TASK 2 RESEARCH

FINAL TECHNICAL MEMORANDUM

Prepared for
Federal Highway Administration
Office of Safety
Under Contract DFTH6116D00004
Task Order T0004

Submitted by
Ioannis Tsapakis
Subasish Das
William Holik
Shawn Turner
Edgar Kraus
Texas A&M Transportation Institute

Anita Vandervalk
Herb Weinblatt
Cambridge Systematics

Submitted on
December 20, 2016

TABLE OF CONTENTS

List of Figures

ACRONYMS

AADT	Annual average daily traffic
ADT	Average daily traffic
AASHTO	American Association of State Highway and Transportation Officials
CART	Classification and regression tree
CMF	Crash modification factor
CCS	Continuous count site
DOT	Department of Transportation
EB	Empirical Bayes
FAST-Act	Fixing America's Surface Transportation Act
FDE	Fundamental Data Elements
FHWA	Federal Highway Administration
GIS	Geographic information system
HPMS	Highway Performance Monitoring System
HPSJB	Highway Policy Steven Jessberger Battelle
HSIP	Highway Safety Improvement Program
HSM	Highway Safety Manual
ITE	Institute of Transportation Engineers
MADT	Monthly average daily traffic
MAP-21	Moving Ahead for Progress in the 21st Century
MAPE	Mean absolute percent error
MIRE	Model Inventory of Roadway Elements
MPO	Metropolitan Planning Organization
NCHRP	National Cooperative Highway Research Program
NFAS	Non-Federal-aid system
OD	Origin destination
OLS	Ordinary Least Squares
OLR	Ordinary linear regression
PTR	Portable traffic recorder
PVA	Property valuation administrator
RMSE	Root mean square error
SPF	Safety performance function
SVR	Support vector regression
TAZ	Transportation analysis zone
TDM	Travel demand model
TMAS	Travel Monitoring Analysis System
TMG	Traffic Monitoring Guide
USDOT	United States Department of Transportation
VMT	Vehicle miles traveled

SUMMARY

One of the most widely used traffic parameters is the annual average daily traffic (AADT). Transportation agencies use AADT to make informed decisions, meet data reporting requirements, and support safety analysis and other functions related to planning, pavement design and maintenance, highway design, operations, and environmental analysis. With the development of new safety analysis tools such as AASHTOWare's Safety Analyst™ and the *Highway Safety Manual* (HSM), AADT has become one of the most important data inputs in safety analysis.

The Federal Highway Administration (FHWA) requires states to report AADT through the Highway Performance Monitoring System (HPMS) for the full extent of mainlines, samples and ramps on all Federal-aid roads. In March 2016, the United States Department of Transportation (USDOT) published the Highway Safety Improvement Program (HSIP) Final Rule. According to the new requirements, by 2016 states will be required to collect AADT data along with other *Model Inventory of Roadway Elements (MIRE) – Fundamental Data Elements (FDEs)* for all public paved roads, including non-Federal aid system (NFAS) roadways.

Most agencies tend to focus their data collection efforts on high-volume roads that typically pose significant safety challenges compared to NFAS roads. The high expense and time consuming nature of data collection tasks has raised concerns regarding states' ability to conduct extensive short-duration counts on NFAS roads. FHWA's HPMS does not require any explicit procedure for sampling traffic volumes on NFAS roads. Selecting a sampling technique and an AADT estimation method is left at the discretion of respective Departments of Transportation (DOTs), which often use historical traffic volume data or modeling methods for these roads. When traffic count data are inaccessible, estimates are based on comparisons to similar types of roadways. This approach may not be repeated often enough to remain current and can result in major errors due to inaccurate assumptions.

The objective of this Task Order is to develop an informational guide on collecting traffic volume data on NFAS roads and computing and estimating AADT values for use in data-driven safety analysis. This study will highlight current research into the estimation of AADT values for lower volume roads and bridge the gap between safety and traffic data communities at State, tribal, Metropolitan Planning Organizations (MPOs), local and Federal agencies.

The goal of Task 2 (Research) is to provide recommendations for the proposed scope of the guide. To achieve this goal, the project team reviewed available documentation on this topic. Overall, the review reveals different types of AADT development methods that have significant differences in structure, data inputs, and accuracy of results. For clarity, the project team divided and presented these methods into three major groups (Figure 1):

A. Traffic count based methods. The traffic count based methods use traffic volume data obtained from continuous count sites (CCSs) or portable traffic recorders (PTRs) or both.

B. Non-traffic count based methods. These methods use non-traffic data such as socioeconomic, census, land use, and temporal variables.

C. Travel Demand Models (TDMs). These models are based on network, land use, and socioeconomic data to calculate the expected demand for transportation facilities.

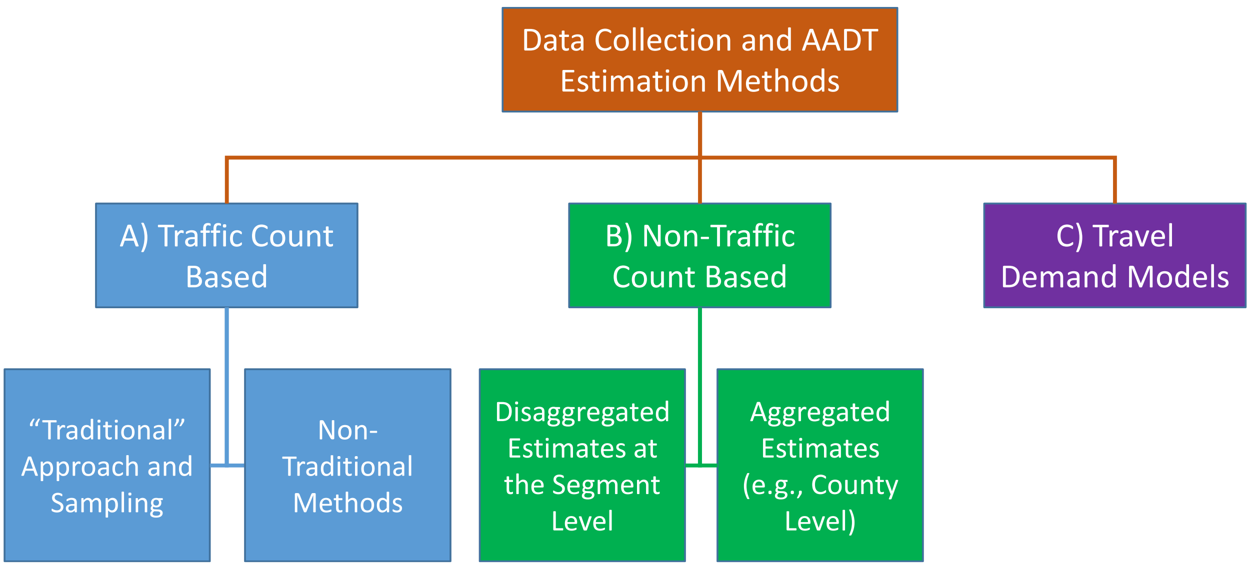

Figure 1. Types of AADT Estimation Methods.

The most important findings from Task 2 are provided below:

Traffic Count Based Systems

- The method of using sample (short-duration) counts combined with estimates for uncounted roads is the most prevalent nationally and internationally. Some states use a single AADT value for all uncounted segments within a group of roads that have common characteristics. The AADT value is typically derived from the counted segments.

- Literature shows that a more accurate method for uncounted segments is to develop regression models using existing counts along with socioeconomic data, network connectivity, and other information to predict an AADT value for each uncounted segment. The downside of these approaches are increased calculation complexity, need for non-traffic data, and good statistical knowledge. Further, the assumptions made for each method may affect the accuracy of the results to some degree.

- In many cases, NFAS roads have limited or no coverage of continuous count sites, thus determining accurate seasonal factors can be a challenging task. The absence or inadequate coverage of permanent counts often limits options that a transportation agency has to estimate AADT. As a result, transportation agencies often simply apply adjustment factors from higher functional classes within the vicinity of a road. This practice tends to introduce high AADT estimation errors because NFAS roads have significantly different characteristics and travel patterns compared to higher functional classification roads.

Non-Traffic Count Based Systems

- Non-traffic count based models can produce AADT estimates either for individual roadway segments or for a group of roads that have common characteristics. In general, the models that generate disaggregated estimates at the segment level have superior performance compared to the models that produce aggregated estimates. However, models for individual roadway segments require disaggregated data, which are often difficult to collect or not readily available. Models for a group of roads that have common characteristics are suitable for groups of uncounted segments and applications that do not require high levels of AADT accuracy.

- While roadway characteristics can have a significant role in predicting AADT for higher functional classification roads, they do not adequately explain AADT variations on NFAS roads. The main reason is that the roadway characteristics of NFAS roads are relatively homogenous compared to higher classifications roads. The AADT on NFAS roads is typically a function of other factors (e.g., socioeconomic, population, and land use).

- The traffic on local roads may not be highly correlated with that on a freeway, even if the roadways are closely located. Traffic patterns on local roads largely depend on characteristics of the local communities, whereas traffic patterns on exclusive facilities such as freeways that have limited access control are more related to regional through traffic.

Travel Demand Models

- In general, studies that developed TDMs reported improvements in AADT accuracy over other non-traditional methods.

- A limitation of TDMs is the number of links that can be built into a network, which is limited. TDMs tend to create bias toward major roads and omit most minor roads. This results in incomplete models that do not contain all links of the network.

- The cost to develop and implement TDMs is higher compared to non-traditional models, and the implementation of TDMs is also more difficult compared to regression models. Creating transportation analysis zones (TAZs) for large rural areas can be time consuming and the model implementation requires knowledge of the theory behind travel demand modeling and model calibration.

CHAPTER 1—INTRODUCTION

1.1 BACKROUND

One of the most widely used traffic estimates is annual average daily traffic (AADT). The Federal Highway Administration (FHWA) defines AADT as the "*total volume of vehicle traffic of a highway or road for a year divided by 365 days*" (*1*). Transportation agencies use AADT to make better informed decisions, meet data reporting requirements, and support various planning, safety analysis, pavement design and maintenance, highway design, operations, and environmental functions. For example, with the development of several new safety analysis tools such as AASHTOWare's Safety Analyst™ and the *Highway Safety Manual* (HSM), AADT has become an increasingly important data input in safety analysis. In addition, agencies use AADT to calculate vehicle miles traveled (VMT), which is the basis for allocating funds toward roadway maintenance and construction of safety improvement projects. FHWA requires states to report AADT through the Highway Performance Monitoring System (HPMS) for the full extent of mainlines, samples and ramps on all Federal-aid roads (*2*).

Because AADT is used in numerous applications, including safety analysis, states traditionally devote significant resources to improve their data collection programs and AADT estimation practices. Transportation agencies typically employ permanent and portable traffic counting equipment for collecting traffic volume data. The permanent stations, known as continuous count sites (CCSs) collect data 24 hours a day and 7 days a week for either all days of the year or at least for a seasonal collection (*1*). The high cost associated with the installation, operation, and maintenance of CCSs allows states to install CCSs only on select locations of the roadway network. For segments without a CCS, agencies might collect seasonal or short-duration count data using portable traffic recorders (PTRs) or counters. Strategic placement of CCSs in combination with strategic seasonal counts allows agencies to adjust and expand these short-duration counts to AADT estimates.

Most states estimate AADT using variations of a "traditional" method that was first introduced by Drusch in 1966 (*3*) and is recommended by FHWA's *Traffic Monitoring Guide* (TMG) (*1*). This method involves the following major steps:

- Gathering and processing traffic data from CCSs and computing axle, seasonal, monthly, and day-of-week adjustment factors for each site.
- Establishing monthly pattern groups that are reasonably homogenous. Creating groups of continuous sites, also known as factor groups, can be based on traditional grouping techniques (i.e., functional classification and geographical stratification), cluster analysis, and volume factor groups.

- Computing temporal adjustment factors (e.g., monthly average daily traffic [MADT]) for each group of CCSs. The group adjustment factors are computed from the individual factors of the sites contained in each group.

- Assigning short-duration counts to the previously determined factor groups. States typically base the assignment task on the location, functional class or other characteristics of the roadway section where a count was taken (e.g., daily patterns and ADT).

- Multiplying the average daily traffic (ADT) of a short-duration count with the appropriate group adjustment factor(s) to generate an AADT estimate for the counted segment (*1*).

1.2 PROBLEM STATEMENT

The United States Department of Transportation (USDOT) published the Highway Safety Improvement Program (HSIP) and the Safety Performance Management Measures Final Rules in the Federal Register on March 15, 2016. The new rules address the requirements of the Moving Ahead for Progress in the 21st Century Act (MAP-21) and the Fixing America's Surface Transportation Act (FAST Act). The HSIP Final Rule updates the existing HSIP requirements under 23 CFR 924 (*4*). The *Model Inventory of Roadway Elements (MIRE) – Fundamental Data Elements (FDEs) Cost Benefit Estimation* (FHWA-SA-16-035) developed the cost estimate for the HSIP Final Rule provision for the collection of the MIRE – FDEs (*5*).

According to the new rules, states must collect and use a subset of MIRE FDEs, including AADT, for all public paved roads. "All public roads" include Federal aid, non-Federal aid, on state, and off state system roadways including rural minor collectors (HPMS Functional Class 6) and local roads (HPMS Functional Class 7). States must define anticipated improvements to collect MIRE FDE in their Traffic Records Strategic Plan by July 1, 2017 and, by September 30, 2026, data must be accessible for all public roads.

The high expense and time consuming nature of traffic data collection tasks has raised concerns about states' ability to conduct extensive short-duration counts on lower functional classification roads. The HSIP requires agencies to reduce the number of fatalities and serious crashes on all public roads. Therefore, agencies tend to focus their data collection efforts on high-volume roads such as interstates, freeways, expressways, and arterial roads that typically pose significant safety challenges compared to lower-volume roads.

According to FHWA *Highway Statistics*, higher functional classification roads made up approximately 31 percent of the total road mileage in the U.S. in 2011 and 2012, but they accounted for 80.3 percent of all the fatalities in the country (*4*). By comparison, the remaining

19.7 percent of fatalities occurred on local roads, which covered approximately 69 percent (2,821,867 miles) of U.S. roadway miles. As a result, it is imperative to find a cost-effective solution to accurately estimate AADT on lower volume roads.

FHWA's *HPMS Field Manual* does not require any specific procedure for sampling traffic volumes in lower functional classification roads that include rural minor collectors and both rural and urban local roads. Departments of Transportation (DOTs) often use historical traffic volume data for these roads and select, if any, their own sampling technique and AADT estimation method. When traffic count data are unavailable, agencies typically base their estimates on comparisons to similar types of roadways. This technique may not be repeated often enough to remain current and can result in major errors due to inaccurate assumptions.

Further, states supplement actual traffic volumes with AADTs estimated for uncounted segments. However, the estimated AADTs are generated using various estimation techniques, which can potentially produce significant errors. Since transportation agencies use AADT to assess safety needs and evaluate the effectiveness of safety projects, the reliability of any type of safety analysis depends on the accuracy of AADT estimates. Some estimation methods do not provide link based AADTs which is what is needed for safety analysis and implementation of corrective safety measures.

1.3 STUDY OBJECTIVE

The objective of this Task Order is to develop an informational guide on collecting traffic volume data and computing/estimating AADT for lower-volume roads. According to the Task Order Proposal Request (HSA 16-13), the lower-volume roads include rural minor collectors and local roads. However, research has shown that roads belonging to the same functional class may not necessarily have similar traffic volumes. Further, several states have expressed concern that it is not always possible to define low-volume roads based on an AADT threshold (e.g., 400 vehicles), because many of these roads have not been counted. In response to these concerns, FHWA will replace the 400 AADT threshold with the roadway functional classification to determine which MIRE FDE requirements apply to a roadway (*4*). To avoid potential confusion to the readers, this Technical Memorandum refers to rural minor collectors and rural and urban local roads as non-Federal-aid system (NFAS) roads.

This study will bridge the gap between safety and traffic data communities at State, tribal, Metropolitan Planning Organizations (MPOs), local and Federal agencies on collecting traffic volume data and estimating AADT values for use in data-driven safety analysis. The results of this project will supplement relevant research studies such as the National Cooperative Highway Research (NCHRP) Project 08-110, which aims to develop a process to analyze and improve the accuracy, reliability, and utility of project-level traffic forecasts (*6*).

This Technical Memorandum includes three chapters, which cover the following:

- Chapter 1 – Background on data collection and AADT estimation, problem statement, and study objective.
- Chapter 2 – Review of literature on data collection and AADT estimation methods for NFAS roads.
- Chapter 3 – Proposed scope of the Guide.

CHAPTER 2—DATA COLLECTION AND AADT ESTIMATION METHODS

2.1 INTRODUCTION

The goal of Task 2 (Research) is to develop recommendations for the proposed scope of the guide. To achieve this goal, the researchers gathered and reviewed available documentation related to data collection and AADT estimation for NFAS roads (i.e., rural minor collectors and local roads). The aim of this activity is twofold. First, to develop a better understanding of current practices and methods in this area. Second, to identify key elements as well as potential gaps and needs that the guide must address. This chapter provides a review of the literature focusing on data collection and AADT estimation methods for NFAS roads.

Overall, the literature review reveals many AADT estimation methods that significantly vary on sample size requirements, data inputs, level of development effort and complexity, software needs, runtime, use of geographic information system (GIS) technologies, application area, anticipated accuracy, and algorithms and statistics used. Because of these differences, the research team organized and presented these methods to allow readers to easily identify the information they need. The team divided the data collection and AADT estimation methods into three major groups as shown in Figure 2.

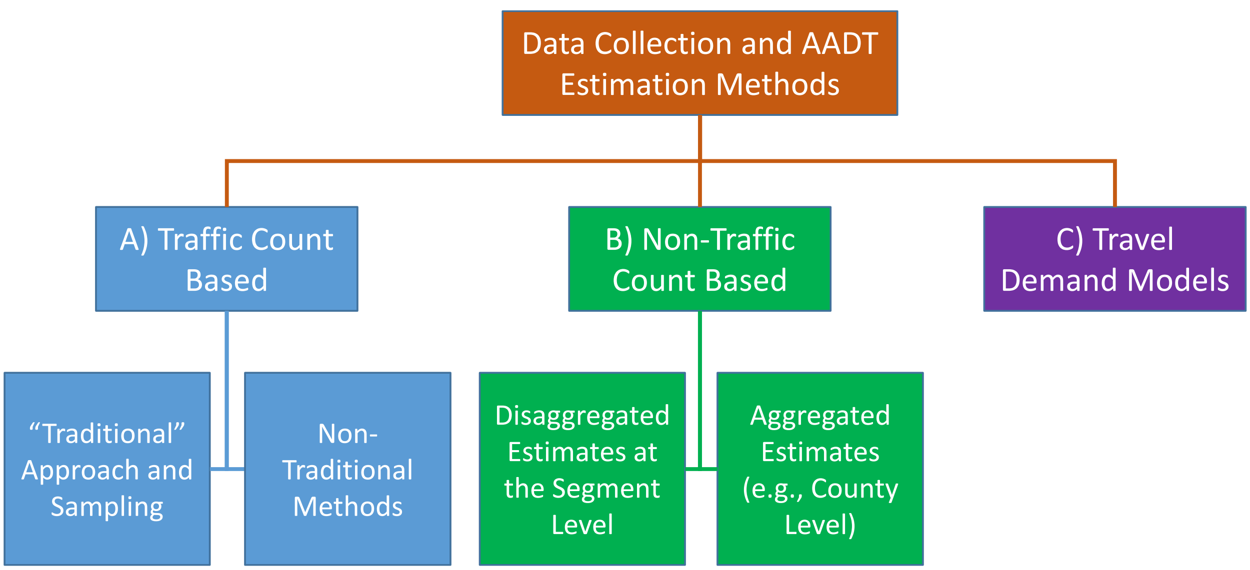

Figure 2. Types of AADT Estimation Methods.

The three groups of methods include the following:

D. Traffic count based methods. These methods use traffic volume data obtained from CCSs or PTRs or both. For clarity, we further divided these methods into two subgroups (Figure 2) as follows:

   1. Traditional approach and sampling – These methods are based on the traditional AADT estimation process. Sampling refers to the process of selecting how many and which roadway sections need to be counted. Sampling is often applied when there is limited or no availability of short-duration counts within a group of roads such as local roads and rural minor collectors.

   2. Non-traditional methods – These methods directly estimate AADT values by avoiding the errors inherent in each step of the traditional approach. This subgroup primarily includes statistical methods such as regression and machine learning methods such as neural networks.

E. Non-traffic count based methods. These methods use non-traffic data such as socioeconomic and census data (e.g., population, household size, income, number of registered vehicles, number of licensed drivers, employment), land use, topological variables (e.g., distance from permanent sites and distance from major roads), and temporal variables (e.g., day of week, weekday, weekend, month, season). As Figure 2 shows, the project team further divided the methods in this group into two subgroups as follows:

   1. Disaggregated estimates – These methods produce AADT values at the roadway segment level.

   2. Aggregated estimates – These methods produce a single AADT estimate for a group of roads that are stratified by one or more variables such as county, area, region, roadway functional class, population size, land use type or other attributes.

F. Travel Demand Models (TDMs). These models are based on network, land use, and socioeconomic data to calculate the expected demand for transportation facilities. They are widely used in many applications including estimating traffic volumes. The traditional travel demand modeling process consists of four steps: trip generation, trip distribution, mode choice, and trip assignment. TDMs incorporate mathematical equations to capture travelers' decisions and allocate trips to roads.

We identified several items for each document we reviewed:

- Type(s) of AADT estimation methods (Figure 2).

- Types of traffic data collected and/or gathered (e.g., vehicle count data, vehicle classification data, vehicle weight data).
- Traffic equipment used to collect data (e.g., CCSs, PTRs, weigh-in-motion stations).
- Other types and sources of data (e.g., U.S. Census Bureau, American FactFinder, other national, state or local agencies, etc.).
- Dependent variable(s) used in the model(s), if different than AADT (e.g., log(AADT)).
- Independent variables used, if any, (e.g., day of week, weekday, weekend, month, season, existence of holidays and special events, roadway functional class, geographical area, demographic data, socioeconomic data, distance from permanent sites, distance from major roads).
- Seasonal adjustment factors and axle correction factors used.
- Basic assumptions and parameters used in the model(s).
- Number of permanent sites and sample size used to develop and validate the model(s).
- Tools developed or used to process data and develop AADT estimation models.
- Performance measures used to quantify the accuracy of AADT estimation method(s) (e.g., variance, standard deviation, mean absolute percent error [MAPE], standard deviation of absolute error, root mean square error [RMSE], and coefficient of variation).
- Precision and confidence level associated with each method and the produced AADT estimates.
- Methods for determining recreational traffic.

The remaining portion of this chapter includes five sections:

- Sections 2.2 and 2.3 provide a review of available traffic-based and non-traffic based AADT estimation methods, respectively.
- Section 2.4 presents previous studies that developed TDMs to determine AADT estimates for NFAS roads.
- Section 2.5 describes methods for determining recreational traffic.
- Section 2.6 summarizes the most important findings from this chapter.

2.2 TRAFFIC COUNT BASED METHODS

This group of AADT estimation methods requires traffic volume data collected from CCSs, PTRs, or both. The segments where traffic volume data exist are called counted segments, whereas the remaining segments, where no continuous or short-duration count data have been collected, are known as uncounted segments. We divided these methods into traditional-based and non-traditional methods that are described in the following two sections, respectively.

2.2.1 Traditional Approach and Sampling

The traditional AADT estimation process involves the following major steps:

a) Computing adjustment factors for each CCS.

b) Establishing monthly pattern groups that are reasonably homogenous. These groups are widely known as factor groups or groups of CCSs.

c) Compute temporal adjustment factors (e.g., MADT) for each group of CCSs.

d) Assigning short-duration counts to the previously determined factor groups.

e) Estimating AADT by multiplying the ADT of a short-duration count with the appropriate group adjustment factor(s).

Sampling refers to the process of selecting the number and the location of seasonal or short-duration counts to be conducted within a roadway stratum that contains a large number of uncounted sections (e.g., local roads and rural minor collectors). Count-based approaches rely on probability sampling theory to provide ADT that is representative of roadways with certain characteristics within a study area. A probability sample is a sample drawn from a population of interest such that every object in the population has a known likelihood of being selected.

2001 – Analysis of Traffic Growth Rates

Barrett et al. developed a random sampling procedure for collecting traffic count data on local roads and streets (7). Historically, traffic counts on local roads in Kentucky are conducted for specific projects, and are not randomly selected. If these counts are used for VMT estimation, the results would be overestimated since local road counts are taken on more heavily traveled roads. The authors presented a GIS-based method to apply a grid to cut road segments into small sections and randomly select segments as locations for traffic counts.

The authors selected three counties as a case study to apply the methodology and considered grid sizes of 0.2, 0.15, 0.1, and 0.05 miles to reduce bias resulting from some segments that were longer than others. Comparing coefficients of the slope variable in the best-fit equation indicated that the grid segmentation process is better suited for rural local roads than urban local roads because they are longer and less dense. The authors recommend that grid sizes of 0.2 miles be used for rural areas and 0.1 miles for urban areas, with the caveat that attempts to use grid sizes below 0.05 miles were not successful due to computer issues. Computing power has increased significantly since this study was conducted, so it is likely that these issues could be avoided now.

2005 – Cost-Effective Reporting of Travel on Local Roads

Blume et al. developed a VMT estimation methodology that used census data from Florida and a random sampling technique (*8*). This study performed correlation analysis among travel, population density, job density, and roadway density. The density factors were used to group comparable zip code tabulation areas into sub-regions. This method allowed random samples to be taken in one sub-region to represent statewide zip code tabulation areas by considering population, job, and roadway density as stratification variables (instead of traffic volume groups recommended by HPMS). To determine the sample size, the study used the following HPMS formula:

$$n = \frac{Z^2C^2/d^2}{1 + \left(\frac{1}{n}\right)[(Z^2C^2/d^2) - 1]}$$

Where:

n = Required sample size

Z = Value of the standard normal statistic for an alpha confidence level (two-sided)

C = AADT coefficient of variation

d = Desired precision rate

N = Universe or population stratum size.

The proposed estimation methodology was able to estimate AADT with more accuracy than Florida DOT's current approach without significant increases in cost and labor. The method allowed adjustments (e.g., re-stratification) that improved the results from year to year.

2006 – Stratification of Locally Owned Roads for Traffic Data Collection

Lloyd and French determined a sampling method to collect data required for producing VMT estimates on local roads, owned by municipalities in Pennsylvania (*9*). The authors initially gathered census data at the municipal level and tried to relate them to AADT levels. The data included population, housing units, employment, total area, and roadway miles. Based on these variables, the analysts computed basic statistics for four measures: population density, household density, employment density, and housing units plus employment/roadway miles. The results of the analysis did not allow the researchers to associate these measures with local roads at a high confidence level. Instead, they stratified the segments by county and population according to which VMT is typically estimated. The end result was four strata of roads for each county:

- Rural
- Small Urban (population between 5,000 and 49,999)
- Urbanized (population between 50,000 and 199,999)
- Urbanized (population 200,000 and greater)

This study used an HPMS sampling formula to perform the analysis. The authors used Microsoft Excel to randomly select a sample of roadway segments. At a 90 percent confidence level and 10 percent precision, the final scheme contained 7,171 count stations spread proportionally over 152 strata.

2009 – Estimating Traffic on Local Road Systems (Phase 3 Report)

As part of a larger study, Michigan DOT conducted a study to assess various options for estimating VMT on local roads and the costs associated with each option (*10*, *11*, *12*). The authors noted that the lowest cost and least accurate alternatives are appropriate for HPMS reporting, while high cost alternatives are appropriate for other planning and funding uses. The six options presented are as follows:

- Retain existing method for estimating minor road VMT. At the time the study was published, Michigan grouped local road data by county, rural/urban code, urban area code (if urban), national functional system, Michigan legal system (state owned, rural major, rural local, city major, and city local). Michigan has not taken counts on these roads since the 1980's. Average AADT values and growth rates are applied to estimate AADT and VMT. According to the report, this method would be accurate for HPMS uses and would not result in additional costs.

- Develop a non-count based program. According to this method, the estimate of total road VMT is found and the higher system VMT based on counts is subtracted. The remaining VMT is for uncounted roads. VMT may be estimated statewide from vehicle emission monitoring programs or from fuel consumption and fuel efficiency estimates. This method is accurate for HPMS uses and Michigan estimated they would need to invest $50,000 to $100,000 in a research study to further refine this method.

- Count based, low cost, least accuracy. Michigan would retain the roadway classification scheme and take counts on rural minor collectors only. Segments would be randomly selected using a grid system. Counts would be taken for 24 hours and factored using standard practices. A simple average AADT value would be applied to uncounted segments. This method is accurate for HPMS uses. The study estimated annual costs of $50,000 to $80,000 for collecting additional counts and a one-time project to correlate local counts with permanent counts costing between $100,000 and $150,000.

- Count based, mid-cost, good accuracy. Michigan would retain the roadway classification scheme and further disaggregate road segments by commercial business district, commercial, urban, suburban, and rural. Traffic counts would be taken on randomly selected segments of rural minor collectors, rural local, and urban local roads. Counts would be taken for 24 hours and factored using standard practices. A simple average AADT value would be applied to uncounted segments. According to this study, this method is accurate for HPMS uses. The study estimated annual costs of $125,000 to $200,000 for minimum accuracy from 1,000 counts and costs between $200,000 and $300,000 for higher accuracy from 1,650 counts.

- Count based, mid-cost, better accuracy. This option is similar to the previous, except that AADTs for uncounted segments would be based on socioeconomic, network connectivity, and road characteristics data for each segment. According to the study, this method would be accurate for HPMS uses and would provide a relatively accurate method for estimating VMT at the jurisdiction level. The authors estimated annual costs of $125,000 to $200,000 for minimum accuracy from 1,000 counts and costs between $200,000 and $300,000 for higher accuracy from 1,650 counts. A one-time project costing between $150,000 and $200,000 would be required to develop and validate the AADT estimation models.

- Count based, accurate to jurisdiction level. Michigan would retain the roadway classification scheme and further disaggregate road segments by volume group. Traffic counts would be taken on randomly selected segments of rural minor collectors, rural local, and urban local roads. Counts would be taken for 24 hours and factors would be developed for each of Michigan DOT's 14 districts. AADTs for uncounted segments would be based on socioeconomic, network connectivity, and road characteristics data for each segment. This method would be accurate for HPMS uses and would provide the most accurate estimations of VMT based on the options described. The study estimated taking 8,300 counts at a cost of $1,000,000 to $1,600,000 annually and a one-time project costing between $150,000 and $200,000 to develop and validate the AADT estimation models.

2014 – HPMS Field Manual

Chapter 6 of the HPMS *Field Manual* discusses sampling techniques, but does not addresses local roads or rural minor collectors (*2*). HPMS uses 12 volume groups based on AADT ranging from less than 500 to more than 250,000. The volume groups are the same for all sample functional classes, which do not include local roads or rural minor collectors. Precision levels are developed for rural, small urban, urbanized areas with populations less than 200,000, and urbanized areas with populations greater than 200,000 for all functional classes except local roads. There is no precision level given for rural minor collectors either. Procedures for determining sample size based on confidence level are also provided.

2016 – Improved Annual Average Daily Traffic Estimation Processes

Jessberger et al. evaluated a new AADT estimation method that incorporated any time increment counts (*13*). This study examined the bias and precision of four different AADT estimation methods from hundreds of permanent traffic counting sites. It used more than 48 million records from the FHWA Travel Monitoring Analysis System (TMAS), covering 14 years of data from 2000 to 2013. The four methods include: simple average; the American Association of State Highway and Transportation Officials (AASHTO) method; AASHTO with day-of-week, month, and year adjustment factors; and the Highway Policy Steven Jessberger Battelle (HPSJB) Method:

Simple average:

$$AADT_{SA} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n} VOL_i$$

Where,

VOL_i = total traffic volume for i^{th} day of the year.

i = occurrence of a day in a year (i = 1, . . . , n) for which traffic volume is available.

n = number of days of the year for which traffic volume is available (n ranges from 1 to 366, depending on number of days in the year and data availability).

In this method, the variance increases significantly with the increase of missing counts.

The AASHTO method uses the advantage of the known periodicity of traffic volume by both month and day of week.

AASHTO method:

$$MADT_{AASHTO_m} = \frac{1}{7}\sum_{j=1}^{7}\left[\frac{1}{n_{jm}}\sum_{i=1}^{n_{jm}} VOL_{ijm}\right]$$

$$AADT_{AASHTO} = \frac{1}{12}\sum_{m=1}^{12} MADT_{AASHTO_m}$$

Where

VOL_{ijm} = total traffic volume for i^{th} occurrence of the j^{th} day of the week within the m^{th} month,

i = occurrence of a particular day of the week in a particular month (i = 1,…, n_{jm}) for which traffic volume is available,

j = day of the week (j = 1, 2,…,7),

m = month (m = 1, 2,…, 12),

n_{jm} = the count of the j^{th} day of the week during the m^{th} month for which traffic volume is available (n_{jm} ranges from 1 to 5, depending on day of the week, month, and data availability).

The mathematical bias of this method compared with the simple average method is minor because it equally weights each day of the week within each month, and then equally weights each month. The AASHTO method overweights shorter months and underweights longer months by equally weighting all MADT values.

AASHTO with day-of-week, month of year adjustment factors:

$$MADT_{AASHTO(MOD)_m} = \frac{\sum_{j=1}^{7} W_{jm}\left[\frac{1}{n_{jm}}\sum_{i=1}^{n_{jm}} VOL_{ijm}\right]}{\sum_{j=1}^{7} W_{jm}}$$

$$AADT_{AASHTO(MOD)} = \frac{\sum_{m=1}^{12} d_m * MADT_{AASHTO(MOD)_m}}{\sum_{m=1}^{12} d_m}$$

Where

i = occurrence of a particular day of the week in a particular month (i = 1, … , n_{wm}) for which traffic volume is available.

W_{jm}= weighting for the number of times the j^{th} day of the week occurs during the m^{th} (4 or 5); the sum of the weights in the denominator is the number of calendar days in the month (28, 29, 30, or 31).

d_m = weighting for the number of days (28, 29, 30 or 31) for the m^{th} month in the particular year.

HPSJB Method:

$$MADT_{HP_m} = \frac{\sum_{j=1}^{7} W_{jm} \sum_{h=1}^{24} \left[\frac{1}{n_{hjm}} \sum_{i=1}^{n_{hjm}} VOL_{ihjm}\right]}{\sum_{m=1}^{7} W_{jm}}$$

$$AADT_{HP} = \frac{\sum_{m=1}^{12} d_m * MADT_{HP_m}}{\sum_{m=1}^{12} d_m}$$

Where:

VOL_{ihjm} = total traffic volume for the i^{th} occurrence of the h^{th} hour of day within j^{th} day of the week during the m^{th} month.

i = occurrence of a particular hour of day within a particular day of the week in a particular month (i = 1, …, n_{hjm}) for which traffic volume is available.

h = hour of the day (h = 1, 2, …, 24) or smaller time increment.

n_{hjm} = number of times the hth hour of day within the j^{th} day of the week during the m^{th} month has available traffic volume (n_{hjm} ranges from 1 to 5, depending on hour of day, day of week, and data availability).

The study showed that the AASHTO method does not take full advantage of all data because traffic volumes are often collected in hourly increments. For partial missing values, the traffic volume for the entire day is excluded from this method. The HPSJB method used a natural extension of the AASHTO modified method by using traffic volume data for each available hour of the day or smaller time increments. The HPSJB method performed better than the AASHTO method in terms of bias and precision.

2.2.2 Non-Traditional Methods

Several studies have developed non-traditional models for estimating traffic volumes on NFAS roads. The main characteristic of these methods is that they allow the direct estimation of AADT by avoiding the main steps, hence the errors, of the traditional approach. These methods typically use a combination of traffic variables such as ADT and non-traffic factors such as socioeconomic data.

The main types of non-traditional AADT estimation methods include statistical models and machine learning. Statistical models rely on assumptions. Any distortion from the assumptions can generate biased results. On the other hand, non-parametric tools like machine learning are based on parameters that users need to define.

2001 – Application of Neural Networks to Estimate AADT on Low-Volume Roads

Sharma et al. developed neural network models to predict AADT on low-volume roads and determine the most effective duration and frequency of short-duration counts in Alberta, Canada (*14*). The study data included hourly traffic volumes from 55 CCSs located on lower functional classes. The authors divided the sites into three groups based on their AADT (AADT≤500, 500<AADT≤750, and 750<AADT≤1000) and created one model for each group. The training dataset contained hourly volumes from sample counts extracted from the available CCSs.

The models were based on a multilayered, feedforward, and back-propagation scheme for supervised learning. The number of neurons in the input layer was equal to the total number of hourly volumes included in a sample count. The number of neurons in the hidden layer was equal to half of that in the input layer. There was one neuron in the output layer of each model.

The study concluded that a single 48-hour count resulted in the highest average 95th percentile error (103). However, conducting two 48-hour counts within a year can reduce the 95th percentile errors to about 25 percent in most cases. The results of a similar analysis using three 48-hour counts indicated that the accuracy of AADT estimates was not improved significantly. Another finding of this study was that two 48-hour counts produced better estimation results than those using two 24-hour counts annually. However, two 72-hour counts did not offer an advantage over two 48-hour counts (*14*).

2004 – The Effects of Errors in Annual Average Daily Traffic Forecasting: Study of Highways in Rural Idaho

Dixon used the classification and regression tree (CART) method to reduce the variability in the AADT annual growth rate (*15*). The study used data from different sources, including CCS stations, portable traffic counts, Idaho DOT's entire AADT database, and economic and demographic data. The CCS datasets displayed four different patterns:

- High growth in a few counties.
- Dissimilar growth within a single county.
- Dissimilar growth within a single functional class.
- Higher traffic growth with lower population.

CART outperformed other models with variables such as functional class, county population, annual county population growth rate, and AADT. The results showed that the CART method provided acceptable mean absolute percent errors, which were less than 10 percent for a 10-year forecast.

The study concluded that low-volume roadways showed lower traffic annual growth rates than higher volume highways. The finding also showed that the CART method performed well in classifying the CCS stations into groups with similar characteristics while reducing the variability of the AADT annual growth rates. This method showed acceptable performance with limited data and it can be used to update growth factors more frequently compared to other methods.

2011 – Data Mining Methods for Traffic Monitoring Data Analysis: A case study

Gecchele et al. performed a comparative analysis of clustering methods to estimate AADT (*16*). The authors performed a comprehensive analysis of the TMG procedures for more precise results. The researchers collected data from 54 CCS sites located on the rural road network of the Province of Venice, Italy in 2005. The network links were mostly two-lane roadways. This study adopted the methodology from Sharma et al. (*14*).

According to the results, the estimation errors for passenger cars were acceptable, but those for trucks were higher. The study showed that the model with variable volumes, equal shape and orientation of the coordinate axes performed better than other models.

2012 – Annual Average Daily Traffic Estimation from Seasonal Traffic Counts

Gastaldi et al. (2012) presented an approach of estimating AADT from a one-week seasonal traffic count in the Province of Venice, Italy (*17*). The research team collected data from 50 CCS sites located on rural roads.

The method preserved the framework of the FHWA procedure that includes four major steps:

- Grouping CCS sites with the Fuzzy C-means algorithm.
- Assigning road segments with counts to predefined groups.
- Calculating uncertainty within each group.
- Estimating AADT as a weighted average of seasonal count values.

The study results showed that traffic volumes collected on weekdays resulted in the most accurate estimates. The authors also analyzed recreational travel counts. The results showed that recreational roads have larger errors due to higher unpredictability of traffic volumes.

2014 – Spatial Interpolation of Traffic Counts Based on Origin–Destination Centrality

Lowry developed an ordinary least squares (OLS) model to predict segment-specific AADT values for the transportation network of a small community in Idaho (*18*). The OLS model included explanatory variables derived through network analysis using GIS. Lowry constructed the variables based on a modified network analysis metric, called stress centrality, that quantifies the topological importance of a link in a network. The original formula of this metric is:

$$Stress\ Centrality_e = \sum_{i,j \in V} \sigma_{i,j}(e)$$

Where:

V = the set of all nodes in a network

$\sigma_{i,j}$ = the shortest path from node i to node j

$$\sigma_{i,j}(e) = \begin{cases} 1, if\ link\ \mathrm{e}\ is\ used\ in\ \sigma_{i,j} \\ 0, otherwise \end{cases}$$

In the context of a street network, stress centrality is the number of times a link would be used if someone were to travel from every node to every other node via the shortest path. As with other forms of centrality, the calculation can be made for links or nodes. The author modified this metric by: a) limiting the calculation to include only a subset of nodes that the analyst identified as origins and destinations; and b) augmenting the calculation with multipliers associated with each origin and destination. The author defined multiplier values for every parcel throughout the network based on values provided in the Institute of Transportation Engineers *Trip Generation Manual*. The form of the origin destination (OD) centrality used in the study was as follows:

$$OD\ Centrality_e = \sum_{i \in I, j \in J} \sigma_{i,j}(e)\, M_i M_j$$

Where:

$I \in V$ and $J \in V$ = subset origin and destination nodes

M_i and M_j = multipliers

Lowry identified three possible pairs of origins and destinations and used them as predictors in the OLS model: 1) internal-to-internal OD centrality; 2) internal-to-and-from-external OD centrality; and 3) external-to-external OD centrality. To estimate OD centrality and AADT, he created a GIS toolbox, which required four input files: 1) the street network; 2) the observed AADT counts; 3) land use parcels; and 4) external gate points (the boundary locations on the street network). The study used 90 percent of the data to train the OLS model and the remaining 10 percent for validation purposes. The dependent variable was a Box-Cox transformed form of the AADT.

The R^2 values of the model were close to 1.0 and the median absolute percent error of the calibration and validation datasets were 34 percent and 22 percent, respectively. The study also presented the results of sub-sampling validation, which revealed that similar level of AADT accuracy could be achieved using approximately one-fifth of the traffic count data. However, the method is suited for small and medium-sized community networks. Gathering input data for all NFAS roads within a state's network could require extensive time and resources.

2.3 NON-TRAFFIC COUNT BASED METHODS

These methods use non-traffic data such as socioeconomic, census, land use, topological, and temporal variables. We divided the non-traffic count based methods into two broad categories:

a) Disaggregated estimates – These methods produce AADT estimates at the roadway segment level using segment-specific data and information from larger geographical units.

b) Aggregated estimates – These methods produce a single AADT estimate for a group of roads that are stratified by one or more variables such as county, area, region, roadway functional class, population size, land use type, census block or other attributes. These methods are suitable for cases where there is limited availability of disaggregated information (e.g., at the segment or corridor level).

The review showed that several studies have applied statistical models and machine learning algorithms to generate both disaggregated and aggregated AADT estimates.

2.3.1 Disaggregated Estimates

This section reviews methods that provide estimates at the segment level.

1998 – An Annual Average Daily Traffic Prediction Model for County Roads

Mohamad et al. developed a traffic prediction model for county roads in Indiana (*19*). This study used count data from 40 counties with three to four 48-hour traffic counts taken in each county. By using seasonal adjustment factors, this study converted short-duration counts to AADT values. The research team used three variable selection methods to determine significant predictors for the prediction model.

The analysts considered location type, easy access to highways, county population, and total arterial mileage of a county as predictor variables. A reduced model with only four significant predictors yielded a model with an R^2 value of 0.77 at a 95 percent confidence interval. The model was:

$$log_{10}(AADT) = 4.82 + 0.82 * \text{location type} + 0.84 * \text{easy access to highway} + 0.24 * \text{county population} - 0.46 \times log_{10}(\text{total arterial mileage of a county})$$

The research team validated the models using data from eight randomly selected counties that were not used in the model development process. The mean squared prediction error (MSPR) was 0.051. The study concluded that the model was adequate for predicting traffic volumes.

1999 – Estimation of Annual Average Daily Traffic for Nonstate Roads in a Florida County

Xia et al. developed a methodology to estimate AADT for non-state roads in urbanized areas in Florida (*20*). For this study, the research team used a multiple linear regression model to estimate AADT on non-state roads. The model used a large sample size (data from 450 count stations in Broward County) and considered up to 12 preliminary variables. GIS was used to aggregate and convert digital socioeconomic data into forms suitable for statistical analysis.

The results of the analysis showed that the most important contributing predictors were roadway characteristics, such as the number of lanes, functional classification, and area type. Socioeconomic variables such as population, housing units, vehicle ownership, employment status, and school enrollment in the surrounding area, have an insignificant impact on AADT. The final model was:

$$ADT = -10759 + 4737.44 * L + 5071.13 * FCLASS1 + 1274.17 * AREA1 + 0.15 * AUTO - 816.21 * ACCESS2 - 0.15 * SEREMP$$

Where,

$ACCESS2$= accessibility to county roads,

L = number of lanes,

$FCLASS1$ = function classification,

$AREA1$= area type,

$AUTO$ = automobile ownership, and

$SEREMP$= service employment.

The study separated the data into two groups: one group for model development (training dataset), and another group for model validation (testing dataset). The results showed that 50 percent of the training data produced errors less than 20 percent, while 85 percent of the test data results in errors lower than 40 percent. The final model explained 63 percent of the ADT variability.

2003 – Estimating Annual Average Daily Traffic from Satellite Imagery and Air Photos: Empirical Results

McCord et al. used high resolution satellite imagery as a source of vehicle information for AADT estimation (*21*). The researchers determined vehicle density from images and converted density to a short-duration volume. The equation used to estimate AADT is the following:

$$AADT^{image} = \frac{(N^C + N^T)\dfrac{N^C U^{C'} + N^T U^{T'}}{N^C + N^T}}{L \times 24 \times F^H(h;\ f) \times F^{MD}(m,\ d;\ f)}$$

Here,

N^C= number of cars

N^T= number of trucks

$U^{C'}$= weighted average of car speed

$U^{T'}$= weighted average of truck speed

$F^H(h;\ f),\ F^{MD}(m,\ d;\ f)$= adjustment factors

The authors compared differences from satellite images and air photos to the corresponding ground-based estimates. The results showed that the empirical errors were small, meaning that one can reduce AADT estimation errors and spatial sampling efforts by including satellite image data with traditional data. The differences between the image-based and the ground-based estimates with higher confidence were smaller, implying that the image-based estimates may be better than what is indicated in the distribution of differences. The findings also suggested that the differences are smaller for images leading to longer equivalent traffic count duration.

2008 – Assignment of Estimated AADT on All Roads in Florida

Pan developed six linear regression models to predict AADT estimates for state, county, and local roads in Florida (22). Based on the population in each county, the author divided the counties into three groups:

1) Large metropolitan area (population > 400,000)

2) Small-medium urban area (100,000 < population < 400,000)

3) Rural area (population < 100,000)

Within each group, Pan developed one model for state highways and county roads, and a second model for local roads. The dependent variable of all models was the AADT of traffic counts. After compiling several socioeconomic, roadway, and land use variables, Pan applied stepwise regression to select the most significant independent variables for each model. The final equations and the adjusted R^2 of the models for rural state/county highways and rural local streets were as follows:

Rural area, state/county highway model (adjusted R^2 = 0.378):

*AADT = 3015.747 + 3878.551*LOCATION + 17.722*VEHICLE + 57.072*MUNICIPALITIES − 1656.733*AGRICULTURE + 22.293*LABORFORCE − 1.931*SALES − 3312.919*RECREATION – 2324.493*INDUSTRIAL + 33.239*POPULATION – 748.708*RESIDENTIAL*

Rural area, local street model (adjusted R^2 = 0.418):

*AADT = 1225.505 + 62.168*POPULATION + 1458.501*LOCATION -1445.085*AGRICULTURE – 1017.873*RESIDENTIAL*

To validate the models, Pan compared actual traffic count data from three randomly selected counties against AADT estimates from each model. According to the results, the MAPEs varied between 32 percent (rural area, state/county highway model) and 160 percent (large metropolitan area, local street model). The two rural models resulted in the lowest MAPEs. However, the validation datasets used to test the two models had significantly fewer sample counts compared to the other four validation datasets. This limitation did not allow a direct and non-biased comparison of the models' performance.

2013 – Spatial Prediction of Traffic Levels in Unmeasured Locations: Applications of Universal Kriging and Geographically Weighted Regression

Selby and Kockelman used universal kriging for spatial prediction of AADT in uncounted locations in Texas (*23*). Universal kriging connects known non-state information (lane count, population, etc.) influencing counts and road networks. The research team obtained traffic counts from Texas DOT's 2005 roadway network database. The dataset included counts from 28,000 geocoded locations. ArcGIS was used to include variables such as speed limit, number of lanes, and functional class. The authors developed several population accessibility indices for each of the 4,388 U.S. Census tracts in Texas by using population and GIS files. The models included data from both large metropolitan areas (Houston and Dallas-Fort Worth) and sparsely populated lands (primarily West Texas).

This study compared results based on Euclidean distances to those using network distances. Both types of distances allowed for strategic spatial interpolation of count values while controlling for each roadway's functional classification, lane count, speed limit, and other site attributes. The findings showed that errors tended to be lower at locations with higher counts and more nearby count locations.

2015 – Wyoming Low-Volume Roads Traffic Volume Estimation

Apronti et al. developed a linear regression model and a logistic regression model for predicting ADT on low-volume roads in Wyoming (*24*). The authors estimated the ADT (response variable) from 72-hour short-duration traffic counts that were conducted in 22 randomly selected counties from 2012 to 2014. They used GIS tools to process U.S. Census data and satellite images (e.g., Google maps and a satellite image downloaded from ESRI's website) to gather pavement, roadway accessibility, and land use data. The authors initially considered in the model development process 13 independent variables that included the following:

- Pavement type – paved roads were coded as 1 and unpaved roads as 0.
- Access – roads with direct access to primary or secondary roads were coded as 1, whereas roads without direct access were coded as 0.
- Land use – the land use categories included industrial areas, agricultural cropland, subdivisions, and agricultural pastureland. Subdivision was the reference category, while dummy variables (0 and 1 values) were used for the other three categories.
- Population by census block.
- Population by census block group.
- Number of households by census block group.
- Number of employed civilians by census block group.
- Number of housing units by census block group.
- Per capita income by block group (in ten-thousands).
- Employment density – employment number divided by the area of each block group.
- Housing unit density – total housing units in a census block group divided by the area of the block group.
- Income density – per capita income of a census block group divided by the area of the block group.
- Population density – population within each census block group divided by the area of the block group.

The analysts created a training and a validation dataset using data from 13 and 9 counties, respectively. Both regression models were developed using the Minitab statistical software. After exploring several combinations of independent variables the authors developed a linear regression model that included four predictors: pavement type, access to highways, land use, and population by block group. The dependent variable was the log transformation of the ADT. The adjusted R^2 of the model was 0.64 and its percentage root mean square error was 73.4 percent. The correlation between the actual and the estimated log(ADT) values were 0.69 and 0.61 for the training and the validation data sets, respectively.

The goal of the logistic regression model was to determine the probability of a road belonging to one of five ADT thresholds: ADTs less than 50, 100, 150, 175, and 200. For this purpose, the authors developed five equations; one for each threshold. The independent variables included land use, pavement type, total household, employment, employment density, per cap income density, house density, and population density. The output of each equation fed a sixth equation that converted the odds into probabilities, as follows:

$$Probability(AADT < Threshold\ i) = \frac{Odds_{Ti}}{1 + Odds_{Ti}}$$

Where:

$Odds_{Ti}$ is the calculated odds for threshold i.

Roads with calculated probabilities of 50 percent or greater were considered to be falling within the threshold of interest and vice versa. According to the validation results, the percentage accuracies in predicting the ADT thresholds ranged between 78 percent and 89 percent. However, the study did not report other types of errors such as RMSE.

In general, the study concluded that the two regression models are inexpensive, simple, and easy to implement. The authors recommended the use of the models when quick estimates of traffic volumes are required, but not when a high level of prediction accuracy is needed. The regression models relied heavily on the independent variable pavement type to predict ADT, however, this can raise questions because the decision to pave a road depends on the ADT.

2.3.2 Aggregated Estimates

To meet HPMS reporting requirements and overcome potential lack or limited availability of traffic counts on lower functional class roads, several studies have developed models that produce a single AADT estimate for a group of roads, known as strata of roads. The objective in identifying one or more stratification variables is to:

a) Determine a good surrogate for AADT, which can yield a stratified population that can be similar to the one directly stratified by AADT; or

b) Stratify the roads according to the structure in which the results need to be reported (*2*).

The second type of stratification yields groups that can aid reporting processes and relevant requirements, but may contain roads that have different AADTs creating high variability within each group. Based on previous works, the most common ways of stratifying NFAS roads are by county, region, census block, urban and rural roads, paved and unpaved roads, land use, and functional class.

1999 – Estimation of AADT for Off-System Roads in Florida

Shen et al. developed four multiple linear regression models to estimate AADT values for off-system roads in Florida (*25*). The dependent variable was the AADT derived from 118 short-duration count stations. The predictors included types of land uses, urban population sizes, and unique characteristics shown by some of the cities. Each model produced aggregated estimates at a different geographical scale. The four regression equations and the corresponding adjusted R^2 values are:

Statewide model (adjusted R^2 = 0.2538):

*ADT = 9643. 704 + 0.0146*POP - 0.155*LABOR - 0.18123*INCOME + 0.00000514*TAXABLE + 0. 05871*VEHICLES*

Rural area model (adjusted R^2 = 0.3486):

*ADT = 4853.489 + 0.1226*POP + 0.2619*LABOR - 18.9302*LANEMILE - 0.003234*VEHICLES*

Small-medium, urban area model (adjusted R^2 = 0.6937):

*ADT = -13418 + 6770.23*LANES + l580.14*ATYPE1 + 2.85*COM_EMP + 1.78*HOT_OCC*

Large metropolitan area (Broward County) model (adjusted R^2 = 0.6069):

*ADT = -12886 + 4689.86*LANES + 5227.57*FCLASS1 + 1388.27*AREA1 + O.15*AUT0 - 1224.06*ACCESS2*

Where:

POP = the total population within a county.

LABOR = the total labor force within a county.

INCOME = the per capita income of a county.

TAXABLE = the taxable sales of a county.

LANEMILE = the total lane miles of state roads in a county.

VEHICLES = the number of automobile registration a county.

LANES = the number of lanes in both directions.

ATYPE1 = land use type.

COM_EMP = commercial employment in a TAZ.

HOT_ OCC = population in hotel/motels in a TAZ.

FCLASS1 = functional classification of a roadway.

AUTO = the estimated total number of automobiles within a certain distance of a count station.

ACCESS2 = this variable was coded as 1 when other county roads were nearby, and 0 otherwise.

According to the adjusted R^2 values and the estimation errors, the first two models did not perform adequately. Some important variables that could possibly explain ADT variations within a county were not included in the models. Another possible reason for the low performance of the rural area model was associated with the limited sample size (27 data points) of the training dataset. By comparison, the statewide model was trained using 107 records. The large variances in the size and characteristics of the counties also contributed to the low R^2 (*25*). The small-medium urban area and the Broward County model performed significantly better than the first two, because the training datasets were larger and the predictors were disaggregated at the transportation analysis zone (TAZ) level. The Broward County model demonstrated that roadway characteristics were more important in explaining variation in ADT compared to socioeconomic factors.

2000 – Estimation of Traffic Volume on Rural Local Roads

Seaver et al. developed a method to determine traffic volumes on non-state roads in Georgia (*26*). The research team randomly selected 80 of the 159 counties in Georgia with four road types (Non-Atlanta urban areas, small urban areas, all rural areas-paved road, or rural areas-unpaved road). The authors initially considered 45 variables from 80 counties. The variables were grouped to eight categories:

- Population demographics (population density, population percentage changes, persons per household).
- Education.
- Transportation (travel time to work, means of transportation to work, leaving time for work).
- Income (per capita income and median household income).

- Employment (types of employment, unemployment, and place of employment).
- Farming.
- Urbanization (rural or urban counts of people and distances to MSAs).
- Housing.

The conventional models for rural roads performed poorly. To improve the performance, the study stratified the counties based on their location inside and outside a metropolitan statistical area to anticipate differences. The study examined the models for each road type. Finally, the team selected the optimal multiple regression model for each road type within or outside a metropolitan statistical area by considering factors such as good interpretations, minimal collinearity, robustness, low number of independent variables as possible, and maximum predictability by using a jackknife coefficient of determination called $R^2_{predict}$:

$$R^2_{predict} = 1 - \left[\sum_{i=1}^{n} \left(y_i - \hat{y}_{(-i)}\right)^2 / \sum (y_i - \bar{y})^2 \right]$$

Where:

i = the i^{th} observation being deleted n times.

This R^2 withheld each observation, fitted the model on the remaining (n – 1) observations like the leave-one-out method, and predicted the withheld each observation with the model based on only (n-1) observations. The fitting and prediction tasks were done *n* times for each observation. The $R^2_{predict}$ value was 0.89 for some clusters. The research concluded that the results were helpful in estimating AADT for certain subgroups of traffic volumes on local roads and the models were capable of estimating AADT in some counties that were not included in this study.

2001 – Analysis of Traffic Growth Rates

In addition to the development of a random sampling procedure, Barrett et al. developed regression models to estimate ADT using county level variables that include population, average per capita income, employment, county-wide total earnings, and licensed drivers (7). The R^2 values for the variables in the urban model were all less than or equal to 0.002, and the R^2 values for the variables in the rural model were all greater than 0.41 with the exception of income, which had an R^2 of 0.263. Based on these results, the urban model was not considered useful, whereas the rural model was more promising in predicting ADT values on local roads. The study presented another set of regression models that included variables specific to a location along a road. These variables included local road densities within 1, 2, and 5 miles, as well as the straight line distance between the road location and the nearest freeway, city, and state road. The regression models were not significant for urban or rural areas.

2004 – Alternatives for Estimating Seasonal Factors on Rural and Urban Roads in Florida

Zhao et al. performed regression analyses to identify possible variables that affect monthly adjustment factors in selected rural areas in Florida (27). The independent variables of the analysis included the following:

- Roadway data: number of lanes, truck factor, functional class, and direction of travel.
- Demographic data: population, population density, population by age group.
- Socioeconomic characteristics: households, hotel/motel rooms, 13 employment-related variables, and average income by census block.
- Other variables: 15 variables describing the distance from a CCS to coastlines, metropolitan areas, state borders, and highway interchanges; accessibility to interstates; network density; and geographical locations of CCSs.

The dependent variable consisted of monthly adjustment factors of 73 CCSs located in two rural districts. The authors retrieved most of the demographic and socioeconomic data from the 2000 census and the employment data from a proprietary database purchased by the Florida DOT. Applying two different buffer methods, the analysts created circular areas around each CCS and aggregated the variables at the buffer level. According to the first buffer method, the radius of each buffer varied according to the functional class of the roadway on which a CCS was located. The buffer size of CCSs located on rural minor collectors was 1 mile. The second buffer method defined the impact area based on the average travel time (approximately 30 minutes) to workplace. Zhao et al. applied stepwise regression to select the most significant variables.

Considering different buffer sizes resulting from the two methods mentioned above, the authors developed six sets of models. Each set includes 12 models, one model for each month of the year. The adjusted R^2 values of the models ranged between 0.0941 (model for January) and 0.5581 (model for December). Among all variables examined, the most significant ones were roadway functional class, percentage of seasonal households, agricultural employment, and truck factor.

Based on the results, the buffer methods did not adequately capture the effects of land use and other activities that typically cause traffic fluctuations over time. For instance, the impact of through traffic, which was not generated or destined in the examined buffer areas, was not well captured in the variables of the analysis. However, the through traffic may have larger effects on higher rural functional classes such as rural interstates and principal arterials than rural minor collectors.

2006 – Improving the Prediction of Annual Average Daily Traffic for Nonfreeway Facilities

Eom at al. developed a spatial regression model to predict AADT for non-freeway facilities in Wake County in North Carolina (*28*). The study employed the kriging geostatistical approach to model spatial trends and correlations of different types of variables that included:

- AADT values derived from 200 randomly selected traffic counts.
- Roadway characteristics: area type, number of lanes, posted speed limit, highway functional classification, signal density, and existence of median left turn.
- Socioeconomic variables: total population (block group level), number of households, household size, number of households with a child under 6 years old, medium income, workers living within block boundary, workers living outside of block boundary.

After exploring different Box-Cox transformations of the AADT, Eom et al. selected the $(AADT)^{0.25}$ as the response variable. By applying stepwise regression, the authors included in the final model the following independent variables: longitude, latitude, speed, median income, number of lanes, urban area, suburban area, urban major arterial, urban minor arterial, rural arterial, local road, and collector. The coefficients of the variables were estimated using ordinary least squares regression and two estimation methods that account for spatial dependency, weighted least squares and restricted maximum likelihood.

In the case of urban arterials and local roads, weighted least squares and restricted maximum likelihood provided more accurate predictions than ordinary least squares regression. However, weighted least squares and restricted maximum likelihood showed no significant improvement in the case of collector roads. The main reason was because stations on collectors were sparsely located and only a few neighbors of each station were within the range and their contribution to the prediction was relatively weak.

2015 – Developing a Method for Estimating AADT on All Louisiana Roads

This study estimated AADT focusing on non-state rural roadways of eight parishes in Louisiana (*29*). The authors developed a support vector regression (SVR) model using demographic and topological data. SVR is a machine learning or pattern recognition method that executes the structural risk minimization inductive principle to get good generalization on a limited number of patterns. The main concept of SVR is to transform data into a high-dimensional feature space *F* via a nonlinear mapping and to perform linear regression in this space. For the model development, the analysts used an open source library in R, called "e1071," by considering four variables: population; number of jobs; distance between a traffic count and the closest intersection; and distance between a traffic count and the closest access point of a major state highway.

The main data sources were the American FactFinder, the longitudinal employer-household dynamics, and the topologically integrated geographic encoding and referencing database. The analysts used an ArcGIS tool (Closest Facility Analysis) to identify the distance between a count station and the closest intersection or major highway. To create the final dataset, the analysts merged four datasets (census, traffic, employment, and shortest path data) at the census block level for each parish. Because some parishes did not have direct access to interstates, the authors developed parish-specific and integrated parish models. The integrated parish models accounted for all parishes with and without direct interstate access.

The results showed that the SVR models tend to underestimate high AADT values and somewhat overestimate the low AADT values. The percent of the AADT estimates that were within ±100 vehicles from the observed traffic volumes ranged between 64 percent and 82 percent among the parishes examined. The corresponding percent of the estimates that had a difference of ±200 vehicles or less from the observed values, spanned from 78 percent to 91 percent. The study also presented a sensitivity analysis, which indicated that a parish-specific model performed better than an aggregated single model.

Despite the promising results presented in this study, SVR and other non-parametric methods suffer from certain limitations. For example, one of the main shortcoming of SVR and all machine learning methods is the lack of probability estimation of the results (*29*). Further, the applicability of such models can be limited to groups of roads for which socioeconomic, demographic or other types of data are readily available. For example, Sun and Das showed that if major highways are not contained within a group, the aggregated model may not produce estimates at the anticipated level of accuracy. Another limitation is the difficulty in understanding the model structure and the relative importance of the variables included in the model. Also, developing a model requires knowledge of computer programming and understanding of statistics and machine learning theory.

2016 – Estimation of AADT on Local Roads in Kentucky

Staats developed six models to estimate aggregated traffic volumes of local roads in Kentucky (*30*). The author initially developed an ordinary linear regression (OLR) model that included five predictors: direct access to an expressway (*DIRECTAC*), employment buffer (*EMPBUFFER*), population buffer (*POPBUFFER*), distance to population center (*DISPOPCNTR*), and accessibility to regional employment centers (*REACCESS*). The model produced one aggregated AADT estimate for each TAZ included in the study area. The form of the model was:

*AADT = 357.23*DIRECTAC + 0.02*REACCESS – 0.63*POPBUFFER – 0.05*EMPBUFFER + 0.09*DISPOPCNTR*

Due to high prediction errors (MAPE=125 percent and average absolute error=417), Staats modified the first model by incorporating property valuation administrator (PVA) data. The PVA data described different types of land uses such as residential, agricultural, commercial, and educational. The final form of the second enhanced OLR model was:

*AADT = 4622.68 -0.01*REACCESS – 0.75*DISPOPCNTR + 0.35*POPBUFFER – 0.92*EMPBUFFER – 0.56*RESIDENTIAL – 0.47*AGRICULTURAL + 17.92*COMMERCIAL – 3.81*EDUCATIONAL*

The second model yielded a small decrease in prediction errors (MAPE=82 percent and average absolute error=402). To further improve the AADT accuracy, Staats developed a third model, called "Rooftop". This model included four variables: connectivity, and number of small, medium, and large properties. By visually examining an aerial map in ArcMAP, the author identified the number of properties adjacent to local roadways and classified them as small, medium and large based on their size. The regression equation for the third model was:

*AADT = 113.8*CONNECTIVITY + 2.1*SMALL + 49.3*MEDIUM + 138.8*LARGE*

The Rooftop model produced less accurate estimates compared to the second model (MAPE=93 percent and average absolute error=332) and also required significant development time. As a result, Staats modified the PVA model by using the number of residential and commercial properties that he extracted from the "911" database; so, the fourth model was called "911" model was:

*AADT = 43.5*RESIDENTIAL+ 16.4*COMMERCIAL*

Overall, the "911" model provided the lowest MAPE (61percent) and the second lowest absolute error value (352). However, the 911 data were not readily available at the statewide level and this limited the applicability of the model to other parts of the state. The fifth OLR model, called "AVIS-HERE", included two variables: probe counts and residential vehicle registrations. Both training and validation datasets contained data from rural two-lane local roads with known AADT:

*AADT = 168.32 + 2.06*PROBE + 1.04*RESIDENTIAL*

The main limitation of this model is that the lowest AADT value is equal to the constant (168), limiting the predictive power of the model in cases where AADT was very small (e.g., <50). The average absolute error (153) of the model was the lowest among all five models, but the MAPE (192 percent) was the highest. Using AVIS-HERE and roadway curvature data, Staats developed a generalized linear model with a Poisson distribution and a log link function of the following form:

$$Y = e^{a+\beta_1 PROBE_1+\beta_2 CURVE_2 \ldots + \beta_3 RESIDENTIAL_3}$$

The variable CURVE expressed the ratio of the actual length of a segment to the straight length between the start and end points of the segment. The author divided the state into three regions to account for spatial and socioeconomic variations. Within each region, Staats built one model for rural roads and a second model for urban roads. In general, the non-linear models performed better than the other five models, producing MAPEs between 97 percent and 102 percent (*30*). One limitation of these models is that they did not take into account residential and commercial properties associated with vehicles registered outside of Kentucky. Also, the study did not report the degree of correlation or the coefficient of determination (R^2) between observed versus predicted AADT estimates.

Staats also performed a sensitivity analysis to evaluate how the model's AADT estimation errors could potentially impact the number of predicted crashes that are used in safety analysis. The sensitivity analysis included the following steps:

a) Identified all intersections in Kentucky using GIS tools.

b) Classified intersections based on their characteristics (rural or urban roads, number of approaches, unsignalized or signalized, and number of lanes in each direction). The analysis considered intersections where State roads intersect local roads.

c) Estimated predicted average crash frequency using safety performance functions (SPFs) from the HSM for different types of intersections such as three-leg and four-leg stop-sign controlled intersections. The AADT of the major road of an intersection was obtained from Kentucky's *Highway Information System*, while the AADT of the local intersecting road was estimated from the AVIS-HERE model.

d) Estimated the expected number of crashes based on the following Empirical Bayes (EB) formula:

   $$Expected\ Crashes\ in\ X\ years = Overdispersion\ parameter * N * CMF * X + (1 - Overdispersion\ parameter) * Previous\ crashes$$

   Where:

   Overdispersion parameter = obtained from the HSM and calibrated for each SPF

   N = the number of crashes predicted by the SPF

CMF = crash modification factor (from HSM or CMF Clearinghouse)

X = the number of years

Previous crashes = the number of crashes at the intersection in the past X years

e) Adjusted AADTs using the following equation:

$Adjusted\ AADT = Estimated\ AADT/(1 + Percent\ Error)$

Where:

Estimated AADT = the AADT generated by the model

Percent Error = maximum percent error (797 percent), minimum percent error (-94 percent), average positive error (134 percent), or average negative error (-38 percent).

f) Revised SPF values using the adjusted AADTs. The EB method incorporated the updated SPF values and used crash data over a 10-year period assuming a CMF of 0.15. Staats estimated that the average crash cost was $54,051.

g) Estimated maximum countermeasure construction cost for each prediction error and intersection assuming that the benefit-to-cost ratio was equal to five.

h) Calculated percent errors for maximum countermeasure costs between the original AADT, the estimated AADT, and the adjusted AADTs. This range of errors captured the variation in the number of predicted crashes due to different AADT estimation errors.

The average AADT estimation errors ranged between 134 percent (overestimated AADTs) to 38 percent (underestimated AADTs). After accounting for the two errors and adjusting the AADT values, the authors found that the derived crash prediction errors were much smaller: 28 percent for overestimated AADTs and 22 percent for underestimated AADTs. The maximum AADT estimation error was 797 percent and the lowest was 94 percent. The derived crash prediction errors were 54 percent and 253 percent respectively. In general, the study showed that the AADT errors had a limited impact on the final crash predictions. This happened because the local road AADT only influences the number of crashes predicted by SPFs. However, intersection crash predictions must take into account both SPFs and historical crash rates and the latter have a higher weight in the equation (*30*).

2016 – Methods to Improve Traffic Flow and Noise Exposure Estimation on Minor Roads

Morley presented a method to predict AADT at the national scale for minor roads. This study used a routing algorithm within a GIS to rank roads by importance based on simulated journeys through the road network (*31*). The researchers used a set of known minor road AADT values as the training set. The outcomes of the routing importance was used to predict AADT using a regression model along with the road class, urban or rural location, and AADT on the nearest major road. The generalized linear model was:

$$AADT = \log(route\ importance) + Ordance\ Survey\ Meridian\ (OSM)\ Road\ Type \\ + \log(AADT\ on\ nearest\ major\ road) + Urban\ or\ Rural$$

The routing method was found to moderately under-predict the observed AADT for both major and minor roads. However, in the context of this study (to estimate road noise exposure) the routing model represented a substantial improvement over the fixed and classified AADT assignment methods.

2.4 TRAVEL DEMAND MODELS

Transportation agencies have widely used TDMs for roughly six decades. The earlier TDM models were mainly used to evaluate major infrastructure investments and prepare long range, regional transportation plans. However, recent technological advances significantly improved travel demand modeling, enabling MPOs and other local agencies to model small urban networks. Unlike older TDMs, modern models have diverse applications such as:

- Project-level studies that use hourly traffic data in geometric design.
- Subarea traffic circulation studies that require peak hour turning movements.
- Analysis to determine the feasibility of transportation investment strategies.
- Evaluation of land use patterns and their impact on transportation systems.
- Regional and localized air quality analysis.
- Analysis and evaluation of travel demand and congestion management systems.
- Analysis of road pricing options.

The traditional travel demand modeling process consists of the following four steps:

- Trip generation – estimating the productions and attractions for each trip purpose.

- Trip distribution – pairing trip productions and attractions by applying the gravity model or one of its variations.
- Mode choice – converting productions-attraction round trips into origins and destinations by the various modes available. This step is not critical for rural models because they do not have alternative modes of travel such as buses and trains.
- Trip assignment – assigning the vehicle trips developed in the mode choice step to the road network. The most common assignment method is the user equilibrium method.

The main data inputs in travel demand modeling are transportation network data, land use data, and socioeconomic data. Travel behavior data such as trip rate tables, vehicle occupancy factors, and time of day factors are often used as well. Socioeconomic data are typically aggregated in transportation analysis zones and used in computing trip generation in the first step. TAZs serve as the basic geographical unit of land use. The network data enable estimation of travel times for the trip distribution and trip assignment steps. Overall, only a few studies have developed TDMs to estimate traffic volumes on NFAS roads. Most studies have focused on higher functional classes.

2009 – GIS-Based Travel Demand Modeling for Estimating Traffic on Low-Class Roads

Zhong and Hanson developed TDMs for estimating traffic volumes for NFAS roads in two regions (York County and Beresford Census Consolidated Subdivision) in the province of New Brunswick, Canada (*32*). The analysts built the TDM in TransCAD following the traditional steps of travel demand modeling by omitting the mode choice step. They used three GIS roadway networks: the national network, the provincial network, and the network covered by the provincial traffic counting program. Dissemination areas were the smallest census units used as TAZs for which socioeconomic data (from the 2000 census) were readily available. Zhong and Hanson used the Quick Response Method for trip generation, trip attraction, and trip balancing.

For trip production, they considered three trip purposes: home-based work, home-based non-work, and non-home-based. The total number of households in the zone and the average income per household were the two data inputs. The authors used a regression equation to estimate the number of person trips attracted to a zone based on the number of dwelling units in a zone and the zone's work activity. The data required for this step included the total number of households, retail employment, and non-retail employment in the zone.

The study used the gravity model to distribute trips among the zones. The analysts created an impedance matrix based on the distance between zone centroids. The result of this step was an OD matrix for all zone-to-zone trips for each type of trip. The assignment of trips to roads and generation of traffic volumes involved:

- Cross-referencing the centroid nodes to the nodes they represent in the road network.
- Indexing the OD matrices with the road network nodes (rather than centroids nodes).
- Replacing links in the road network that had a length of zero with a non-zero value.

The study applied the STOCH method, according to which trips were distributed between OD pairs based on the shortest path. Also a portion of the trips were assigned to other reasonable routes based on probabilities calculated by a logit route choice model. The output of the model was total traffic volumes on each link for each trip type. The sum of traffic volumes for all trip types represented the AADT estimate of the link.

In the case of the York County, the results showed that AADT estimates for arterial roads had the lowest errors (9 percent), followed by collectors (45 percent), and local roads (160 percent). Overall, data aggregated at the dissemination area level are too coarse for applications involving local roads, which typically attract most trips generated from abutting land uses. The study reported an overestimation trend of AADT values across many sites, meaning that traffic was not effectively distributed to collectors and local roads. After applying linear regression to calibrate the model, the overall average error was reduced to 36 percent for collector roads and was lower than 40 percent for local roads. In the case of Beresford Census Consolidated Subdivision, the estimation errors for local roads were limited to less than 40 percent, and the average error was reduced to 25 percent.

2012 – A Travel Demand Model for Rural Areas

Berger developed a TDM for Gallatin County in Montana to estimate total daily vehicle miles traveled (*33*). For the trip generation step, Berger used parameter estimates from the Institute of Transportation Engineers (ITE) *Trip Generation* report (*34*) and the NCHRP Report 365 (*35*). The trip distribution step was based on the gravity model and friction factors from the NCHRP Report 365. The mode choice step considered only personal vehicle travel and used the vehicle occupancy rates recommended in NCHRP 365. In the trip assignment step, Berger initially used the all or nothing traffic approach to assign trips to the shortest trip route based on free flow speeds. The results from the all or nothing traffic assignment method were not satisfactory and Berger replaced it with the user equilibrium method.

The study reported a RMSE of 111.3 percent. After applying a ratio of total estimated ADT to actual ADT to each of the trip rate equations (for the trip generation step), the calibrated TDM yielded significantly lower RMSE (49.0 percent).

2013 – Estimating AADT for Local Roads for Highway Safety Analysis

Wang et al. developed a tax parcel-level TDM to estimate AADTs for local roads in Broward County, Florida (*36*). The main difference between a traditional TDM and the parcel-level model was that the latter simulated choices that travelers made in response to a given local street system, as opposed to an entire trip from the origin to the destination. The parcel-level TDM involved the following steps:

- Network modeling – Defined the boundaries of the study area and processed and linked the roadway network with parcel and traffic count data.
- Parcel-level trip generation – Estimated the number of vehicle trips generated by each parcel in the study area based on the land use type of each parcel and its corresponding ITE trip generation rate (*34*).
- Parcel-level trip distribution – Determined the number and the destination of the trips generated by each parcel based on traffic count data and the shortest travel time between a parcel and a traffic count site.
- Parcel-level trip assignment – Predicted the routes travelers took to reach the traffic count sites on major roads, resulting in the estimated AADTs of local roads.

The mode choice step was omitted because of the insignificant existence of transit trips or those of other travel modes on local roads. The authors developed the model using ArcGIS from ESRI and Cube from Citilabs. The study compared the AADT estimates of the parcel-level TDM against those collected from short-duration counts. The MAPE ranged from 39 percent to 66 percent. The average MAPE was 52 percent. The authors also developed a set of regression models for comparison purposes. The MAPE of the regression models were 211 percent.

One advantage of using tax parcel data was that the data are updated at least annually allowing frequent updates to AADT estimates in response to land use changes. One disadvantage of the model was the large number of parcels that had to be preprocessed. This can be a challenge for applications involving large areas. Another drawback of the model was that it required sufficient traffic count data to cover the entire study area. Further, the counts had to be evenly spaced because uneven coverage of traffic count sites may result in inaccurate distribution of trips.

2015 – Wyoming Low-Volume Roads Traffic Volume Estimation

Apronti et al. developed a TDM for four counties in Wyoming (*24*). The authors used ArcGIS to compile network data and socioeconomic or land use data from 2013. The roadway attributes included posted speed limit, number of lanes, functional class, and existence of one or two directions of travel (one-way vs. two-way roads). The analysts also developed TAZs from census block shapefiles and aggregated the socioeconomic data at the TAZ level. After processing the data, the authors inserted them into Cube 6 and developed the model in the four traditional TDM steps:

- Trip generation – Using TAZ attributes and trip rates obtained from the NCHRP Report 365 (*35*), the analysts calculated trip production and attraction rates and trip purpose percentages for 832 analysis zones for three trip purposes: home-base work, home-base other, and non-home base trips.

- Trip distribution – Use of a gravity model to pair productions and attractions that were estimated from the previous step. The gravity model estimated that the relative number of trips made between two zones are directly proportional to the number of productions and attractions in each zone and inversely proportional to a function of the spatial separation or travel time between the two zones. The output was a matrix table with assigned production and attraction trips between all pairs of zones in the study area.

- Mode choice – Converted the production-attraction output of the previous step to an origin-destination format. The only mode of travel considered in this study was personal vehicle travel because other modes (e.g., transit) were not significantly represented in the study area.

- Trip assignment – Distributing the origin-destination trips for all pairs of zones among the network links connecting the zones. The user equilibrium method assigned traffic to the network. This method is based on the principle that as congestion delay increases, road users seek alternative routes until all vehicles traveling between two zones use the shortest routes available. The output of this step included a shapefile of the road network with the assigned ADT for each roadway segment.

After calibrating the model, the authors compared the estimated ADT values against actual traffic volume data collected on 100 roads in the study area from 2012 to 2014. The analysts estimated a RMSE of 50.3 percent and reported a strong linear relationship (R^2=74 percent) between the actual and the estimated values. The analysts did not perform additional calibrations such as modification of the travel time distribution factors (friction factors), vehicle occupancy and time of day distribution factors. Such calibrations can further improve the predictive power of the model.

Another limitation of this model is the inability to estimate ADT values for roadway segments that fall entirely within a zone. Such roads were excluded from the analysis because they did not provide a link between two zones necessary for the trip distribution process. This problem also affected the accuracy of the estimates, because the trips generated in a zone were assigned to fewer links than they should in reality.

The authors recommended the TDM for implementation (as opposed to the two regression models developed in the same study) due to its high prediction accuracy and its ability to explain traffic volume changes using purely independent variables. The main disadvantage of the TDM is its high development and implementation cost.

2.5 RECREATIONAL TRAFFIC

As part of Task 2, we extended the review of the literature to studies that examined ways of determining recreational traffic. This focused review identified potential methods of capturing recreational patterns on NFAS roads and assessed the need for further examining these methods in Task 4 and potentially including them in the guide.

1969 – Recreational Trip Generation: A Cross Section Analysis of Weekend Pleasure Trips to the Lake District National Park

Mansfield analyzed the demand for day and half-day recreational trips to a holiday area. The study developed a model describing the generation of trips to alternative recreation facilities and holiday resorts (*37*). The aim was to make a cross-section analysis of a single year's traffic data to determine how much of the observed variations in trip demand during a single year can be attributed to decision making.

For modeling purposes, the study assumed that unobserved variables, $C_2, \dots C_N, t_2 \dots t_N, r_2 \dots r_N$ would remain constant and would not affect the analysis. These variables were expressed as a function of a single variable, R_{kl} that measured the overall competitiveness of the Lake District when compared with alternative recreation zones available to the population of the kth zone:

$$R_{kl} = f(C_2, \dots C_N, t_2 \dots t_N, r_2 \dots r_N)$$

The form of the general model was:

$$\gamma_{kl} = f(C_{kl}, t_{kl}, W_k, R_{kl}, P_k) \qquad [K = 1, 2, \dots \dots \dots 15]$$

This study was not statistically rigorous due to the small sample size of the original data. The results in this study are not precise estimates of future predictions and the results can be interpreted as general indications of the responsiveness of demand to changes in the variables considered.

1980 – Progress and Problems in the Development of Recreational Trip Generation and Trip Distribution Models

Ewing discussed problems with recreational spatial interaction models, which estimate the expected number of recreational trips based on different factors such as population, destination attractiveness, and costs (*38*). The author raised several issues with the models including difficulties in estimating travel costs, not being able to compare competing recreational activities, and unreliable estimates of facility attractiveness. The author noted the need to perform sensitivity analysis because parameter estimates are not exact.

2003 – Traffic Effects of Fairs and Festivals on Low-Volume Roads

Eck and Montag conducted a study to evaluate the effects of recreational traffic like fairs and festivals on low volumes roads in West Virginia (*39*). West Virginia Tourism Calendar of Events and Festivals indicates that slightly more than 1,000 events are scheduled each year. The research team selected four sites where the impact of recreational travel was significant.

This study calculated trip generation rates and compared the results with regular periods. The authors divided the recreational trips into five major groups: food and drink, arts and crafts, cultural heritage, historical, and performing art. The results showed that the average rates for cultural heritage and historical events were much lower than other events, whereas the highest rate occurred at food and drink related events. The results also showed that weather played a significant role (a decrease of 15 percent to 21 percent for severe weather conditions) in recreational travel.

2010 – Traffic Monitoring in Recreational Areas

The Office of Federal Lands Highway of FHWA conducted a study to assess recreational traffic data collection by reviewing literature, conducting surveys, and holding a workshop (*40*). The authors categorized their findings as they relate to national guidance for traffic monitoring in recreational areas, vehicle classification, practices in literature, and observed practices.

The authors noted that recreational traffic varies greatly and this is recognized in national guidance documents. However, these documents provide little guidance to address recreational areas. Recreational traffic consists of a higher percentage of recreational vehicles, buses, and vehicles with trailers than typical traffic. As a result, the FHWA vehicle classification scheme may not adequately cover recreational traffic. The authors found wide variability in traffic monitoring practices and effectiveness in the literature and in practice. The study noted that any changes to existing recreational traffic monitoring practices cannot be a one size fits all approach.

2013 – Traffic Monitoring Guide

The TMG (2013) does not recommend any single method for determining recreational traffic (*1*). A general recommendation is that analysts have to rely to a large extent on their judgment and knowledge of the travel characteristics of the state to take the following actions:

a) Determine the routes or general areas where a given recreational pattern is clearly identifiable. Toward this direction, the analysts need to examine continuous count data and subjectively identify and select the recreational roads or patterns.

b) Establish a set of locations where short-duration counts need to be conducted.

c) Allocate factors to short-duration counts by applying engineering judgement (*1*).

The TMG acknowledges that statistical procedures are not directly applicable in all cases. For example, the TMG states that:

A roadway is likely a recreational road when the difference between the ratio of the highest hourly volume to AADT and the ratio of the thirtieth highest hourly volume to AADT is greater than one. No single method exists for determining recreational patterns. The best way to determine trip purpose absolutely is to conduct intercept surveys.

Regarding the variability of recreational patterns, the TMG reports that the coefficient of variation of typical monthly patterns for rural areas ranges between 10 and 25 percent. Values higher than 25 percent are indicative of highly variable travel patterns, which may be due to recreational traffic or other factors that are not related to recreational travel (*1*). Overall, the TMG focuses on major commuting and through-traffic routes because:

a) Recreational traffic has greater variability and specific procedures for monitoring it are less formulaic on a national basis; and

b) Recreational traffic comprises a small percentage of the total VMT that must be monitored by resource-constrained State agencies.

2.6 SUMMARY OF FINDINGS

This chapter reviewed previous studies on data collection and AADT estimation for NFAS roads, namely rural minor collectors and local roads. For clarity, we divided and presented the AADT estimation methods into three groups: traffic count based methods, non-traffic count based methods, and travel demand models. The review also involved methods of determining recreational traffic on NFAS roads with the aim to assess the need for further examining these methods in Task 4 and potentially including them in the guide. The most important findings from Task 2 are summarized below.

2.6.1 Traffic Count Based Systems

- Most of the literature available on traffic volume estimation are concentrated on urban roads, highways and freeways or expressways. Most of the studies that dealt with AADT estimation for NFAS roads are county or state based and may not be transferable to other study areas without requiring appropriate calibration (*1*).

- While the majority of the research studies have developed various statistical and machine learning methods, few agencies have adapted advanced estimation methods, and most agencies still use variations of the traditional approach, including network stratification and sampling (*10*).

- The method of using a sample of short-duration counts combined with estimates for uncounted roads is the most prevalent nationally and internationally (*10*). Many states use a single AADT value for all uncounted segments within a group of roads that have common characteristics (e.g., functional class, county, population size, land use). The AADT value is typically derived from the counted segments. While this approach is simple and easy to apply, it is not as accurate as using non-traditional methods, and provides questionable usefulness for safety analysis (*10*).

- Literature shows that a more accurate method for uncounted segments is to develop regression models using existing count data, socioeconomic data, network connectivity, and other information to predict an AADT value for each uncounted segment (*11*). Neural networks can also improve the accuracy of AADT estimates. However, it is difficult to interpret the structure of the models and fully understand the role and contribution of all variables. Not surprisingly, the effectiveness of non-traditional models increases as more short-duration counts are included in the training dataset (*18*).

- Research results on data collection and AADT estimation for NFAS roads vary with respect to sampling techniques, data inputs, level of development effort and complexity, runtime, use of GIS technologies, study area, anticipated accuracy, and software, algorithms and statistics used. Therefore, drawing conclusions for the effectiveness and appropriateness of one method is not a straightforward task, but relies on several factors.

- Sampling based on a volume group stratification scheme provides an efficient method for minimizing sample size and maximizing accuracy because the variation within each of the volume groups is well understood. However, many states do not have enough counts on NFAS roads to designate volume groups. In these cases, other non-volume based classification schemes are needed to serve as a proxy for volume groups until several years of traffic data have been collected (*12*).

- Roads belonging to the same functional class may not necessarily have similar traffic patterns and volumes (*18*). The effectiveness of the sampling process depends on the amount of traffic volume variability within each group of roads. For example, Wang et al. found that ADTs on local roads in Broward County (Florida) ranged from 2,000 to 25,000 (*36*). Overall, methods that rely only on functional classification tend to produce high estimation errors.

- While county boundaries are readily available in various forms and make the stratification of the network a relatively easy task for agencies, there is high variability in ADT among the NFAS roads within a county. Therefore, agencies should stratify their network using other variables (e.g., population, land use) that may be better surrogates for AADT.

- In many cases, low-volume local roads have limited or no coverage of continuous count sites, thus determining accurate seasonal factors can be a challenging task. The absence or inadequate coverage of permanent counts often results in applying adjustment factors from higher functional classes within the vicinity of a road. This tends to introduce high AADT estimation errors because NFAS roads possess rather different characteristics and travel patterns compared to high-class roads (*25*). For this reason, some states use ADT values computed from short-duration counts without applying adjustment factors. This topic needs further examination.

- Mohamad et al. found that traffic volumes on local roads do not substantially vary within a day or a week. This finding suggests that factoring short-duration counts on these roads does not provide significant benefits (*19*, *24*). Additional research is needed on this topic.

- Short term traffic counts on NFAS roads are typically carried out for a period of 48 hours. However, some states use less expensive 24-hour counts with the goal to cover larger portion of their local network with counts at the expense of having less accurate estimates. A benefit-cost analysis is needed to quantify the effectiveness and the costs of 24-hour counts versus 48-hour counts on NFAS roads.

- High levels of AADT accuracy for low-volume roads may not be necessary in many applications. For instance, a 30 percent error in AADT estimates would rarely affect major decisions pertaining to the design of low-volume roads. On the other hand, large errors may be of great concern to those who deal with allocation of funds. For example, underestimating an AADT value by 30 percent could make a road ineligible for financing if the estimated AADT falls short of a predefined minimum threshold (*14*).

- The Highway Policy Steven Jessberger Battelle method of calculating AADT provided less biased and more accurate results than other methods such as the AASHTO method (*13*). States should consider the HPSVB method for calculating AADT from continuous count data.

2.6.2 Non-Traffic Count Based Systems

- Non-traffic count based models can produce AADT estimates either for a) individual roadway segments or b) a group of roads that have common characteristics. In general, the models that generate disaggregated estimates at the segment level showed superior performance compared to the models that produce aggregated estimates. However, the first type of models require disaggregated and detailed data, which are often difficult to collect or not readily available within transportation agencies. The second type of models result in less accurate AADT estimates because they fail to capture ADT variability at the segment level. For example, it is difficult to disaggregate demographic information at the segment level, especially when the segments are short. As a result, these models are suitable for groups of uncounted segments and applications that do not require high levels of AADT accuracy.

- The review revealed gap in knowledge related to how AADT accuracy can impact the results of safety analysis. A study by Staats (2015) that dealt with this topic, found that the AADT errors, produced by regression models, had a limited impact on the number of predicted crashes at intersections in Kentucky (*30*). However, the study did not analyze potential impact of estimation errors on roadway segments. Additional research is needed on this topic.

- While roadway characteristics such as number of lanes can have a significant role in AADT prediction models developed for higher functional classification roads, they do not adequately explain AADT variations on NFAS roads. The main reason is that the roadway characteristics of rural minor collectors and local roads are relatively homogenous and less variable compared to those of higher classifications roads. The AADT on NFAS roads is typically a function of other factors such as socioeconomic characteristics, population density, and types of land use (*25*).

- The traffic on local roads may not be highly correlated with those on a freeway, even if the roadways are closely located. Traffic patterns on local roads largely depend on characteristics of the local communities, whereas those on exclusive facilities (e.g., freeways) that have limited access control are more related to regional through traffic (*28*).

- One limitation of some regression models is that they can have a rather large negative intercept. While the intercepts do not have a physical meaning, large negative values tend to underestimate the AADT. This can be the result of not including all relevant predictors that may have a significant impact on AADT (*25*).

- Additional research is needed to determine whether the independent variables used in various studies are applicable and important to other states. Because of differences in local land use patterns and economies, it is possible that some areas have a different set of variables that explain the patterns on traffic variations (*27*).

2.6.3 Travel Demand Models

- In general, studies that developed travel demand models reported improvements in AADT accuracy over other non-traditional methods. For example, Zhong et al. found that the average estimation errors for NFAS roads were limited to less than 40 percent (*32*). However, there are several limitations associated with the development and implementation of TDMs. Some of the most important limitations are listed below.

- Travel demand models are often limited in the number of roadway links that can be included in the analysis. Especially for large areas, the number of links that can be built into a network is limited, creating bias toward major roads and omitting most minor roads. This results in incomplete models that do not contain all links of the network. This can lead to underestimation of traffic volumes on minor roads (phase 1 report). Regression can be used to calibrate the estimates in order to remove this bias and increase the overall accuracy, particularly for NFAS roads (*32*).

- Low-volume roads by nature are susceptible to higher prediction percentage errors compared to higher functional classes (*32*). For example, Zhong et al. found that TDM predictions for high volume roads resulted in average errors of nine percent, whereas those for low volume collectors and local roads were 44 percent and 174 percent, respectively.

- One limitation of some travel demand models is that the spatial resolution of transportation analysis zones may limit the ability to estimate traffic volumes for certain roads. The TAZ delineation typically requires combining smaller geographical units (e.g., census blocks) to create TAZs. However, some short segments may lie wholly within a zone without providing a link between two zones. These roads are excluded from the trip distribution step since trips cannot be assigned to them. This problem is prevalent in low population areas that can have a significant number of short segments. This limitation can also affect the accuracy of AADT estimates because trips generated in a zone are eventually assigned to fewer links compared to what happens in reality.

- The cost to develop and implement travel demand models is relatively higher compared to non-traditional models. The implementation of TDMs is also more difficult compared to regression models. Creating TAZs for large rural areas can be time consuming and the model implementation requires some knowledge of the theory behind travel demand modeling and model calibration.

- Trip rate equations must be separately developed for special types of land use such as oil and gas wells, mines and plants, because they tend to generate and attract significant levels of traffic.

- The accuracy of AADT predictions depends on how well a TDM is designed and calibrated, and how much of the roadway network it covers (*10*). The addition of available traffic counts as capacity constraints improves the model accuracy (*32*).

- The accuracy and age of demographic and other input data must be considered when evaluating the quality of model output. The accuracy of the model inputs affects the accuracy of AADT estimates (*10*).

- Most studies omitted the mode choice step to account for the lack of significant transit and other modes of travel on NFAS roads, where the main means of transport is personal vehicles.

2.6.4 Recreational Traffic

- The majority of previous works that dealt with the determination of recreational roads examined all functional classes without concentrating on NFAS roads.

- Ewing noted several issues with recreational models developed by various studies including, no maximum number of trips, difficulties in estimating travel costs, not being able to compare competing recreational activities, and unreliable estimates of facility attractiveness (*38*).

- The Office of Federal Lands Highway of FHWA conducted a study to assess recreational traffic data collection by reviewing literature, conducting surveys, and holding a workshop. The authors found wide variability in traffic monitoring practices and effectiveness in the literature and in practice. The study noted that any changes to existing recreational traffic monitoring practices cannot be a one size fits all approach and must be flexible to encompass the different needs and goals of the various agencies responsible for monitoring recreational traffic (*40*).

- Distinct recreational patterns cannot be defined based simply on functional class or area boundaries (*1*). Recreational patterns may be obvious for roads at some locations but non-existent for adjacent roads. The boundaries of recreational groups need to be designated based on subjective knowledge. The existence of different seasonal patterns makes the determination of recreational traffic more difficult (*1*).

- The TMG does not recommend any single method for determining recreational traffic. A general recommendation is that analysts have to rely to a large extent on their judgment and knowledge of the travel characteristics of the state network to determine recreational routes (*1*).

- Recreational patterns are usually identifiable by examining continuous count data. The limited availability of continuous and short-duration count data on NFAS roads makes this task more challenging (*1*).

CHAPTER 3—PROPOSED SCOPE OF GUIDE

3.1 INTRODUCTION

This chapter describes the proposed topic, target audience, anticipated use, and main questions that the guide needs to address. The last section of the chapter discusses key elements that the guide should include. The key elements include a general framework, data collection methods, AADT estimation methods, AADT validation, data collection costs, research results, case studies, and noteworthy practices. The selection of these elements relies upon two objectives that are common for most target agencies: 1) to meet relevant HSIP Final Rule requirements (*4*); and 2) to be able to collect data in a cost-effective manner by reducing estimation errors and reaching AADT accuracy levels that are needed to reliably perform safety analysis.

3.2 TOPIC, AUDIENCE, USE, AND QUESTIONS

This section discusses the proposed topic of the guide, along with the target audience, its anticipated use, and the questions that this project should address.

3.2.1 Topic

The guide will focus on data collection and AADT computation/estimation methods for non-Federal-aid system roads that include:

- Functional Class 6 – rural minor collectors (6R), and
- Functional Class 7 – urban local roads (7U) and rural local roads (7R).

3.2.2 Target Audience

The target audience of the guide include safety and traffic monitoring officials at State, tribal, MPOs, and local and Federal land management agencies.

3.2.3 Anticipated Use

The guide will help target agencies to develop work plans for collecting data on NFAS roads and estimating/computing AADT values for use in data-driven safety analysis. Data collection and AADT estimation practices in the country vary significantly from one agency to another. For example, some agencies cover a significant portion of their local roads and rural minor collectors with short-duration counts and are able to estimate segment-specific AADT values for the majority of their local network. Other states do not collect traffic volume data on their non-Federal-aid system and produce less accurate AADT estimates at higher levels of aggregation such as the functional class and county levels. To account for the variability and different maturity levels of state practices on this topic, the guide should be suitable for use by both:

a) Agencies that currently do not collect count data or do not estimate AADT for NFAS roads.

b) Agencies that would like to expand their existing data collection practices and enhance the accuracy of their AADT estimates for NFAS roads.

This distinction will maximize the usefulness and anticipated benefits of the guide by the target agencies. The traffic volumes generated using the method(s) included in the guide will be mainly used in two types of safety analysis: network screening using SPFs, and systemic analysis.

Network screening is the first of six activities included in the roadway safety management process of the Highway Safety Manual (*41*). The network screening process involves reviewing a transportation network, identifying, and ranking sites from most likely to least likely to realize a reduction in crash frequency by implementing a countermeasure. One of the main steps of network screening is the estimation of one or multiple performance measures.

The HSM includes 13 performance measures, of which 8 measures use traffic volume as input. One of the most promising measures is the excess predicted average crash frequency using SPFs. According to this measure, the site's observed average crash frequency is compared to a predicted average crash frequency from an SPF. The difference between the observed and predicted crash frequencies is the excess predicted crash frequency using SPFs.

In addition to estimating performance measures for network screening, traffic volumes can be used in systemic safety analysis. The MAP-21 emphasizes the eligibility of systemic safety improvements and projects to reduce the potential for traffic-related fatalities and serious injuries on all public roads. FHWA's *Systemic Safety Project Selection Tool* (Systemic Tool) provides supporting information for state DOTs and local governments to incorporate a systemic planning component into their existing safety management programs (*42*).

3.2.4 Questions

The main questions that the guide needs to address to achieve the goals described above include the following:

- How should agencies estimate AADT if they do not have count data?
- Should agencies use default AADT values for certain functional classifications?

- Is there a bias associated with roadways selected for actual traffic counts?
- How should agencies select sample locations to conduct short-duration counts?
- What is the required sample size of short-duration counts?
- Should the agencies expand the ADT of a short-duration count to AADT? If yes, what factoring approach should be used?
- What is the anticipated accuracy of AADT estimates for various sample sizes of short-duration counts and different AADT estimation methods?

3.3 KEY ELEMENTS

To address the questions and the goals listed in the previous section, we propose to include in the guide the key elements discussed in the following subsections.

3.3.1 Framework

The proposed framework will allow target agencies to select the most appropriate methods based on their needs and available resources such as the percent of counted segments. The framework will allow agencies to conduct a gap analysis, determine the accuracy of their (existing) AADT estimates, and identify appropriate actions to improve their program. To address different needs and maturity levels of state practices, some methods may be suitable for estimating AADT at the segment level, while others at higher levels of aggregation.

3.3.2 Sampling Methods

Taking into consideration that the TMG and the HPMS *Field Manual* do not provide guidance on selecting sample segments on local roads and rural minor collectors, we propose to include relevant recommendations in the guide. The guidelines will cover topics on network stratification (e.g., by functional class and AADT volume groups or other variables if count data are not available), creating sample panel sections, confidence and precision levels, required sample size, AADT coefficient of variation, and sampling method(s) (e.g., random selection of segments).

3.3.3 AADT Estimation Methods

The guide will describe each AADT estimation method and provide step-by-step instructions on how to apply it. The information will be presented to a depth suitable for engineers to learn and use a simple or a more advanced method. Interpretation and implementation of results will also be included. The team will use HPMS and TMAS data for three states of different size and geographic location to verify the proposed method(s). The AADT values estimated using a particular method will be compared to the actual AADT obtained from CCSs.

For the selection of the methods to be included in the guide, we will focus only on methods of estimating AADT, not on other metrics such as VMT. We will take into consideration potential concerns on the debate between the simplicity in the interpretability of results produced from simple methods versus those obtained from mathematically or statistically based approaches. For example, there may be marginal improvements in AADT accuracy by using fuzzy logic decision trees over traditional approaches. In this case, decision trees may perform slightly better, but are more difficult to develop and interpret. As a result, the general guidance would suggest (assuming similar results) the simpler, easier to interpret model is the better approach.

3.3.4 Validation

The guide will provide instructions on how to determine the accuracy of AADT estimates. It will include examples on how to calculate and interpret different performance measures (e.g., MAPE and RMSE).

3.3.5 Costs

The guide will include anticipated data collection costs for the three states that will be selected (Task 4) to test the AADT estimation methods and validate the accuracy of the estimates.

3.3.6 Research Results

The guide will reference findings from previous studies and the literature review (Task 2), and research results from Task 4.

3.3.7 Case Studies and Noteworthy Practices

The guide will reference case studies and information gathered from the interviews on noteworthy practices in data collection and AADT estimation for NFAS roads.

REFERENCES

1. Federal Highway Administration (FHWA), *Traffic Monitoring Guide*, Washington D.C., September, 2013.

2. Federal Highway Administration, *Highway Performance Monitoring System Field Manual*, Office of Highway Policy Information, Washington, D.C., March, 2014.

3. Drusch R. Estimating Annual Average Daily Traffic from Short-Term Traffic Counts. *Highway Research Record*, Vol. 118, 1966, pp. 85–95.

4. Federal Highway Administration. *Highway Safety Improvement Program Final Rule*, Section 934.17 MIRE Fundamental Data Elements. Federal Register, Vo. 81, No. 50, March 15, 2016.

5. Fiedler, R., K. Eccles, N. Lefler, A. Fill, and E. Chan. *MIRE Fundamental Data Elements Cost-Benefit Estimation*. Final Report FHWA-SA-13-018, Federal Highway Administration Office of Safety, Washington D.C., 2015.

6. National Cooperative Highway Research Program (NCHRP). *Traffic Forecasting Accuracy Assessment Research*. NCHRP 08-110 (RFP), 2017.

7. Barrett, M., Graves, R., Allen, D., Pigman, J., Abu-lebdeh, G., Aultman-Hall, L., and Bowling, S. *Analysis of Traffic Growth Rates*. Final Research Report KTC-01-15/SPR213-00-1F, Kentucky Transportation Center, Lexington, KY, August, 2001.

8. Blume, K., M. Lombard, S. Quayle, P. Worth, and J. Zegeer. Cost-Effective Reporting of Travel on Local Roads. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1917, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 1–10.

9. Lloyd, J.F., and M.S. French. *Stratification of Locally Owned Roads for Traffic Data Collection*. Final Report FHWA-PA-2006-009-050210, Pennsylvania Department of Transportation, Harrisburg, Pennsylvania, August, 2006.

10. Carswell, M., D. Schade, D.L. Munn. *Phase 1 Report: Summary of Methods Used to Estimate Local Road VMT*. Michigan Department of Transportation, Bureau of Transportation Planning, Asset Management Division, Data Collection, Traffic Information Unit. September 2009.

11. Carswell, M., D. Schade, D.L. Munn. *Phase 2 Report: Alternative Methodologies for Estimating Local Road VMT in Michigan*. Michigan Department of Transportation,

Bureau of Transportation Planning, Asset Management Division, Data Collection, Traffic Information Unit. September 2009.

12. Carswell, M., D. Schade, D.L. Munn. *Phase 3 Report: Alternative Methodologies for Estimating Local Road VMT in Michigan*. Michigan Department of Transportation, Bureau of Transportation Planning, Asset Management Division, Data Collection, Traffic Information Unit. September 2009.

13. Jessberger, S., R. Krile, J. Schroeder, F. Todt, and J. Feng. Improved Annual Average Daily Traffic Estimation Processes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2593, Transportation Research Board, Washington, D.C., 2016, pp. 103–109.

14. Sharma, S., P. Lingras, F. Xu, and P. Kilburn. Application of Neural Networks to Estimate AADT on Low-Volume Roads. *Journal of Transportation Engineering*, Vol. 127, No. 5, 2001, pp. 426-432.

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

16. Gecchelea, G., R. Rossia, M. Gastaldia, and A. Caprinia. Data Mining Methods for Traffic Monitoring Data Analysis: A case study. *Procedia Social and Behavioral Sciences*, Volume 20, 2011, pp. 455–464.

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

18. Lowry, M. Spatial Interpolation of Traffic Counts Based on Origin–Destination Centrality. *Journal of Transport Geography*, Vol. 36, No. 0, 2014, pp. 98-105.

19. Mohamad, D. An Annual Average Daily Traffic Prediction Model for County Roads. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1617, Transportation Research Board, Washington, D.C., 1998, pp. 99–115.

20. Xia, Q., F. Zhao, Z. Chen, L. Shen, and D. Ospina. Estimation of Annual Average Daily Traffic for Nonstate Roads in a Florida County. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1660, Transportation Research Board, Washington, D.C., 1999, pp. 32-40.

21. McCord, M., Y. Yongliang, Z. Jiang, B. Coifman, and P. Goel. Estimating Annual Average Daily Traffic from Satellite Imagery and Air Photos. *Transportation Research*

Record: Journal of the Transportation Research Board, No. 1855, Transportation Research Board, Washington, D.C., 2003, pp. 136-142.

22. Pan, T. *Assignment of Estimated Average Annual Daily Traffic on All Roads in Florida*. Master Thesis, Civil Engineering Department, University of South Florida, 2008.

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

24. Apronti, D.T., J.J. Herpner, and K. Ksaibati. *Wyoming Low-Volume Roads Traffic Volume Estimation*. Final Report FHWA-WY-16/04F, Wyoming Department of Transportation, October, 2015.

25. Shen, L.D., F. Zhao, and D.I. Ospina. *Estimation of Annual Average Daily Traffic for Off-System Roads in Florida*. Final Report, Florida Department of Transportation, July 1999.

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

27. Zhao, F., M.T. Li, and L.F. Chow. *Alternatives for Estimating Seasonal Factors on Rural and Urban Roads in Florida*. Final Report BD015-03, Florida Department of Transportation, Tallahassee, Florida, June, 2004.

28. Eom, J. K., M. S. Park, T.Y. Heo, and L.F. Huntsinger. Improving the Prediction of Annual Average Daily Traffic for Nonfreeway Facilities by Applying a Spatial Statistical Method. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1968, 2006, pp. 20-29.

29. Sun, X., and S. Das. *Developing a Method for Estimating AADT on All Louisiana Roads*. Final Report FHWA/LA.14/158, Louisiana Department of Transportation, Baton Rouge, Louisiana, July, 2015.

30. Staats, W.N. *Estimation of Annual Average Daily Traffic on Local Roads in Kentucky*. Master Thesis, Civil Engineering, University of Kentucky, 2016.

31. Morley, D., J. Gulliver. *Methods to Improve Traffic Flow and Noise Exposure Estimation on Minor Roads*. Environmental Pollution, 2016, pp. 1-9.

32. Zhong, M., and B. L. Hanson. GIS-Based Travel Demand Modeling for Estimating Traffic on Low-Class Roads. *Transportation Planning and Technology*, Vol. 32, No. 5, 2009, pp. 423-439.

33. Berger, A.D. *A Travel Demand Model for Rural Areas*. Master Thesis, Montana State University, Bozeman, Montana, 2012.

34. Institute of Transportation Engineers, *Trip Generation: An ITE Informational Report, 8th ed.*, Washington, D.C., 2008.

35. Martin, W., and N.A. McGuckin. *NCHRP Report 365: Travel Estimation Techniques for Urban Planning*. Final Report, American Association of Transportation Officials, Washington D.C., 1998.

36. Wang, T., A. Gan, and P. Alluri. Estimating Annual Average Daily Traffic for Local Roads for Highway Safety Analysis. *Transportation Research Record, Journal of Transportation Research Board*, No. 2013, Vol. 2398, pp. 60-66.

37. Mansfield, M. Recreational Trip Generation: A Cross Section Analysis of Weekend Pleasure Trips to the Lake District National Park. *Journal of Transport Economics and Policy*, Vol. 3, No. 2, 1969, pp. 152-164.

38. Ewing, G.O. Progress and Problems in the Development of Recreational Trip Generation and Trip Distribution Models. *Leisure Sciences*, Volume 3, Number 1, 1980.

39. Eck, R., and D. Montag. Traffic Effects of Fairs and Festivals on Low-Volume Roads. *Transportation Research Record, Journal of Transportation Research Board*, No. 1819, 2003, pp. 260-264.

40. Turner, S., J. Carson, C.A. Zimmerman, L.J. Wilkinson, K. Travis. *Traffic Monitoring in Recreational Areas*. Federal Highway Administration, Western Federal Lands Highway Division, Vancouver Washington, August, 2010.

41. American Association of State Highway and Transportation Officials (AASHTO), *Highway Safety Manual*, Washington D.C., 2010.

42. Preston, H., R. Storm, J.D. Bennett, and B. Wemple. *Systemic Safety Project Selection Tool*. Final Report FHWA-SA-13-019, Federal Highway Administration Office of Safety, Washington D.C., July, 2013.