Research Article

TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Transportation Research Record
1–14

DOI: 10.1177/03611981211043544
journals.sagepub.com/home/trr
SAGE

Improving Stratification Procedures and Accuracy of Annual Average Daily Traffic (AADT) Estimates for Non-Federal Aid-System (NFAS) Roads

Ioannis Tsapakis[1], Subasish Das[1], Paul Anderson[1], Steven Jessberger[2], and William Holik[1]

Abstract

The 2016 safety Final Rule requires states to have access to annual average daily traffic (AADT) for all public paved roads, including non-Federal aid-system (NFAS) roads. The latter account approximately for 75% of the total roadway mileage in the country making it difficult for agencies to collect traffic data on these roads. Many agencies use stratified sampling procedures to develop default AADT estimates for uncounted segments; however, there is limited guidance and information about how to stratify the network effectively. The goal of this paper is to enhance transportation agencies' ability to improve existing stratification schemes, design new schemes, and ultimately develop more accurate AADT estimates for NFAS roads. The paper presents the results from five pilot studies that validated and compared the performance of current, updated, and new (traditional and decision-tree-based) schemes using readily available data. According to the results, the median absolute percent error of existing AADT estimates, developed by state agencies, ranged between 71% and 120%. Updating these schemes using recent counts resulted in an AADT accuracy improvement of 25%. The best-performing schemes were developed using DTs that improved the AADT accuracy of existing schemes by 41%. Overall, having more strata and very homogenous strata is better than having fewer strata and more samples within each stratum. The analysis revealed that a key to selecting an effective scheme is to determine a critical point, beyond which creating more strata improves the AADT accuracy marginally but increases the required sample size exponentially.

Annual average daily traffic (AADT) is one of the most widely used traffic parameters in transportation engineering. The Federal Highway Administration (FHWA) Traffic Monitoring Guide (TMG) 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 meet data reporting requirements, better inform decision-making, and support various functions related to planning, pavement and highway design, operations, safety, maintenance, and environmental analysis. State Departments of Transportation (DOTs) are required to report AADT every year to the Highway Performance Monitoring System (HPMS) for all Federal-aid facilities (*2*). In addition, the 2016 Highway Safety Improvement Program (HSIP) Final Rule requires states to have access to AADT for all public paved non-Federal aid-system (NFAS) roads and use the AADT in data-driven safety analysis (*3*).

The NFAS roads include rural minor collectors (6R), rural local roads (7R), and urban local roads (7U). Together, these three functional classes account approximately for 75% of the total roadway mileage in the United States (U.S.) making the collection of traffic volume data on these roads a challenging and expensive task (*4*). Unlike higher roadway functional classifications that agencies routinely count to meet federal reporting requirements (i.e., AADT reported every year to HPMS), many agencies have limited or no traffic volume data for NFAS roads. For example, based on Travel Monitoring Analysis System (TMAS) data that state DOTs submit to FHWA on an annual basis, the authors determined that the average number of continuous count sites (CCSs) that measure traffic 24/7 throughout a year is 1.6 CCSs per

[1]Texas A&M Transportation Institute, San Antonio, TX
[2]Federal Highway Administration, Office of Highway Policy Information, Travel Monitoring and Surveys, Washington, D.C.

Corresponding Author:
Ioannis Tsapakis, i-tsapakis@tti.tamu.edu

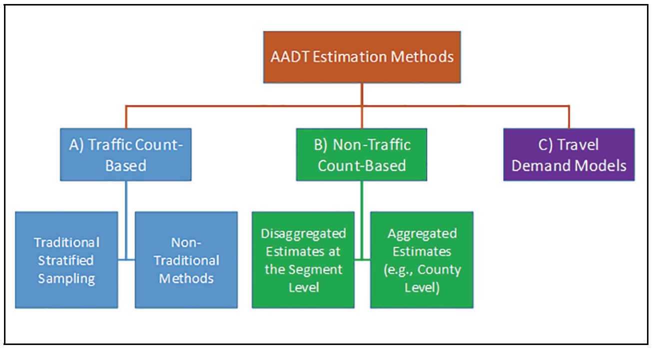

Figure 1. Types of annual average daily traffic (AADT) estimation methods for non-Federal aid-system (NFAS roads)
Source: Tsapakis et al. (5).

state—this number is significantly smaller than the corresponding average number of CCSs located on higher functional classes. Further, many states use portable traffic recorders (PTRs) to conduct short-term counts (STCs) (e.g., for a few hours up to a few weeks) on predetermined roadway segments that constitute a small portion of their NFAS. Covering the extensive NFAS road network with STCs can be financially difficult.

The literature reveals different types of AADT estimation methods that can be broadly divided into three major groups (Figure 1): count-based, non-count-based, and travel demand models (TDMs).

The count-based methods are divided into traditional stratified sampling procedures and non-traditional approaches such as statistical and machine/deep learning methods (*6–20*). Most agencies use stratified sampling procedures to estimate AADT for uncounted NFAS roadway segments (*5*). Though these procedures vary from one state to another, the main steps are the same:

- Stratify NFAS road network to create as homogenous groups (or strata) of roads as possible with respect to AADT. In the absence of traffic volumes for all NFAS roads, the stratification is typically based on one or multiple non-traffic variables (e.g., roadway functional classification, rural/urban designation, county boundaries) that are used as surrogates for AADT.
- Select a sample of segments from each group (or stratum) to conduct STCs. The goal is to select a sample that is representative (in terms of AADT) of all the roads within each stratum.
- Expand or convert each STC into an AADT estimate by multiplying its average daily traffic (ADT) with one or more adjustment factors (e.g., hour of day, day of week, month of year, axle correction factors) that are calculated from CCSs.
- Calculate a default AADT for each stratum as the average of all AADT estimates developed in the previous step. The default AADT of each stratum is assigned to every uncounted segment within the stratum.

The non-traditional count-based methods use a combination of traffic, probe, U.S. Census, land use, and other types of data. Their main advantage is that they directly estimate AADT by avoiding the errors stemming from each step of the traditional approach. Though some studies have reported promising results for some non-traditional methods, the latter have several requirements and shortcomings making it difficult for practitioners to implement them (*15–17*). For example, they:

a) Require statistical software
b) Require knowledge in statistics and machine learning
c) Require non-traffic data that may not be readily available and easily accessible to practitioners
d) Can be difficult to develop, implement, and interpret their results because some methods (e.g., neural networks) are considered to be “black boxes”
e) Can produce different results every time a model is run or by slightly changing some model parameters

The second group includes non-count-based methods that allow the direct estimation of AADT using exclusively non-traffic data such as those mentioned above. These methods can produce estimates at the segment level and at larger geographical areas (e.g., county, zip code) (*21–33*). In general, these methods are not as accurate as count-based methods because the non-traffic variables that are used as predictors tend to have weak-to-moderate correlations with AADT (*5*).

The third group includes TDMs that incorporate mathematical equations to capture travelers’ decisions and allocate trips to roads (*26, 34–38*). Metropolitan Planning Organizations (MPOs) and local agencies typically use TDMs to model primarily small- to medium-size areas and regions. TDMs require significant amounts of development time, resources, and data, making them less favorable over other simpler and less-demanding methods.

Despite the significant number of research studies that have developed non-traditional methods and TDMs, many transportation agencies continue to use traditional sampling approaches to estimate AADT for NFAS roads. The adherence to traditional approaches can be largely attributed to the aforementioned practical challenges and implementation considerations associated with non-traditional methods that cannot be easily overcome. To put things in perspective, as of the date of this paper, the authors are not aware of any state or local transportation agency that currently estimates AADT using non-traditional methods—even for higher functional classes. In addition, existing federal guidelines such as the 2016 TMG and the 2016 HPMS Field Manual do not prescribe how agencies should collect data and estimate AADT for NFAS roads (*1, 2*).

Note that many existing stratified sampling schemes were originally designed a long time ago using simplistic non-statistical approaches to estimate vehicle miles traveled (VMT) for federal reporting purposes. However, VMT estimation methods have fewer data and accuracy requirements compared with (segment-specific) AADT estimation methods, because VMT is typically aggregated and reported at higher geographical levels (e.g., by state, county, functional class) allowing more room for error.

Likewise, the majority of past studies that examined stratified sampling approaches focused on improving the accuracy of VMT estimates (*6–14*). Though some findings may still be applicable for AADT estimation purposes, there are several topics that need further investigation. For example, to the best of the authors' knowledge, there is limited-to-no information about (*a*) the accuracy of default AADT estimates for NFAS roads, and (*b*) the factors that affect the accuracy of these estimates.

As a result of the requirements, gaps, and considerations described above, the largest portion of the transportation network nationwide includes uncounted segments, some of which have been assigned default AADT estimates of unknown accuracy. Nevertheless, many agencies continue to share and diversely use these estimates to support decision-making and feed various analyses, as described earlier. Potentially inaccurate estimates can affect the reliability of these applications and lead to poor results and decisions, including ineffective allocation of funds and under- or over-designed transportation projects, among others.

Goal and Objectives

To address various gaps in the body of literature, the goal of this paper is to provide information that will help researchers, but primarily practitioners, to improve current stratification schemes, design new more effective schemes, and develop more accurate AADT estimates for NFAS roads. To address this goal, the authors conducted five pilot studies with state DOTs (the names of the DOTs are not mentioned for confidentiality purposes). This paper addresses the most important questions that DOT officials raised at the beginning of the pilot studies:

1) Which non-traffic variables should be used to stratify the network effectively? Are administrative variables, such as county boundaries, effective stratification variables?
2) What is the accuracy of existing default AADT estimates that some DOTs have already developed for NFAS roads?
3) Can the accuracy of these estimates be improved without developing a new stratification scheme? If yes, how and what is the expected improvement?
4) Do simple statistical methods such as decision trees (DTs) perform better than manual stratification approaches? What are their advantages and disadvantages?
5) How can different stratification schemes be compared and the most effective one ultimately be selected?
6) Which factors affect the stratification process and ultimately the accuracy of default AADT estimates?

Exploratory Data Analysis

The authors initially conducted an exploratory data analysis to better understand traffic and non-traffic data, identify potential trends and outliers, and determine which non-traffic variables are good surrogates (or proxies) for AADT and, therefore, can be used to stratify the network effectively.

Exploratory Analysis of Traffic Data

The five DOTs provided traffic data that included AADT values derived primarily from short-term counts conducted over the last 3–5 years. The number of CCSs on NFAS roads in the five pilot states was 1, 1, 2, 2, and 11, respectively. Figure 2 shows AADT frequency histograms, box plots, and descriptive statistics disaggregated by functional class and rural/urban code (FC_RU) for two pilot states. Similar trends and results (not shown here because of word limitations) were obtained for the other three pilot states.

The exploratory analysis revealed some general trends that are consistent across all five pilot studies:

- Some suspiciously high AADT values are often observed. The authors investigated and removed some values that appeared to be outliers. Common causes of outliers were incorrect (functional) classification of roadway segments, unusual traffic volumes because of special events, human errors, and equipment malfunction.
- Partly because of outliers, the median AADT is lower than the mean AADT. Overall, the distributions are skewed to the right.
- The standard deviation tends to be higher than the mean AADT. In other words, the coefficient of

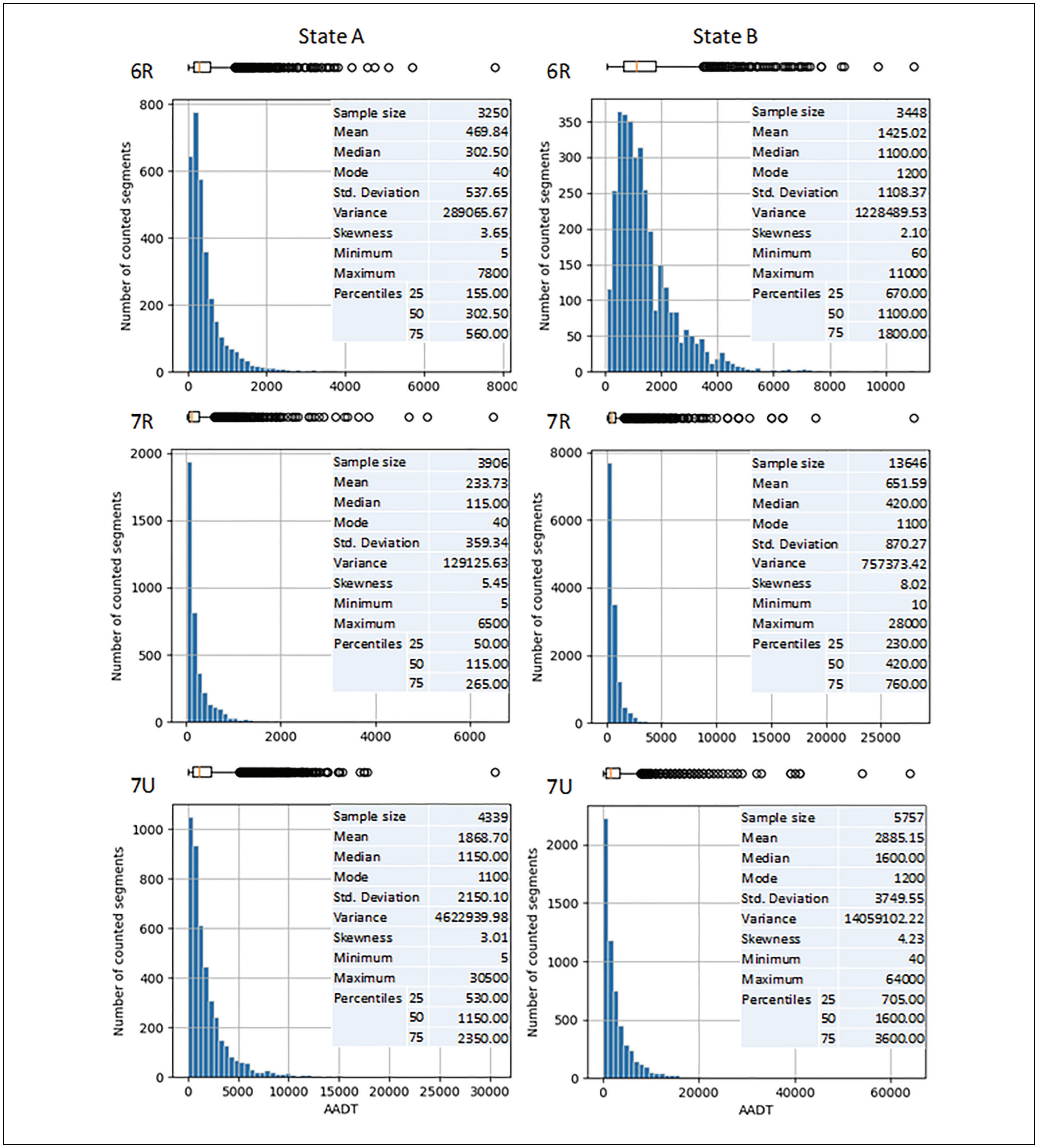

Figure 2. Frequency histograms, box plots, and descriptive statistics for two pilot states.
Note: 6R = rural minor collectors; 7R = rural local roads; and 7U = urban local roads.

variation (= standard deviation/mean) is greater than one.

- Not surprisingly, among the three functional classifications, 7U has the highest traffic volumes, followed by 6R, and 7R.

Despite these similarities, the average AADT of a specific functional class may significantly vary from one state to another. For example, the three mean AADT values of State B (Figure 2) are considerably higher than those of State A. Likewise, other

Variable Description (Abbreviation, Data Source)	State A			State B		
	6R	7R	7U	6R	7R	7U
U.S. Census Bureau and Distance Variables						
Aggregate Earnings (Agg_Earn, ACS)	0.21	0.10	0.10	0.28	0.11	0.10
Aggregate Income (Agg_Inc, ACS)	0.22	0.11	0.09	0.27	0.11	0.08
Aggregate Number of Rooms (Agg_Rooms, ACS)	0.17	0.10	0.10	0.20	0.09	0.02
Workers (Workers, ACS)	0.24	0.11	0.11	0.23	0.11	0.08
Aggregate Number of Vehicles (Agg_Veh, ACS)	0.27	0.17	0.10	0.19	0.06	0.00
Employment (Empl, ACS)	0.25	0.12	0.11	0.23	0.11	0.08
ACS Housing Units (HU, ACS)	0.16	0.11	0.08	0.17	0.07	0.04
Occupied Housing Units (OHU, ACS)	0.19	0.12	0.09	0.22	0.09	0.05
ACS Population (Pop, ACS)	0.20	0.10	0.11	0.22	0.09	0.04
Census Population (C_Pop, Census)	0.19	0.09	0.09	0.19	0.08	0.02
Census Housing Units (C_HU, Census)	0.13	0.10	0.07	0.14	0.06	0.03
Jobs by Workplace Area (WAC, LEHD)	0.33	0.28	0.20	0.21	0.18	0.18
Jobs by Residence Area (RAC, LEHD)	0.20	0.09	0.10	0.25	0.09	0.06
WAC + RAC (WAC_RAC, Calc.*)	0.35	0.26	0.21	0.29	0.19	0.18
Pop + Empl (Pop_Empl, Calc.)	0.22	0.11	0.11	0.23	0.10	0.06
Distance to Closest Interstate (Dist_IH, Calc.)	-0.09	-0.06	-0.09	-0.10	-0.04	-0.09
Distance to Closest U.S. Highways (Dist_US, Calc.)	-0.17	-0.12	0.00	-0.13	-0.11	-0.03
Density (den) of U.S Census Bureau Variables						
Earning Density (Agg_Earn_Den, Calc.)	0.39	0.21	0.05	0.38	0.27	0.22
Income Density (Agg_Inc_Den, Calc.)	0.43	0.23	0.04	0.40	0.26	0.22
Room Density (Agg_Room_Den, Calc.)	0.39	0.20	0.05	0.38	0.29	0.27
Worker Density (Worker_Den, Calc.)	0.54	0.31	0.06	0.41	0.27	0.28
Vehicle Density (Agg_Veh_Den, Calc.)	0.37	0.20	0.06	0.41	0.29	0.29
Employment Density (Empl_Den, Calc.)	0.57	0.32	0.06	0.40	0.27	0.28
ACS Housing Unit Density (HU_Den, Calc.)	0.39	0.22	0.04	0.36	0.26	0.26
Occupied Housing Unit Density (OHU_Den, Calc.)	0.40	0.22	0.04	0.38	0.26	0.27
ACS Population Density (Pop_Den, Calc.)	0.44	0.23	0.06	0.39	0.27	0.27
Census Population Density (C_Pop_Den, Calc.)	0.38	0.20	0.05	0.38	0.26	0.27
Census Housing Unit Density (C_HU_Den, Calc.)	0.34	0.20	0.02	0.35	0.26	0.26
WAC Density (WAC_Den, Calc.)	0.53	0.45	0.11	0.18	0.21	0.15
RAC Density (RAC_Den, Calc.)	0.41	0.22	0.05	0.40	0.26	0.27
WAC + RAC Density (WAC_RAC_Den, Calc.)	0.53	0.42	0.11	0.29	0.25	0.18
Pop + Empl Density (Pop_Empl_Den, Calc.)	0.48	0.26	0.06	0.40	0.27	0.28

Notes: The black bold numbers (e.g., **0.37**) indicate statistically significant correlations (two-tailed) at the 99% confidence level (p<0.01). The black (not bold) numbers (e.g., 0.04) suggest statistically significant correlations (two-tailed) at the 95% confidence level (p<0.05). The red numbers (e.g., 0.01) indicate that correlations are not statistically significant at the 95% confidence level.
* Calc. = Calculated variable.

Figure 3. Correlation coefficients calculated for two pilot states.
Note: 6R = rural minor collectors; 7R = rural local roads; and 7U = urban local roads.

descriptive statistics may vary by state. These differences show that traffic can significantly vary by geography, highlighting the need to perform this analysis separately for each state.

Correlation Analysis

The authors identified several readily available non-traffic variables that have been used in past studies for roadway stratification or AADT prediction purposes (*5–33*). The main data sources are the American Community Survey (ACS), the Census Longitudinal Employer-Household Dynamics (LEHD), and the decennial Census. The first column of Figure 3 shows the variables included in the analysis along with their data source. The authors also calculated composite variables (e.g., population + employment), the geographical density of each original Census variable, and two distance variables: the distance of a count to the closest interstate and U.S. highway, respectively. Note that the variables listed in Figure 3 were separately downloaded/calculated at four geographical units: Census block, block group, tract, and county levels.

After integrating traffic and non-traffic data using GIS tools, a correlation analysis was separately performed for each geographical unit to determine which variables have the strongest relationship with AADT and at what geographical resolution. Indicatively, Figure 3 shows the Pearson product-moment correlation coefficients calculated between AADT and all variables aggregated at the block group level by 6R, 7R, and 7U for two states. The black bold numbers (e.g., **0.37**) indicate statistically significant correlations at the 99% confidence level (two-tailed). The black (not bold) numbers (e.g., 0.04) suggest statistically significant correlations at the 95% confidence level (two-tailed). The red numbers (e.g., 0.01) indicate that correlations are not statistically significant.

General observations and findings from Figure 3 and other similar correlation analyses conducted in the other three pilot studies are summarized below:

- The density of a Census variable typically has a higher correlation with AADT than the raw Census variable that does not account for the size of the geographical unit of interest.
- Though not shown here, among the four geographical units, higher correlations were obtained in the case of Census variables aggregated at the Census block group and tract levels.
- Overall, the variables that exhibit the highest correlations with AADT are:
  - ACS employment density
  - Worker density
  - ACS population density
  - Occupied housing unit density
  - Density of jobs by workplace area
  - Density of jobs by residence area
- However, the correlations may vary from one state to another. A variable that may be strongly correlated with AADT within a specific functional classification of a given state may have a weaker relationship with AADT in another state. This confirms the need to perform a correlation analysis separately for each state. Note that no single variable can be used as is, as a proxy for AADT. Some variables can potentially be used as predictors in statistical and machine learning models to estimate AADT; however, no transportation agency, to the best of the authors' knowledge, has developed and is currently using non-traditional models to estimate AADT for NFAS roads.
- The correlations tend to decrease in the following order: 6R, 7R, and 7U.
- Correlations calculated by FC_RU are typically higher compared with those calculated for all NFAS roads together at the state level.

The variables *ACS Population* and *ACS Housing Units* that are estimated every year by ACS tend to have higher correlations with AADT than the corresponding decennial census variables (*Census Population* and *Census Housing Units*) when the analysis is performed for the 9-year period (e.g., 2011–2019) following the last census year (e.g., 2010).

Existing, Updated, and New Stratification Schemes

This section describes existing schemes developed by three DOTs, schemes updated based on recent counts, and new schemes developed using both the traditional stratification approach and DTs.

Existing and Updated Schemes of DOTs

Of the five pilot states, two have not stratified their network. The other three states stratified their NFAS roads several years/decades ago as follows:

- State A developed 11 strata based on a combination of functional classification, district, surface type (paved/unpaved roads), urban/rural code, and county.
- State B developed 94 strata based on a combination of route system, population, and county.
- State C developed 3 strata that correspond to the three functional classes: 6R, 7R, and 7U.

Every stratum has been assigned a default AADT that each DOT estimated using historical counts that were available when the schemes were developed. The authors used recent traffic data from each stratum to calculate three performance metrics that are presented in section Performance Metrics.

In addition, without modifying the existing strata, the authors updated each scheme by calculating the median AADT within each stratum using the same (recent) traffic data. The median AADT was considered to be the new default AADT of each stratum. Then the performance of each updated scheme was validated as well. The results and the comparison of existing, updated, and new schemes are shown in section Results and Scheme Comparison.

Traditional Stratification – New Schemes

A key to developing an efficient scheme is to use stratification variables that are good surrogates for AADT. The traditional approach involves manually dividing the network into strata using one or more variables, which can be discrete or continuous. Each continuous variable was discretized into four bins (low, low-medium, medium-high, and high) using the 25th, 50th, and 75th quartile values as break points.

In each pilot study, the authors developed 75 traditional or manual (M) schemes using different combinations of the continuous variables listed in Figure 3 along with two nominal variables: *County* and *FC_RU*. However, for simplicity, this paper presents only 28 schemes that, nonetheless, capture all the important trends and findings from the analysis. The schemes are:

Schemes with One Variable:

- County
- FC_RU
- Pop

- HU
- Empl
- RAC
- WAC
- Pop_Den
- HU_Den
- Empl_Den
- RAC_Den
- WAC_Den
- RAC_WAC_Den

Schemes with Two Variables:

- FC_RU, Pop
- FC_RU, HU
- FC_RU, Empl
- FC_RU, RAC
- FC_RU, WAC
- FC_RU, Pop_Den
- FC_RU, HU_Den
- FC_RU, Empl_Den
- FC_RU, RAC_Den
- FC_RU, WAC_Den
- FC_RU, WAC_RAC_Den

Schemes with Three Variables:

- FC_RU, Pop_Den, HU_Den
- FC_RU, Pop_Den, Empl_Den
- FC_RU, Pop_Den, RAC_Den
- FC_RU, Pop_Den, WAC_Den

For example, the scheme developed using *FC_RU* and *Pop* can contain up to 12 strata (= 3 functional classes $\times$ 4 population bins) depending on whether the four population bins exist within each functional class.

Decision Trees (DTs)—New Schemes

DTs belong to the larger umbrella of classification and regression trees (CART) that were first introduced by Breiman et al. (*39*). The goal of DTs is to predict the value of a dependent variable based on several predictors. Classification trees are appropriate when the dependent variable is nominal or discrete, whereas in DTs the target variable is continuous such as AADT. Among several purposes, DTs can be used to create homogeneous strata with respect to AADT, discretize continuous variables, and identify the most influential variables.

Note that there are several machine learning methods that can be used for similar purposes. A typical example is random forests that are simply a collection of multiple DTs that are generated using random subsets of data.

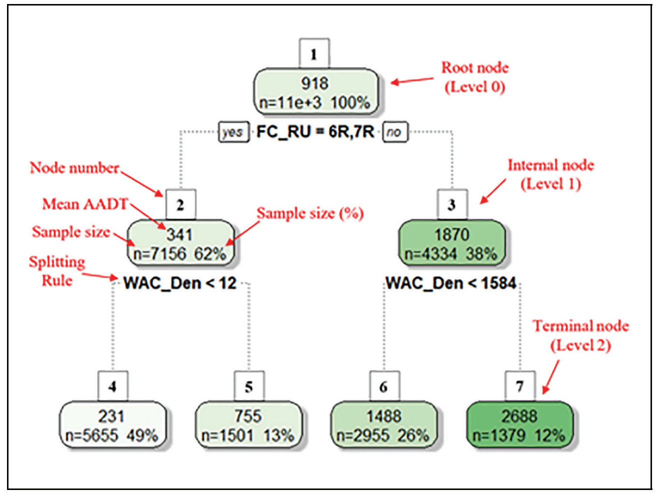

Figure 4. Example of a decision tree (DT) developed in R.
Note: AADT = annual average daily traffic

These methods are more complex, and many consider them as "black boxes" that cannot be easily interpreted. Because of implementation concerns expressed by pilot state officials, DTs were selected to be examined in these studies as they are easier to understand, present, and apply, and have several benefits over the traditional stratification approach. The trees were developed using two open-source R packages: rpart, and rattle (*40*, *41*).

Figure 4 illustrates an example of a simple tree that includes two variables: *FC_RU*, and *WAC_Den*. DTs are flow charts that include nodes interconnected with branches. The first node in a tree is the root node. The nodes that are divided into two nodes are called internal nodes. The nodes that are not split further are the terminal nodes (i.e., strata) of the DT. For example, the tree shown in Figure 4 includes seven nodes: the root node (Node 1 is at Level 0), two internal nodes (Nodes 2 and 3 are at Level 1), and four strata (Nodes 4 through 7 are at Level 2). Annotations are shown in red letters.

DTs are constructed following a top-down process, starting with the first node. The CART-development algorithm scans all independent variables and determines the value (or classes) of the variable that results in the largest possible reduction in AADT heterogeneity. After splitting the root node into two nodes, the algorithm continues to split each new node until a predetermined stopping criterion is met. The tree predicts a default AADT value for each stratum.

Though many splitting rules can be established, most of them are based on the notion of "impurity," which is a measure of the degree of heterogeneity within a node. When all cases within a node have the same value for the dependent variable, the node is considered to be homogenous or "pure" and requires no further splitting. The impurity measure used to develop the trees was the least-squared deviation (LSD), which is computed as the within-node variance. To interpret a tree, two basic principles must be followed:

- The root node and each internal node has a single splitting rule. For example, in Figure 4 the rule of Node 1 is FC_RU = 6R or 7R.
- The left branch of a node means that the rule is met, whereas the right branch means the opposite. For example, the stratification criteria of Stratum 7 are FC_RU = 7U and WAC_Den > 1,584.

Similar to the traditional schemes, the authors developed several DTs, but, for simplicity, only six trees are presented in Section 5. The development of each tree stopped at a different level of growth, starting with Level 2 all the way through Level 7. The AADT derived from recent counts was used as the dependent variable and the independent variables included those listed in Figure 3 along with *County* and *FC_RU*.

Performance Metrics

To validate and compare the performance of the existing, updated, and new (traditional and DT-based) stratification schemes, the following metrics were computed for each scheme:

- Accuracy of default AADT estimates
- AADT variability within each stratum, as well as the overall variability of a scheme
- Number of samples required for a scheme

AADT Accuracy

Among these metrics, the AADT accuracy is of the utmost importance for safety analysis because the AADT estimates must be as close (in magnitude) as possible to actual traffic volumes at individual segments. Initially, the absolute percent error (APE) was separately calculated for each counted segment, as follows:

$$APE_k = \frac{\left|AADT_{Default,i} - AADT_k\right|}{AADT_k} \times 100 \tag{1}$$

where:
APE_k = absolute percent error at counted segment k;
$AADT_{Default,\,i}$ = default AADT of stratum i, which contains counted segment k;
$AADT_k$ = AADT calculated from a CCS or other traffic monitoring device or derived from a short-duration count at segment k.

Then the median APE was computed to quantify the overall accuracy of the default AADT estimates of a scheme.

AADT Variability

The coefficient of variation, C_i, (= standard deviation/mean) was calculated using available traffic data within each stratum i to capture the within-stratum-homogeneity (or AADT variability). The overall variability of a scheme was computed as a weighted average coefficient of variation (WACV):

$$WACV = \frac{\sum_{i=1}^{s}(n_i \times C_i)}{\sum_{i=1}^{s}(n_i)} \tag{2}$$

where:
$WACV$ = weighted average coefficient of variation of a stratification scheme;
s = total number of strata within the scheme;
i = 1, 2, …, s;
n_i = number of counted segments within stratum i;
C_i = coefficient of variation within stratum i.

Sample Size

The sample size required for each stratum was calculated as follows (*2*):

$$n_i = \frac{\frac{Z_i^2 C_i^2}{d_i^2}}{1 + \left(\frac{1}{N_i}\right)\left[\left(\frac{Z_i^2 C_i^2}{d_i^2}\right) - 1\right]} \tag{3}$$

where:
n_i = required sample size within stratum i;
Z_i = value of the standard normal statistic for a specific confidence level (two-sided);
C_i = coefficient of variation (described in the previous section) within stratum i;
d_i = desired precision level (or allowable error) as a proportion of the AADT;
N_i = total number of all counted and uncounted segments within stratum i.

Note that in the case of NFAS roads, N tends to be very large, and the denominator of equation 3 approaches 1.0 and becomes insignificant. As a result, n_i largely depends on the numerator that includes three parameters: C_i, Z_i, and d_i. Previous guidelines and studies simplified formula 3, by retaining only its numerator (*6*, *7*, *9*, *10*):

$$n_i = \frac{Z_i^2 C_i^2}{d_i^2} \tag{4}$$

However, analysts select a confidence (Z_i) and precision (d_i) level based on their needs, therefore the sample size depends (almost entirely) on the coefficient of variation of each stratum.

Figure 5 depicts how the sample size varies with respect to C_i for five pairs of confidence-precision level.

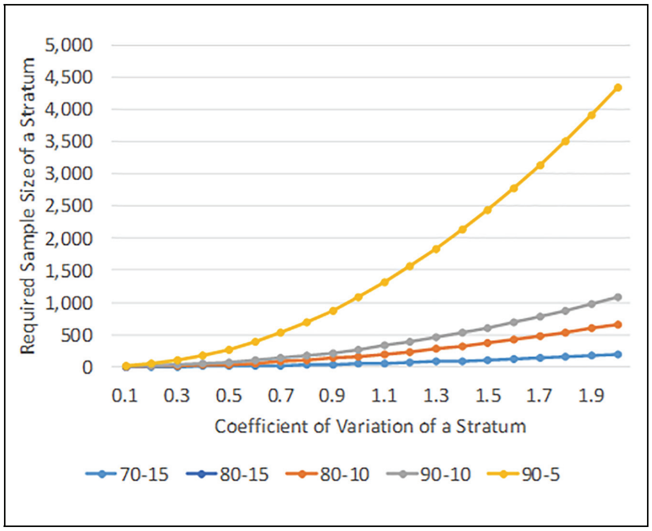

Figure 5. Sample size versus coefficient of variation for different confidence-precision levels.

For any given pair, the more homogenous a stratum, the lower the within-stratum variability (i.e., lower coefficient of variation), and, therefore, the lower the number of required samples, yielding cost savings. This highlights the importance of stratifying the network with the aim to reduce the within-stratum variability rather than to exercise administrative convenience (e.g., to easily manage future counts), which may yield internally heterogeneous strata that will require more counts. The total number of counts required per scheme was calculated as follows:

$$n_{Scheme} = \sum_{i=1}^{s} (n_i) \tag{5}$$

where:
n_{Scheme} = total sample size per scheme;
s = total number of strata within the scheme;
$i = 1, 2, \ldots, s$.

Results and Scheme Comparison

This section shows and compares the performance of existing, updated, traditional, and DT-based schemes. The results from three pilot studies are shown in Figure 6 through Figure 8, respectively. The horizontal axis of each figure shows the number of strata within each scheme. The schemes are shown above each chart and are sorted by the number of their strata to facilitate the comparison. The six DTs are highlighted in blue bold letters and their corresponding data points are enlarged to easily distinguish them from those of the manual (M) schemes. The current scheme of each DOT and the updated scheme are highlighted in black bold letters. The three vertical axes from left to right show the median APE, the sample size required at 70–15 confidence-precision level, and the WACV of each scheme, respectively.

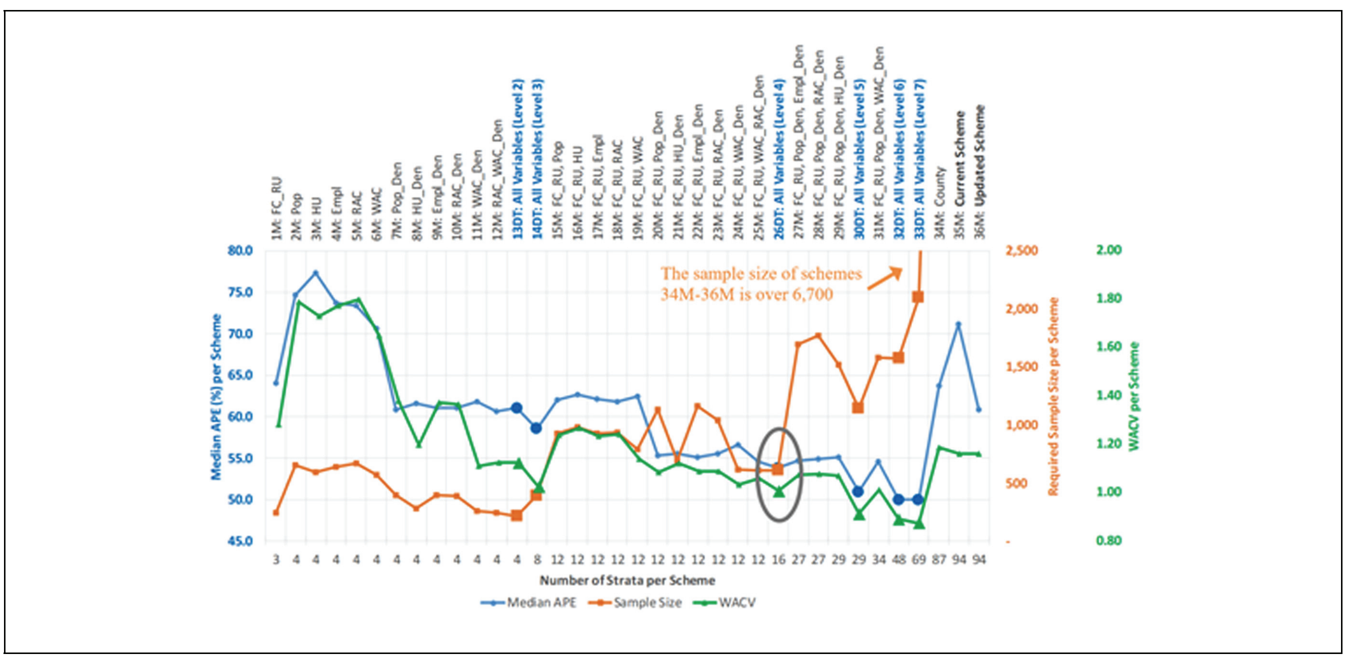

Figure 6. Performance metrics of existing, updated, and new stratification schemes for State A.
Note: APE = absolute percent error; WACV = weighted average coefficient of variation.

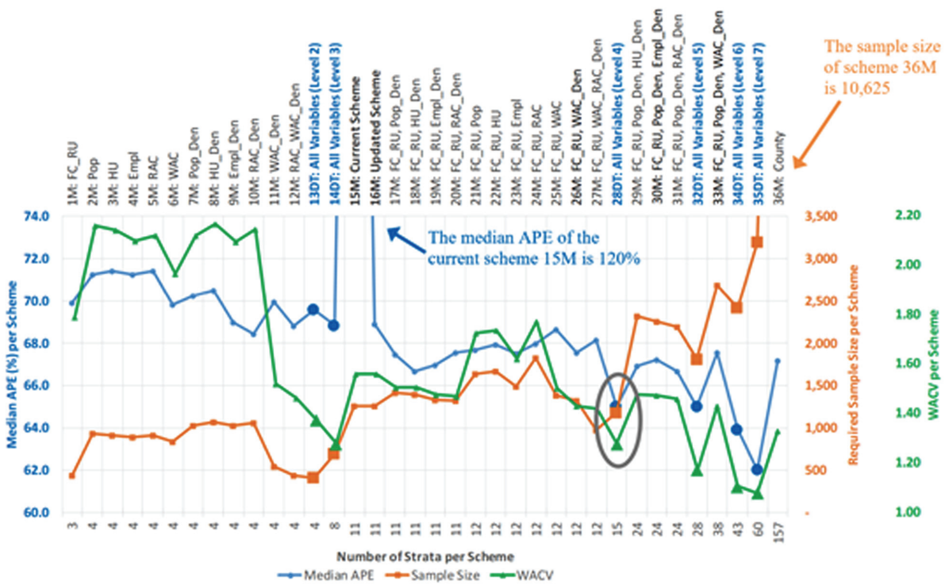

Figure 7. Performance metrics of existing, updated, and new stratification schemes for State B.
Note: APE = absolute percent error; WACV = weighted average coefficient of variation.

General findings from all three figures and similar analyses conducted in the other two pilot studies are discussed below:

- The median APE of the three existing DOT schemes was relatively high: 120%, 71%, and 84%, respectively.
- Updating the current DOT schemes using recent counts improved on average the AADT accuracy by 25%.
- DTs were consistently the best-performing schemes. They produced more accurate AADT estimates than the three existing DOT schemes by 30%, 48%, and 44%, respectively. That is an average improvement of 41%.
- As the number of strata increases, the AADT accuracy tends to improve, but the sample size increases. There is a "critical" point, beyond which creating more strata results in a marginal improvement in AADT accuracy, but a significant increase in the required sample size. For example, in Figure 6, 26DT (highlighted in an oval) is the fourth most accurate scheme and requires considerably fewer counts compared with the three best-performing schemes: 30DT, 32DT, and 33DT. It also contains fewer strata, and, therefore, it is easier to implement. By comparison, 32DT improves the AADT accuracy by only 7%, but requires 158% more counts and contains 32 more strata that require additional implementation effort. Compared with 32DT, the most accurate scheme (33DT) does not provide any improvement in AADT accuracy, yet it requires 33% more counts. This suggests that further dividing the 48 strata of scheme 32DT is meaningless, because the available surrogates cannot capture additional AADT variability within smaller strata.
- DTs tend to produce more homogenous strata and accurate AADT estimates compared with manual schemes that contain similar number of strata. For example, in Figure 8, 27DT results in more accurate estimates than schemes 16M–29M and also requires fewer counts.
- Schemes developed using density variables are more effective than those that include raw Census variables. For example, in Figure 6, 7M–12M are more accurate and have more homogenous strata compared with 2M–6M.
- County-based schemes developed manually (e.g., 36M in Figure 7) tend to produce many internally heterogeneous strata that require a significant number of samples, making these schemes inaccurate and highly expensive.
- Stratifying the network by *FC_RU* is more effective than using a single raw Census variable.

In general, the main factors that affect the effectiveness of the stratification process, and therefore the accuracy of default AADT estimates, are:

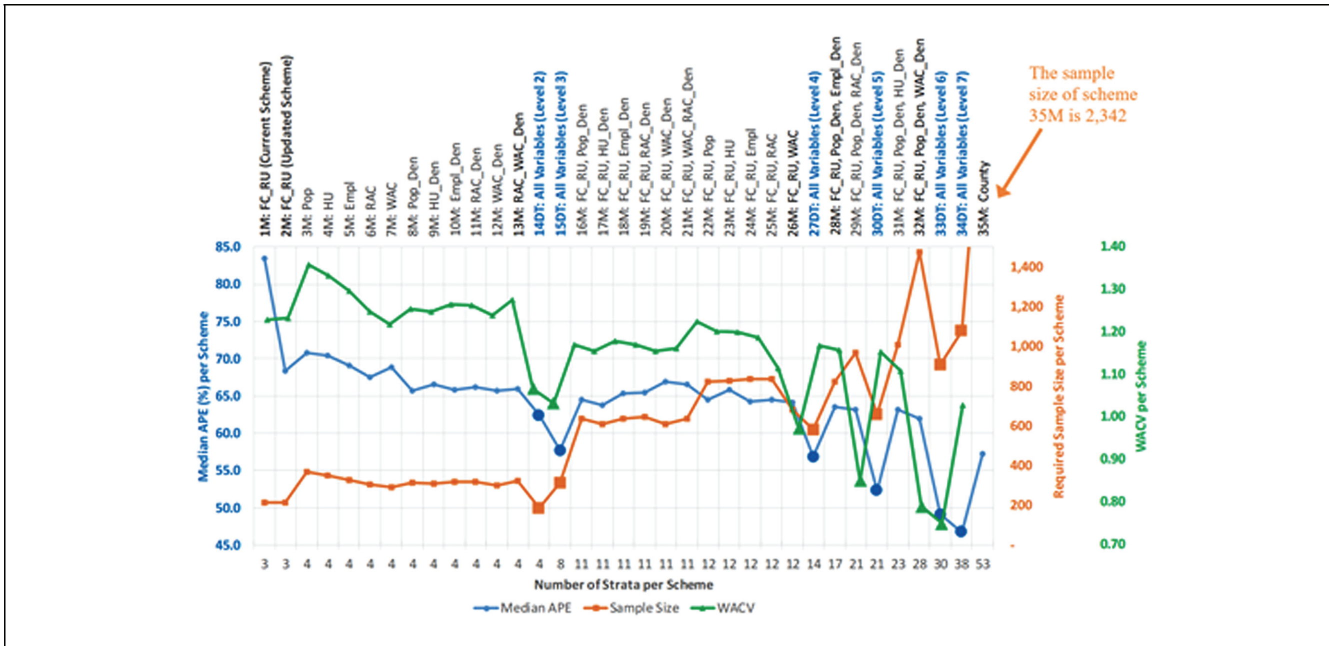

Figure 8. Performance metrics of existing, updated, and new stratification schemes for State C.
Note: APE = absolute percent error; WACV = weighted average coefficient of variation.

- Correlations between stratification variables and AADT. The higher the correlations, the better. As previously explained, there is a "critical" number of strata beyond which the accuracy improves at a much lower rate, but the required sample size increases exponentially. The main reason is that even the best surrogates typically have moderate correlations (e.g., 0.3–0.6 as shown in Figure 3) with AADT. That is, the largest percent of the AADT variability that these variables can explain is often captured during the early stratification stages, when the entire network is divided into the first few strata. Beyond a certain point, further subdividing existing strata does not significantly improve the within-stratum variability, because the surrogates cannot effectively capture the AADT variability within smaller volume ranges. This "critical" number of strata depends on the correlations between AADT and the stratification variables. If the correlations are high (e.g., 0.8–0.9) the critical point is expected to shift to the right. This means that more strata can be created by reducing the within-stratum variability without the sample size increasing exponentially.
- Number of strata. From a safety analysis perspective, the intent is to develop default AADT estimates that are as close as possible to actual traffic volumes at individual segments. In other applications (e.g., VMT estimation), the goal is to predict a precise mean value for an entire group of segments. This distinction is important, because the key to meeting the first purpose is to have more strata and very homogenous strata, whereas, in the second purpose, the key is to have fewer strata but more samples within each stratum (*1*). The results confirmed that the more strata, the higher the accuracy.
- Number of stratification variables. Developing a DT using several independent variables (as opposed to only one) increases the likelihood of explaining more AADT variability within certain volume ranges, where a single variable may not effectively capture. The extra time spent to gather and process more variables is often outweighed by the anticipated improvement in within-stratum-homogeneity and AADT accuracy.
- Stratification method. DTs generally stratify the network more effectively than traditional approaches, and tend to require fewer samples, yielding cost savings.

Conclusions

The goal of this study is to enhance transportation agencies' knowledge and ability to improve potentially existing stratification schemes, design new more effective schemes, and ultimately develop more accurate default AADT estimates for NFAS roads. The paper presents the results from five state pilot studies that aimed to address several questions by validating and comparing the performance of current, updated, and new (traditional and DT-based) schemes. Two important aspects of this research were to use readily available data and examine intuitive stratification methods that are not difficult to understand, develop, and implement. The authors initially performed a data exploratory and correlation analysis that revealed the following:

- The AADT distributions are skewed to the right, therefore the median AADT (not the mean AADT) of a stratum is more appropriate to be used as the default AADT.
- Census variables at the block group and tract levels have higher correlations with AADT than those disaggregated at the block level or aggregated by county.
- The geographical density of a Census variable is typically a better surrogate than the raw Census variable.
- Though correlations vary geographically, the variables that overall exhibit the highest correlations with AADT are Empl_Den, Pop_Den, Worker_Den, OHU_Den, RAC_Den, and WAC_Den. Correlations may vary from one state to another. A variable that may be strongly correlated with AADT within a specific functional classification of a given state may have a weaker relationship with AADT in another state. This confirms the need to perform a correlation analysis separately for each state. No single variable can be used as is, as a proxy for AADT. Some variables can potentially be used as predictors in statistical and machine learning models (e.g., regression models) to estimate AADT; however, no agency, to the best of the authors' knowledge, has developed and is currently using non-traditional models to estimate AADT for NFAS roads.
- The correlations tend to decrease in the following order: 6R, 7R, and 7U.
- Correlations disaggregated by *FC_RU* are typically higher than those aggregated at the state level.

The main conclusions drawn from the comparison of existing, updated, and new schemes are:

- The higher the correlations between stratification variables and AADT, the higher the likelihood that the produced strata will be internally homogeneous.
- From an AADT accuracy standpoint, having more strata and very homogenous strata is better than having fewer strata and more samples within each stratum.
- The median APE of existing schemes ranged between 71% and 120%. A simple way to improve an existing scheme, without modifying the stratification criteria, is to update the default AADT estimates using recent counts—the average improvement was 25%.
- DTs are more effective than traditional approaches as they produce more homogeneous strata and tend to require fewer samples yielding cost savings. In the traditional approach, analysts typically stratify the network using a small number of variables that are arbitrarily discretized. Conversely, DTs can automatically and efficiently scan many variables and determine the optimal threshold within each variable that provides the best separation of a dataset/node. DTs are intuitive and easy to explain to stakeholders, do not require normalization and scaling of data, and are not significantly affected by outliers and missing values, compared with other statistical methods. On the other hand, DTs require statistical programs and knowledge in statistics, can produce different results by changing parameter values, and can be difficult to present to stakeholders if they are large.
- Beyond a critical number of strata, the AADT accuracy improves marginally, but the required sample size increases considerably. The results showed that carefully selected schemes around this critical point not only produced more accurate AADT estimates, but required fewer counts compared with existing, updated, or county-based schemes. Overall, agencies should select one of the best-performing schemes by considering available budgets. If, for example, an agency can afford to conduct the number of counts required by the best-performing scheme, then this scheme should be preferred over other schemes. On average, the best-performing schemes improved the AADT accuracy of existing DOT schemes by 41%.
- County-based schemes developed manually provide for administrative convenience, but often result in a high number of heterogeneous strata that require significantly more counts than other schemes developed using better surrogates. DTs can create more effective county-based schemes. They can identify groups of counties that exhibit similar AADT and then divide each group into more strata based on other variables that are important within each group.

These findings can be used by agencies that currently do not collect data nor estimate AADT for NFAS roads, as well as by those desiring to improve their schemes and the accuracy of their estimates. By feeding various analyses and functions with improved AADT estimates, agencies can produce more accurate results, improve reported statistics, meet new HSIP requirements, enhance decision-making, and make the HSIP project selection, prioritization, and evaluation process more reliable, resulting in more cost-effective HSIP projects.

In addition to the stratification process, the accuracy of default AADT estimates is affected by other factors (not examined in this paper) that are associated with sampling and the traffic counts used in the analysis. Some of these factors that the authors are currently examining include:

- How the counts are selected (i.e., random versus non-random sampling)
- The sampling method used (e.g., simple, stratified, systematic, cluster, multistage sampling)
- Count characteristics (e.g., number, location, quality, duration, timeliness)
- How the counts are annualized (i.e., the reliability and appropriateness of temporal adjustment factors, axle-correction factors, and yearly change factors that are applied to annualize the counts)

Acknowledgments

The authors would like to thank all state DOT officials for providing data and sharing their ideas.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: I. Tsapakis, S. Das, P. Anderson, W. Holik, and S. Jessberger; data collection: I. Tsapakis, S. Das, P. Anderson, and W. Holik; analysis and interpretation of results: I. Tsapakis, S. Das, P. Anderson, W. Holik, and S. Jessberger; draft manuscript preparation: I. Tsapakis, S. Das, P. Anderson, W. Holik, and S. Jessberger. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work was completed as part of FHWA project *Data Collection and AADT Estimation for NFAS Roads (DTFH61-16-D-00004L)*.

ORCID iDs

Ioannis Tsapakis https://orcid.org/0000-0003-4529-4949
Subasish Das https://orcid.org/0000-0002-1671-2753
Paul Anderson https://orcid.org/0000-0003-2833-6668
Steven Jessberger https://orcid.org/0000-0002-3499-4539

References

1. Federal Highway Administration. *Traffic Monitoring Guide*. Washington, D.C., 2016.
2. Federal Highway Administration. *Highway Performance Monitoring System Field Manual*. Office of Highway Policy Information, Washington, D.C., 2016.
3. Federal Highway Administration. *Highway Safety Improvement Program Final Rule*. Docket No. FHWA-2013-0019. Federal Register, Vo. 81, No. 50, 2016.
4. Federal Highway Administration. *Highway Statistics 2016*. Office of Highway Policy Information, Washington, D.C., 2017.
5. Tsapakis, I., W. Holik, S. Das, E. Kraus, and P. Anderson. *Informational Guide on Data Collection and Annual Average Daily Traffic Estimation for Non-Federal Aid System Roads*. Report No. FHWA-SA-20-064. Federal Highway Administration, Washington, D.C., 2020.
6. Mergel, J. *Traffic Estimating Procedures for the Local Functional System*. Prepared by Volpe National Transportation Systems Center for FHWA, Washington, D.C., 1993.
7. Federal Highway Administration, *Travel Estimation Procedures for the Local Functional System*. Office of Highway Information Management, FHWA, Washington, D.C., 1994.
8. Barrett, M., R. Graves, D. Allen, J. Pigman, G. Abu-lebdeh, L. Aultman-Hall, and S. Bowling. Analysis of Traffic Growth Rates. *Final Research Report KTC-01-15/SPR213-00-1F*. Kentucky Transportation Center, Lexington, 2001.
9. 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*, 2005. 1917: 1–10.
10. French, L., and M. S. French. Stratification of Locally Owned Roads for Traffic Data Collection. *Final Report FHWA-PA-2006-009-050210*. Prepared by French Engineering, LLC, Pensylvania Department of Transportation, Bureau of Planning and Research, Harrisburg, 2006.
11. Frawley, W. Random Count Site Selection Process for Statistically Valid Estimations of Local Street Vehicle Miles Traveled. *Transportation Research Record: Journal of the Transportation Research Board*, 2007. 1993: 43–50.
12. Carswell, M., D. Schade, and D. L. Munn. *Phase 1 Report: Summary of Methods Used to Estimate Local Road VMT*. Michigan Department of Transportation, 2009.
13. Carswell, M., D. Schade, and D. L. Munn. *Phase 2 Report: Alternative Methodologies for Estimating Local Road VMT in Michigan*. Michigan Department of Transportation, 2009.
14. Carswell, M., D. Schade, and D. L. Munn. *Phase 3 Report: Alternative Methodologies for Estimating Local Road VMT in Michigan*. Michigan Department of Transportation, 2009.
15. 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.

16. 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.
17. 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*, Vol. 20, 2011, pp. 455–464.
18. Gastaldi, M., R. Rossi, G. Gecchele, and L. Lucia. *Annual Average Daily Traffic Estimation from Seasonal Traffic Counts*. Research Report. University of Padova, Padova, Italy, 2012.
19. Lowry, M. Spatial Interpolation of Traffic Counts Based on Origin–Destination Centrality. *Journal of Transport Geography*, Vol. 36, 2014, pp. 98–105.
20. Federal Highway Administration. *Exploring Non-Traditional Methods to Obtain Vehicle Volume and Class Data*. Transportation Pooled Fund Study TPF-5(384). FHWA, Washington, D.C., 2019.
21. Mohamad, D. An Annual Average Daily Traffic Prediction Model for County Roads. *Transportation Research Record: Journal of the Transportation Research Board*, 1998. 1617: 99–115.
22. 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*, 1999. 1660: 32–40.
23. 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*, 2003. 1855: 136–142.
24. Pan, T. *Assignment of Estimated Average Annual Daily Traffic on All Roads in Florida*. Master thesis. Civil Engineering Department, University of South Florida, 2008.
25. Selby, B., and K. Kockelman. *Spatial Prediction of AADT in Unmeasured Locations by Universal Kriging*. Research Report. The University of Texas, Austin, 2011.
26. 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, 2015.
27. 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, 1999.
28. Seaver, W. L., A. Chatterjee, and M. L. Seaver. Estimation of Traffic Volume on Rural Non-State Roads. *Research Report*. University of Tennessee, Knoxville and University of Georgia, Athens, 2000.
29. 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, 2004.
30. 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*, 2006. 1968: 20–29.
31. 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, 2015.
32. Staats, W. N. *Estimation of Annual Average Daily Traffic on Local Roads in Kentucky*. Master thesis. Civil Engineering, University of Kentucky, 2016.
33. Morley, D., and J. Gulliver. Methods to Improve Traffic Flow and Noise Exposure Estimation on Minor Roads. *Environmental Pollution*, Vol. 216, 2016, pp. 1–9.
34. 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.
35. Berger, A. D. A Travel Demand Model for Rural Areas. *Master thesis*. Montana State University, Bozeman, 2012.
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*, 2013. 2398: 60–66.
37. Zhang, L., A. Nasri, K. Stewart, Q. Sun, J. Fan, H. Younes, F. Habtemichael, and W. Perez. Improved Vehicle Miles Traveled (VMT) Estimation on Non Federal-Aid System Roadways. *Final Report PL-18-031*. Federal Highway Administration, Washington, D.C., 2018.
38. Unnikrishnan, A., M. Figliozzi, M. K. Moughari, and S. Urbina. A Method to Estimate Annual Average Daily Traffic for Minor Facilities for MAP-21 Reporting and Statewide Safety Analysis. *Final Report FHWA-OR-RD-18-17*. Oregon Department of Transportation, 2018.
39. Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1984.
40. Therneau, T., and B. Atkinson. *rpart: Recursive Partitioning and Regression Trees*. package version 4.1-15. 2019. https://CRAN.R-project.org/package=rpart. Accessed August 1, 2020.
41. Williams, G. J. Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery. https://cran.r-project.org/web/packages/rattle/rattle.pdf. Accessed August 1, 2020.

The contents of this paper reflect the views of the author and not the official views or policies of FHWA or TxDOT.