature type code, and LLL is the sequential occurrence)
- **BEGIN AND END LOGMILE**, the beginning and ending logmile from the Control Section of the roadway segment
- **SHAPE_LENGTH (MILES)**, the length of the roadway segment
- **STATE_ROUTE**, if the roadway segment is on a state-maintained roadway
- **ROADWAY_CATEGORY**, the type of roadway (main road, frontage road, etc.)
- **OWNERSHIP**, the owner of the road (State, Parish, or Municipal).

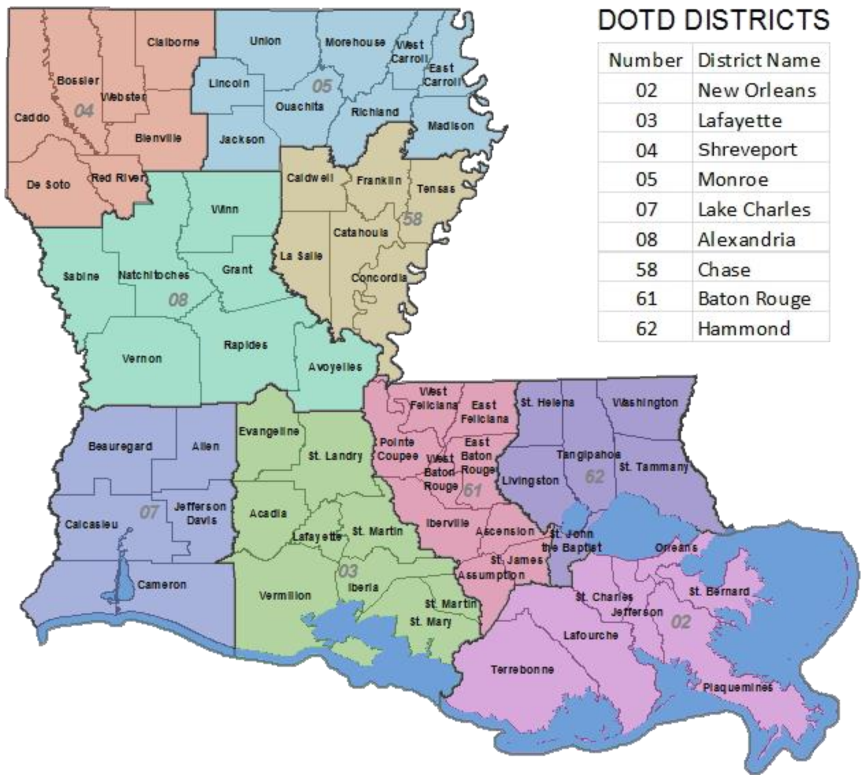
Appendix D: Louisiana Department of Transportation and Development Districts

The Louisiana Department of Transportation and Development (DOTD) operates nine districts throughout the state[3] that are responsible for operations and highway maintenance in a particular region of the state, which include:

- District 02: Southeastern Louisiana south of Lake Pontchartrain (headquarters in Bridge City just west of New Orleans)
- District 03: Acadiana (headquarters in Lafayette)
- District 04: Northwestern Louisiana (headquarters in Bossier City immediately east of Shreveport)
- District 05: Northeastern Louisiana (headquarters in Monroe)
- District 07: Southwestern Louisiana (headquarters in Lake Charles)
- District 08: Central Louisiana (headquarters in Alexandria)
- District 58: East-Central Louisiana (headquarters in Chase)
- District 61: South-Central Louisiana and Capitol Area (headquarters in Baton Rouge)
- District 62: Northshore of Lake Pontchartrain (headquarters in Hammond)

The map on the next page shows the location of each district within Louisiana, including the District number and the location of the headquarters of the District.

[3] http://wwwapps.dotd.la.gov/administration/announcements/DistrictMap.aspx



DOTD DISTRICTS

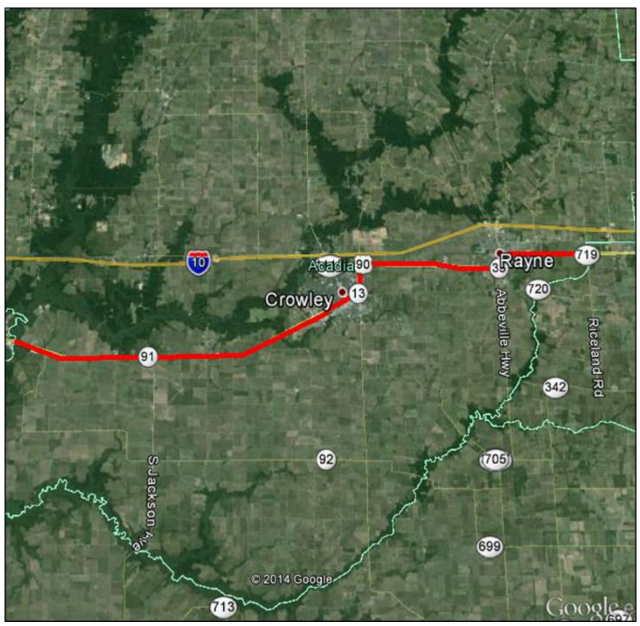
Number	District Name
02	New Orleans
03	Lafayette
04	Shreveport
05	Monroe
07	Lake Charles
08	Alexandria
58	Chase
61	Baton Rouge
62	Hammond

Louisiana DOTD Districts
Image Courtesy of Louisiana DOTD

Appendix E: Major Highways in Louisiana

United States Highways

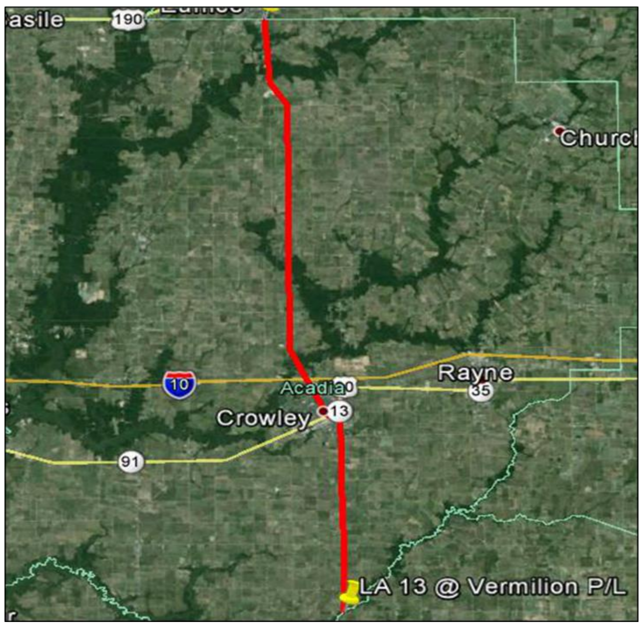
United States Highways (i.e. "US Highways") are generally state-maintained highways that are interstate in nature (serving more than one state) but do not typically meet Interstate Highway standards. In many locations where Interstates are nearby (e.g. Interstate 10 paralleling United States Hwy. 90 from the Texas State Line to Lafayette and New Orleans to the Mississippi State Line), these highways serve predominately local traffic; however, these highways can still be major thoroughfares in other locations where Interstates that are not nearby (e.g. United States Highway 90 between Lafayette and New Orleans).



Example: U.S. Highway 90 in Acadia Parish

Trans Parish Direct Highways

These highways are the direct route between two Parish Lines. The figure below shows Louisiana Highway 13 in Acadia Parish serving as the direct route between Vermilion Parish to the south (particularly between Gueydan, Kaplan, and Abbeville) and St. Landry Parish (Eunice) to the north. Especially in coastal areas, the North-South Trans-Parish highways serve as Hurricane Evacuation Routes from the coastal communities in the south towards North Louisiana.



Example: Louisiana Highway 13 in Acadia Parish from Vermilion Parish to St. Landry Parish

Main Highway to Parish Seat in Neighboring Parish

Some major highways have a terminus within a particular Parish, but serve as the main route to a major population center such as a Parish Seat in the neighboring Parish. The example figure below shows that the southern terminus of Louisiana Highway 9 is in Natchitoches Parish, and this highway serves as the direct route to the seat in Bienville Parish-Arcadia; in this particular example, this highway is more direct, especially for Natchitoches Parish, to reach Arcadia (and Interstate 20 Eastbound) versus using United States Hwy. 71 or Interstate 49 towards Shreveport to reach Interstate 20 Eastbound.

Example: Louisiana Highway 9 in Natchitoches Parish Connecting to Arcadia in Neighboring Bienville Parish

Appendix F: Census Geographic Attributes

The Census geographic dataset includes these attributes, regardless of whether the geographic subdivision is a Tract, Block Group, or Block:

- **FID**, which is the numerical order of the Census geographic subdivision
- **STATEFP10**, the Census's code for each state (The **STATEFP10** for Louisiana is 22.)
- **COUNTYFP10**, the Census's code for each parish (The code follows a *2n-1* formula where *n* is the Parish's number in alphabetical order; for example, Avoyelles Parish's **COUNTYFP10** is *7* since this Parish is the fourth Parish in alphabetical order in Louisiana; this is related to the Parish_FIPS code in Appendix D)
- **GEOID10,** the numerical designation of the Census geography, given in the format USSSCCCTTTTTTBBBB
    - SS-STATEFP10
    - CCC-COUNTYFP10
    - TTTTTT-Census Tract
    - BBBB-Block Number
- **NAME10**, the number of the Census geography which is in the format- *Block GBBB, Block Group G, Census Tract TTTT, (Parish Name), Louisiana*
- **NAMELSAD10**, the full name of the Census geography (e.g. *Census Tract 2*)
- **ALAND10**, the land area of the Census geography, in square meters
- **AWATER10**, the water area of the Census geography, in square meters
- **INTPLAT10**, the Latitude of the centroid of the Census geography
- **INTPTLON10**, the Longitude of the centroid of the Census geography

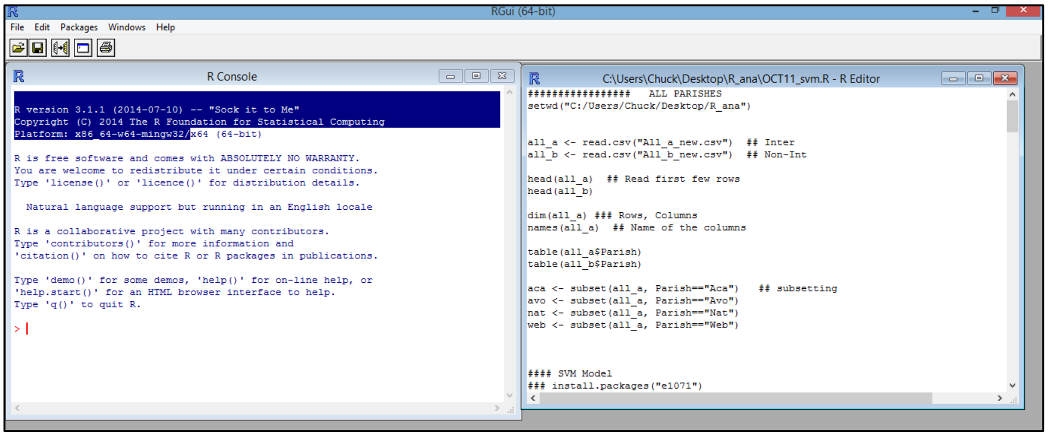
Since the area of a particular Census geographic subdivision (both land and water) is in square meters, a conversion to square miles is necessary to calculate the geographic subdivision's population density, which is shown in this formula:

$$AREA_{Square\ Miles} = \frac{\sum AREA_{\ Square\ Meters}}{(\ 1000\ \times 1.609)^2}$$

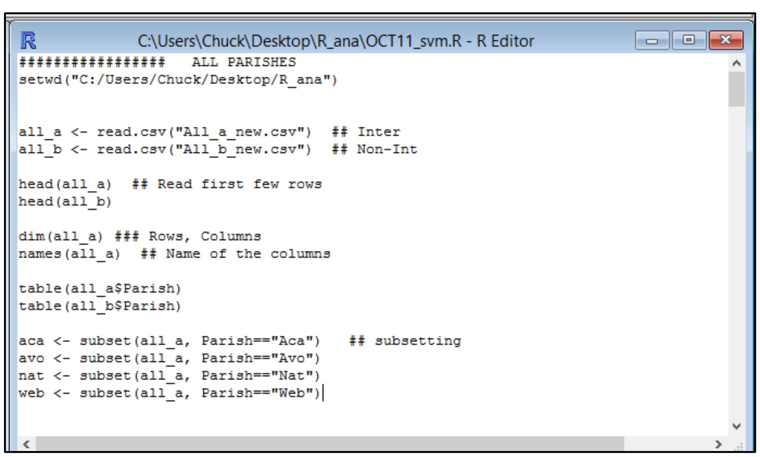
Dividing by $1{,}000^2$ (or 1,000,000) converts the area from square meters to square kilometers while dividing by 1.609^2 (or 2.588) converts the area from square kilometers to square miles.

Appendix G: Using R in SVR and Sensitivity Analysis

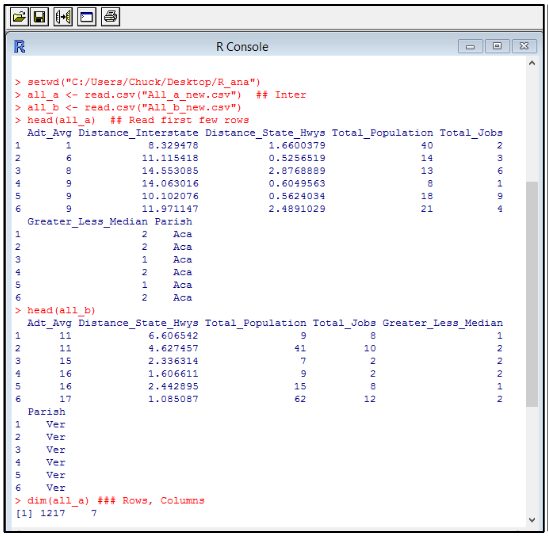
Before starting an analysis, the data source (spreadsheet) has to be in the EXCEL file format *Comma Separated Values (CSV)*, and a script needs to be created. Below is the initial setup of R with the console to the left and script to the right.

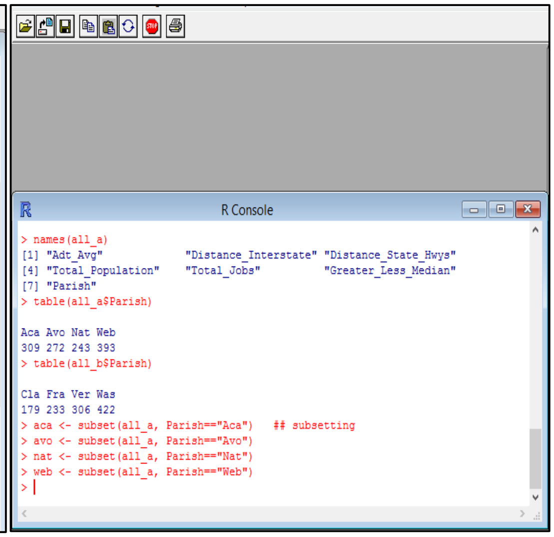


Because this study focuses on whether the parish has direct Interstate access, two CSV files are to be created for both model development and sensitivity analysis. A folder containing all necessary files (script and CSV spreadsheets) is highly recommended, and the resulting CSV spreadsheets can be saved in this particular folder. Next, to run the lines in the script, the particular lines must be highlighted, followed by clicking the "Run Line or Section"; a section is a group of lines between two spaces in the script. The initial lines and/or sections to be run in the script are shown in the script window below.

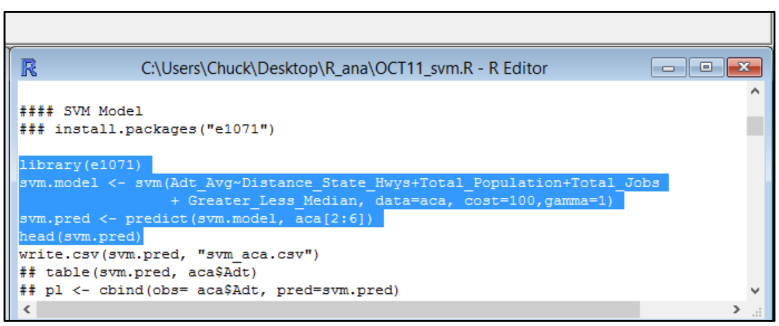


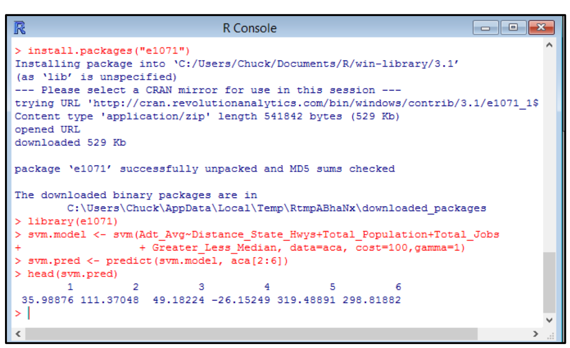
The results from running the initial script are shown in the following two figures in the R console.



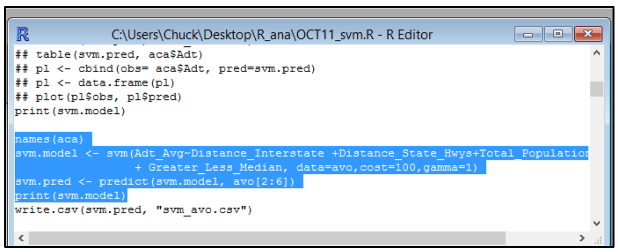


Afterwards, the necessary library for SVM Model must be installed ("e1071") by running the line *install.packages ("e1071")*. The CRAN mirror to be installed is *USA (TX 1)*. Once the library is loaded into the console, then SVM can be used to estimate AADT. The initial lines to be run and the results are shown below.

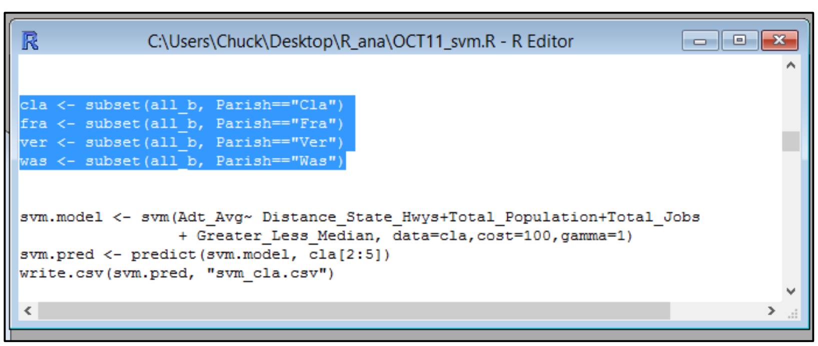




Because R only shows the results from the first few observations, a graphical representation is needed to compare observed and estimated AADT. To create a spreadsheet with the estimated AADT values, the line immediately below the initial highlighted code must be run. The resulting spreadsheet (in CSV format) is saved in the path folder determined when starting the analysis. Because the results are for only one parish selected, the next lines to be run to estimate AADT for the next parish are shown below.

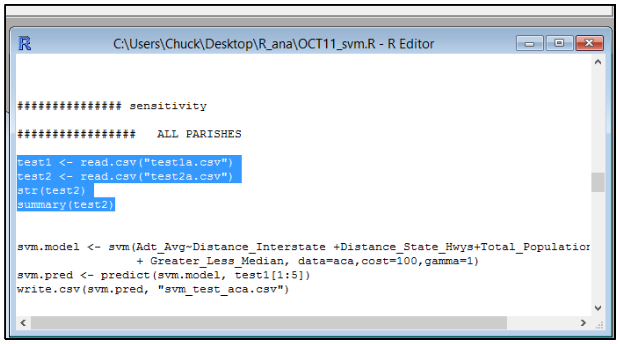


To create the spreadsheet with the estimated AADT values for this parish, the line immediately below the highlighted lines must be run, like in the first parish analysis. This process repeats for the remaining parishes in a particular group (Interstate or Non-Interstate access). The only change in using R for estimating AADT in the second group is shown below.



Once these lines are ran using R, then the succeeding lines in the script can be run to create spreadsheets with the estimated AADT for each parish in the second group.

For the sensitivity analysis, these script lines must be run in the initial run.



Afterwards, the first three lines must be ran to start the sensitivity analysis, and to create the results for each parish, the last line of the section code must be ran. Estimated AADT for each of the eight study parishes is determined from the eight code line sections.

LeBoeuf, Charles W. Bachelor of Science, University of Louisiana at Lafayette, Fall 2012;
Master of Science, University of Louisiana at Lafayette, Fall 2014
Major: Engineering, Civil Engineering option
Title of Thesis: Estimating Annual Average Daily Traffic for Non-State Roads in Louisiana
Thesis Director: Dr. Xiaoduan Sun
Pages in Thesis: 83; Words in Abstract: 241

ABSTRACT

Average annual daily traffic (AADT) is important in transportation engineering and planning, and although the State of Louisiana collects AADT on a regular basis on state-maintained highways, most parishes and smaller municipalities do not have the resources to collect AADT frequently. Because the roads under the jurisdiction of parishes and municipalities account for three-fourths of the entire state road network, a practical method to estimated AADT must be developed. Before model development, previous studies into AADT estimation and their results are to be further analyzed. Roadway, demographic, and economic data for selected parishes in Louisiana is collected and processed to remove any data not necessary in model development, and afterwards, parish-specific and combination data models using this data are developed to compare to the observed AADT at a particular count station. Parish selection is based on population, number of existing count stations within the parish, and if an Interstate Highway traverses the parish. Because of the varying characteristics among the data in the selected parishes, parish-specific models for the rural parish roads are developed, and Poisson is selected as the regression model due to discrete data. Results for all Poisson models developed show that the models tend to overestimate AADT for lower observed AADT and underestimate AADT for higher observed AADT. Because of this, support vector regression (SVR) was used, and this method greatly improved

the estimation of AADT in comparison to the Poisson regression as shown using certain goodness-of-fit parameters.

Biographical Sketch

Charles W. LeBoeuf was born to Charles J. and Tara LeBoeuf on June 5, 1990 in Baytown, Texas. He moved to Houma, Louisiana in October 1997 and to Lafayette, Louisiana in May 2009. LeBoeuf received his Bachelor of Science in Civil Engineering from the University of Louisiana at Lafayette in the Fall of 2012, and he was a member of the American Society of Civil Engineers, Chi Epsilon, and the Institute of Transportation Engineers. LeBoeuf plans to live and seek full-time employment in Lafayette after graduation.