Data-Driven Approach to Quantify and Reduce Error Associated with Assigning Short Duration Counts to Traffic Pattern Groups

Giuseppe Grande¹, Puteri Paramita¹, and Jonathan D. Regehr¹

Abstract
Traffic monitoring agencies collect traffic data samples to estimate annual average daily traffic (AADT) at short duration count sites. The steps to estimate AADT from sample data introduce error that manifests as uncertainty in the AADT statistic and its applications. Past research suggests that the assignment of a short duration count site to a traffic pattern group (TPG), characterized by known traffic periodicities, represents a significant but poorly quantified source of error. This paper presents an approach to quantify the range of errors arising from such assignments and to mitigate these errors using a novel data-driven assignment method. The approach uses simulated 48-hour short duration counts sampled from continuous count sites with known AADT to develop a benchmark of the total error expected when AADT is estimated from such samples. Likewise, the analysis produces a set of AADT estimates using temporal factors from pre-defined TPGs to quantify the range of assignment errors. The data-driven assignment method aims to mitigate these errors by minimizing the absolute mean deviation in AADT estimates produced from multiple short duration counts in a single year. The approach is applied to traffic data collected in Manitoba, Canada, as a case study. The results indicate that the mean absolute error from 48-hour short duration counts is 6.40% of the true AADT and that improper assignment can lead to a range in mean absolute errors of 9%. When applied to previously unassigned sites, the data-driven assignment method reduced mean absolute errors from 10.32%, using a conventional assignment method, to 7.86%.

Background
AADT is a ubiquitous traffic statistic, essential for applications such as infrastructure design and management, road safety assessments, resource allocation, economic appraisals, roadway and transport system planning, operational and environmental analyses, and transportation research (1–8). In the United States, state and municipal transportation agencies submit system-wide estimates of AADT (and vehicle-miles traveled) on their highway networks as required by the Highway Performance Monitoring System (HPMS) (9). Having access to AADT for all paved public roads in the United States by 2026 is also a requirement of the 2016

¹Department of Civil Engineering, University of Manitoba, Winnipeg, Manitoba, Canada

Corresponding Author:
Giuseppe Grande, giuseppe.grande@umanitoba.ca
Table 1. Main Steps and Potential Error Sources Associated with the Estimation of Annual Average Daily Traffic (AADT) at Short Duration Count Sites

| Step | Potential error source | Relevant research |
|------|------------------------|-------------------|
| 1. Calculate temporal factors at each continuous count site. | Temporal factors summarize data by arbitrarily pre-defined periods (e.g., months, days-of-week) and thus do not fully represent non-periodic traffic variations. | Jessberger et al. (3), Grande et al. (4) |
| 2. Group continuous count sites into traffic pattern groups (TPGs) with similar temporal factors. | Continuous count sites may be incorrectly or sub-optimally grouped, producing TPGs with dissimilar temporal factors. | Reimer and Regehr (12), Regehr et al. (13) |
| 3. Calculate average temporal adjustment factors for each TPG. | Grouped sites have similar, but not identical, traffic periodicities; thus, the average factors do not represent the traffic at any individual continuous count site in a group. | Federal Highway Administration (9), Regehr et al. (10), Reimer and Regehr (12) |
| 4. Collect sample traffic data at short duration count sites. | Sampled traffic data may not sufficiently represent the predominant traffic periodicities at the count site. | Sharma and Allipuram (5), Nordback et al. (14), Minge et al. (15), Jackson et al. (16), Gadda et al. (17), Sharma et al. (18), Sharma and Allipuram (5), Gadda et al. (17), Sharma et al. (18), Tsapakis and Schneider (19), Tsapakis et al. (20) |
| 5. Assign short duration count sites to TPGs based on roadway and land use characteristics. | Short duration counts may be assigned to the wrong TPG, particularly if observed roadway and land use characteristics poorly correlate with expected traffic characteristics at the site. | Sharma and Allipuram (5), Krile et al. (11), Gadda et al. (17), Sharma et al. (18), Schroeder (21), Milligan et al. (22) |
| 6. Estimate AADT using temporal adjustment factors and short duration count data. | The average temporal adjustment factors from the assigned TPG may not represent the traffic periodicities at the short duration count site. | Sharma and Allipuram (5), Krile et al. (11), Gadda et al. (17), Sharma et al. (18), Schroeder (21), Milligan et al. (22) |

Highway Safety Improvement Program (HSIP) Final Rule. In Canada, despite a lack of formal data reporting requirements, there is widespread acknowledgment of the value of standardizing and continuously improving highway traffic monitoring practices at all levels of government. This is evident through the recent national-level publication of the Traffic Monitoring Practices Guide for Canadian Provinces and Municipalities (10).

The conventional approach for estimating AADT at short duration count sites, first introduced by Drusch (11), leverages data collected at continuous and short duration count sites (2, 9, 10). Table 1 describes the six main steps of this approach. The first three steps involve data obtained from continuous count sites. These data are used to calculate temporal factors at each site (Step 1), to develop traffic pattern groups (TPGs)—sometimes referred to as factor groups—comprising sites that exhibit similar periodicities (Step 2), and to produce average adjustment factors for each group (Step 3). Because of their high cost of installation, operation, and maintenance, an agency typically deploys a limited number of continuous count sites. Thus, agencies use portable equipment to collect short duration (sample) counts for broader spatial coverage (Step 4) (5). The assignment of each short duration count site to a TPG (Step 5) facilitates the application of the TPG’s temporal adjustment factors and the estimation of AADT at the short duration count site (Step 6). In other words, the assignment step presumes that the average traffic periodicities (monthly, day-of-week, hourly) observed at the sites comprising the appropriate TPG apply to the short duration count site, even though such periodicities are never directly observed at that short duration count site.

Each step of this approach has the potential to introduce error that ultimately manifests as uncertainty in the estimated AADT (2). For the first three steps, error arises because average temporal adjustment factors derived from TPGs do not perfectly represent traffic variations at any particular continuous count site. In the latter three steps, these errors propagate through to the estimation of AADT when assumptions are made about the traffic periodicities at short duration count sites. Overall, the table reveals a major limitation of the conventional approach, namely that the accuracy and precision of the AADT estimated from short duration counts are conditional on multiple sources of error, which may not be quantified in practice even when they are duly recognized.

As cited in Table 1, the literature includes several evaluations of uncertainties arising when estimating AADT from short duration counts. Gadda et al. (17) quantified multiple types of error arising from the use of short duration count data to estimate AADT, including what they refer to as misclassification error (equivalent to the assignment error discussed in Step 5 above). Using data from continuous count sites in Florida, they found that misclassification raised errors from 6.69% (with ideal
factors applied) to 19.35%. They suggested that classifying the count sites into different categories based on the functional classification, lane count, and area types would help to reduce the AADT estimation error.

Past research has identified the assignment step as being the most critical and one that is susceptible to inaccuracies because of human judgment (5, 18–20). Poor assignment may triple the error of AADT estimates from short duration count data (19). Sharma et al. (18) found that proper assignment has more influence on the accuracy of estimates than count duration. They showed that 72 h counts with improper assignment were less accurate than 24 h or 6 h counts using proper assignments. Recent work has focused on the use of data-driven techniques to reduce reliance on human judgment within the assignment process. Tsapakis et al. (20) used a discriminant analysis model for assignment and found a 58% improvement in mean absolute error of AADT estimates. Extending this analysis, Tsapakis and Schneider (19) used machine learning to construct assignment models which improved the mean accuracy of estimates by 65% and reduced standard deviation by 73.7%. Both of these analyses found improvements when using directional traffic volume data in the assignment process.

The research presented in this paper builds on the data-driven assignment techniques explored in the literature. Specifically, it proposes a generic approach to both quantify and reduce assignment errors using existing adjustment factor groupings and available traffic data. The approach recognizes the relative importance of the assignment error when estimating AADT from short duration counts and reveals opportunities to adjust short duration count programs to target error reduction at this step of the AADT estimation process.

Research Objectives and Scope

This paper seeks to improve on the state-of-the-practice in short duration count programs by addressing three research questions:

1. What are the expected errors in current AADT estimates produced from short duration counts? By answering this question, the paper benchmarks errors to provide context for subsequent steps.

2. What portion of AADT estimation errors can be attributed to the assignment of short duration counts to TPGs? Answering this question identifies the component of the error that can be reduced or eliminated by proposing a novel assignment method.

3. What reduction of AADT estimation errors is possible by employing a novel assignment method? The paper proposes and evaluates a data-driven method to reduce the error in AADT estimates that arises when assigning short duration counts to TPGs.

This paper addresses these research questions through a generic approach for quantifying and reducing the errors produced during the assignment step of AADT estimation. The next section describes this approach. The paper then applies the approach using traffic data obtained from Manitoba, Canada as a case study. While the results presented are specific to this case study, the insights generated are considered transferable to other jurisdictions. The final section of the paper discusses these insights and their implications for traffic monitoring programs.

Approach

The proposed approach comprises three steps, which map directly to the three foregoing objectives:

1. Benchmark the total error produced when using the conventional AADT estimation approach for short duration count data.

2. Evaluate the range of errors attributed to the assignment of short duration counts to TPGs by applying factors from multiple TPGs.

3. Develop and apply a data-driven method to assign short duration counts to TPGs and compare the errors produced using this method to those evident from the conventional assignment approach.

The following subsections provide detailed descriptions of each step.

Step 1: Benchmarking Total Error using the Conventional AADT Estimation Approach

This step begins by calculating the (true) AADT at all continuous count sites, by direction, using the AASHTO formulation shown in Equation 1 (9). AADT can only be calculated in this fashion if there exists at least one daily volume for each day-of-week within each month (i.e., \( n_{jm} \geq 1 \) for all days-of-week, \( j \), and months, \( m \)) at a site. In this case, vehicle classification data are not considered (so \( c \) represents all vehicle classes).

\[
AADT_c = \frac{1}{7} \sum_{i=1}^{7} \left[ \frac{1}{12} \sum_{j=1}^{12} \frac{1}{n_{jm}} \left( \sum_{k=1}^{n_{m}} VOL_{ijk,c} \right) \right]
\]  

where

\( VOL_{ijm} = \) total traffic on \( i \)th occurrence of \( j \)th day-of-week within \( m \)th month

...
\[ i = \text{day-of-week, 1–7} \]
\[ j = \text{month-of-year, 1–12} \]
\[ k = \text{occurrence of a particular day-of-week in a particular month} \]
\[ n_{im} = \text{number of times day-of-week} \ i \text{ occurs in month} \ j \text{ with available traffic data} \]
\[ c = \text{vehicle class (optional).} \]

Next, simulated short duration count data are produced by sampling data from continuous count sites. Samples of 48 h duration are taken at each site, by direction, at which AADT can be calculated. To ensure the assumption of normally distributed errors is valid, 100 independent samples are taken per site-direction. The sampling periods are the same for all studied sites and must meet the following criteria, established based on guidance in the literature:

- All short duration counts are conducted between May and September, inclusive.
- All short duration counts begin on Mondays or Tuesdays and cannot include holidays.
- All short duration counts begin between 6:00 a.m. and 6:00 p.m. and last exactly 48 h.

In some cases, sites with the requisite data to produce AADT estimates may have relatively short periods of missing data (i.e., a few hours, but not enough to preclude the use of Equation 1). The sampling criteria require each sample to comprise data from a full 48 h. To address the missing hourly data, the average volume for each hour on each day-of-week within each month is calculated at each site. These averages are used to impute the missing data at all sites at which AADT can be estimated. Existing data are not altered by this method, as only missing data are imputed.

The simulated short duration data are used to estimate AADT using the method recommended in the FHWA Traffic Monitoring Guide (9). Continuous count sites in each TPG are used to develop average temporal adjustment factors for each day-of-week and month. Equation 2 shows how these factors are used to estimate AADT using a single short duration count. Note that the conventional subscripts are presented, whose definitions do not align with those shown in Equation 1.

\[
\text{AADT}_{hi} = \text{VOL}_{hi} \times M_h \times D_h \times A_i \times G_h \quad (2)
\]

where

- \(\text{AADT}_{hi} = \text{the estimated annual average daily traffic at site} \ i \text{ of TPG} \ h\)
- \(\text{VOL}_{hi} = \text{the 48 h volume measured at site} \ i \text{ of TPG} \ h\)
- \(M_h = \text{the applicable seasonal (monthly) factor for TPG} \ h\)
- \(D_h = \text{the applicable day-of-week factor for TPG} \ h \text{ (if needed)}\)
- \(A_i = \text{the applicable axle-correction factor for site} \ i \text{ (if needed)}\)
- \(G_h = \text{the applicable growth factor for TPG} \ h \text{ (if needed)}\)

The analysis applies Equation 1 to the continuous count data at each site to calculate the true AADT and Equation 2 to each sampled short duration count at each site to calculate the estimated AADT. Each estimated AADT is compared with the true AADT for that site. The analysis measures the deviation between the 100 sampled AADT estimates and the true AADT at each site. The range of errors is quantified to establish a benchmark of the errors expected when using the conventional AADT estimation approach. Results are discussed in relation to percent error and absolute percent error relative to the true AADT. Equation 3 shows how the mean percent error (MPE) is calculated for any combination of site and TPG using the 100 sampled short duration counts from the site. Mean absolute percent error (MAPE) is similarly calculated, considering the absolute values of the summands. The analysis identifies the spectrum of errors that are produced using this AADT estimation method for current TPG assignments, thus addressing the first research question: *What are the expected errors in current AADT estimates produced from short duration counts?*

\[
\text{MPE}_{i,h} = \frac{1}{100} \sum_{k=1}^{100} \frac{\text{AADT}_{i} - \text{AADT}_{i,h,k}^{*}}{\text{AADT}_{i}} \quad (3)
\]

where

- \(\text{AADT}_{i} = \text{the ground truth AADT at site} \ i\)
- \(\text{AADT}_{i,h,k}^{*} = \text{the estimated AADT at site} \ i, \text{ based on TPG} \ h, \text{ for the} \ k\text{th data sample}\)
- \(\text{MPE}_{i,h} = \text{the MPE of AADT estimates at site} \ i, \text{ based on TPG} \ h.\)

**Step 2: Evaluating Assignment-Specific Errors**

In Step 2, sampled data are again used to estimate AADTs by applying Equation 2. However, in this step, the analysis considers temporal adjustment factors from all TPGs, not only those to which sites had been previously assigned. This simulates the case in which short duration counts are assigned to different TPGs, including those which would be expected to result in suboptimal AADT estimates.

The sampled data from each continuous count site are used to estimate AADT as if they were assigned to each TPG. Equation 3 finds the MPE for each combination of site and TPG. The results for Step 2 indicate the range of errors that exist when short duration count data are used to estimate AADT. The least of these errors is assumed to be the best case, where the TPG is optimally selected based on the available data. In this case, the remaining error is
also assumed to be the aggregated error produced during each other step of the AADT estimation process. Conversely, the highest of these errors is taken to be the worst case. Thus, the resulting difference of errors in the AADT estimates represents the potential range of assignment errors. This range provides an answer to the second research question: What portion of AADT estimation errors can be attributed to the assignment of short duration counts to TPGs?

Step 3: Developing a Novel Assignment Method

The third step develops a novel method (hereafter referred to as the data-driven assignment method or DDA method) to assign short duration counts to TPGs using short duration count data and available continuous count data. The DDA method is contingent on multiple short duration counts being conducted at each short duration count site during a study year. This is not typical practice for every monitoring agency, but is recommended in traffic monitoring guidance and literature \(5, 18, 19\). Normally, the AADT at a site is estimated by taking the average of all AADT estimates in a year. Theoretically, if traffic patterns at a given site follow the periodicities of a TPG to which it is assigned, these AADT estimates should be close to each other. If instead these AADT estimates are disparate, it suggests that the TPG assignment is inappropriate.

The DDA method applies this theory by measuring and minimizing the absolute mean deviation in AADT estimates when using short duration count data and temporal adjustment factors from multiple TPGs. The method is an optimization problem that selects a TPG by minimizing the absolute mean deviation between all AADT estimates at a site by changing the applied TPG, as expressed in Equation 4.

\[
TPG_{i, \text{assigned}} = \arg \min_{h \in H} \left| \frac{1}{n} \sum_{k=1}^{n} (AADT_{i,k,h} - AADT_{g}) \right|
\]

where

- \(TPG_{i, \text{assigned}}\) = TPG assigned to site \(i\)
- \(h\) = selected TPG
- \(H\) = set of all TPGs
- \(AADT_{i,k,h}\) = AADT estimate for site \(i\) using short duration count \(k\) and TPG \(h\)
- \(AADT_{g}\) = average AADT estimate for site \(i\) using all short duration counts and TPG \(h\)
- \(n\) = number of short duration counts conducted at a site.

The DDA method hypothesizes that minimizing the absolute mean deviation between AADT estimates produced from two or more counts at the same site within a year selects a TPG which also reduces the error produced during the assignment step. In other words, the TPG producing the most precise AADT estimates is also expected to produce the most accurate estimates (note that the accuracy of both the DDA method and conventional assignment methods are limited by the TPG factors available). The following case study tests this hypothesis using data from Manitoba, Canada, thus offering a response to the final research question: What reduction of AADT estimate errors is possible by employing a novel assignment method?

Case Study: Evaluation of Short Duration Count Assignments in Manitoba, Canada

Overview of Manitoba’s Traffic Monitoring Program

The Manitoba Highway Traffic Information System (MHTIS) processes and analyzes traffic data collected on Manitoba’s provincial highway network. Principal traffic volume data sources include 85 continuous count sites and approximately 2,000 short duration count sites. The continuous count sites provide the source data for this case study. Each site records hourly traffic volumes by direction, yielding a possible 170 total site-direction pairs. However, if data are missing from a site for a significant portion of the year, Equation 1 cannot be used to calculate AADT \(9, 10\); thus, some site-directions were unavailable for use in this study. Figure 1 shows
the continuous count sites used in the case study, including those where one or both directions of traffic data did not meet the criteria for estimating AADT in the study year (2018). Missing data are normally the result of temporary equipment malfunctions.

The MHTIS has seven established TPGs. Each TPG comprises a set of sites that exhibit similar temporal traffic patterns and share common roadway and land use characteristics. An unpublished cluster analysis formed the basis for defining these TPGs. The cluster analysis considered the average monthly variation of traffic volume at the continuous count sites to develop initial groupings. A secondary cluster analysis further sorted these sites into groups based on their hourly and day-of-week traffic variations. Finally, the grouping process identified common roadway and land use characteristics shared by the continuous count sites within the clusters formed through the statistical analyses. These characteristics are used to assign short duration count sites to TPGs. An assignment algorithm, based on a sequence of binary-response questions, assists in the assignment process (see Figure 2). Inherent in this process is the assumption that the unmeasured traffic patterns at short duration count sites resemble the measured traffic patterns at the continuous count sites that comprise the TPG to which the short duration count is assigned.

Table 2 shows the defining characteristics for each of the seven TPGs in Manitoba. In general, distinctions between the TPGs depend on their proximity to population centers and recreational destinations, and related predominant trip-making characteristics. For example, sites close to major urban centers generate strong morning and afternoon peaks, higher weekday than weekend traffic, and relatively low seasonal variations (e.g., as in TPG 1). In contrast, sites close to recreational destinations generate relatively high weekend traffic and more pronounced seasonal variations (e.g., as in TPG 6). TPG 7 comprises sites located in northern Manitoba. There are 17 continuous count sites that are

Figure 2. Algorithm for assigning short duration count sites to traffic pattern groups (TPGs) in Manitoba, Canada.
not assigned to any TPG, either because they were installed after the initial cluster analysis took place or because they did not align with the characteristics of any TPG.

For this case study, of the 170 total continuous count site-directions, 146 met the requirements for calculating AADT using Equation 1 (i.e., they measured traffic volumes for at least one day-of-week in every month). Of these, 117 site-directions had pre-assigned TPGs (note that normally TPGs are assigned to sites and not site-directions). Data from these sites were used to develop the temporal adjustment factors for each TPG.

**Results**

**Step 1.** Simulated short duration count data were produced by extracting 48 h samples from the continuous count data at all 146 site-directions for which AADT could be estimated. At first, only data from the 117 pre-assigned site-directions were used to estimate AADT, using Equation 2 and inputting the calculated temporal adjustment factors for the assigned TPGs. Figure 3 shows the resulting percent errors in a histogram, aggregated for all studied sites that have an assigned TPG. Table 3 provides a detailed summary of results in tabular form. The distribution of errors is approximately normal, with a mean of 0.2% and standard deviation of 8.6%. In absolute terms, the MAPE was 6.4% with a standard deviation of 5.7%.

**Step 2.** The analysis was repeated using input temporal adjustment factors from each of the seven TPGs to isolate assignment errors. In this case, unassigned sites are included in the analysis. Table 4 shows the MAPE of AADT estimates using the sampled short duration count

---

**Table 2. Characteristics of Manitoba’s Traffic Pattern Groups**

| Traffic pattern group (TPG) | Temporal characteristics | Roadway, land use, and related trip-making characteristics | Number of site-directions used/total |
|-----------------------------|--------------------------|----------------------------------------------------------|-------------------------------------|
| TPG1                        | High morning and evening peaks on weekdays | Highways near major urban centers | 21/28 |
|                             | Steady weekday traffic with lower weekend traffic | | |
|                             | Low seasonal variation | | |
| TPG2                        | Gradual increase and decrease in hourly traffic | Long-distance trips away from major urban centers | 44/48 |
|                             | Steady weekday traffic with lower weekend traffic | | |
|                             | Moderate seasonal variation | | |
| TPG3                        | Gradual increase and decrease in hourly traffic | Long-distance trips that connect to recreational destinations | 3/6 |
|                             | Low weekend traffic and high Friday/Sunday | | |
|                             | Very high summer peak | | |
| TPG4                        | Evening peak on weekdays | Highways near rural population centers | 23/28 |
|                             | Steady weekday traffic with lower weekends | | |
|                             | Low seasonal variation | | |
| TPG5                        | Evening peak on weekdays and weekends | Highways near population centers that connect to recreational destinations | 6/6 |
|                             | Low weekend traffic and high Friday/Sunday | | |
|                             | High summer peak | | |
| TPG6                        | Steady, high, daytime traffic on weekends | Highways near recreational destinations | 10/10 |
|                             | Low weekend traffic and high Friday/Sunday | | |
|                             | Very high summer peak | | |
| TPG7                        | Evening peak on weekdays | Highways in northern Manitoba | 10/10 |
|                             | Steady weekday traffic with lower weekends | | |
|                             | High summer peak | | |
| No TPG                      | Excluded from other groups | NA | 29/34 |
| Total                       | NA | NA | 146/170 |

**Note:** NA = not applicable.
data from all sites. For comprehensibility, the sites are grouped by their existing TPG assignments and averages are reported across these groupings. The “best” and “worst” columns represent the average errors that are generated by using the TPGs that produce the lowest and highest MAPE, respectively. The “actual” column shows the MAPE when using the site’s assigned TPG (there is no result for this category for the unassigned sites). The average error, aggregated for all assigned sites, ranges from 5.57% using optimal assignments to 14.28% using the worst assignments. As expected, given that some of the unassigned sites have unique temporal characteristics, the average error at these sites is higher, ranging from 10.21% to 19.56%.

**Step 3.** Finally, the DDA method was applied to the unassigned count sites. To align with MHTIS operations, the sampled short duration counts from these unassigned sites were split into two groups: those that start during summer holidays in July or August (53 out of 100 counts), and those that start when school is in session in May, June, or September (47 counts). These counts were used to assign TPGs by applying Equation 3, using one count from each group of counts (i.e., \( n = 2 \) in Equation 3). In total, 2,491 cases were tested for each site (the product of 53 July or August counts and 47 May, June, or September counts). Table 5 shows a sample calculation using the DDA method at site 5 NB with one pair of counts. In this case, the DDA method would assign the site to TPG 1 because that TPG yields the lowest absolute mean deviation between AADT estimates. This process is repeated for all 2,491 combinations of counts at site 5 NB. The 2018 AADT at this site was 186 vehicles per day.

As a comparison case, the assignment algorithm shown in Figure 2 was used to assign the unassigned sites to TPGs, based on their roadway and land use characteristics. Table 6 summarizes the results. The DDA method outperforms the algorithmic method (i.e., it

![Figure 3. Histogram of total errors for annual average daily traffic (AADT) estimates from short duration counts.](image)

Table 3. Summary Statistics of Total Errors for Annual Average Daily Traffic Estimates from Short Duration Counts

| Statistic            | Percent error (%) | Absolute percent error (%) |
|----------------------|-------------------|----------------------------|
| Minimum              | −85.23            | 0.00                       |
| 2.5th percentile     | −16.46            | 0.23                       |
| First quartile       | −4.76             | 2.34                       |
| Median               | 0.38              | 4.95                       |
| Third quartile       | 5.07              | 8.84                       |
| 97.5th percentile    | 17.49             | 20.63                      |
| Maximum              | 50.64             | 85.23                      |
| Mean                 | 0.19              | 6.40                       |
| Standard deviation   | 8.58              | 5.72                       |
produces an AADT that is closer to the true AADT) in 68.3% of tested cases. Both methods produce average errors of 2.34%. However, the overall average MAPE was 7.86% using the DDA method and 10.32% using the assignment algorithm. Using a paired *t*-test, considering each site as an independent sample, the results were statistically significant at the 0.01 level (df = 28, *t*-statistic = 5.3).

When considering individual sites, the assignment algorithm will always assign the same TPG. This static assignment is based on the roadway and land use characteristics at the count site, regardless of the data collected. Conversely, the DDA method produces a dynamic TPG assignment for each site based on the combination of counts used. Table 7 shows the frequency of TPG assignment for each site-direction using both methods and the resulting MAPE. For example, the northbound traffic at site 5 was assigned to TPG 7 using the algorithmic method. When applying the DDA method, it was assigned to TPG 1 in 516 of the tested cases, to TPG 2 in 401 of the tested cases, and so on. As indicated in the final two columns, the MAPE for all test cases using the DDA method for this site was 10.97%, compared with 13.49% for the algorithmic method. Considering all sites examined, the DDA method outperformed the algorithmic assignment method in 23 of the 29 unassigned site-directions (79.3%).

Table 4. Mean Absolute Percent Error by Traffic Pattern Group

| Traffic pattern group (TPG) | Actual assignment (%) | Best assignment (%) | Worst assignment (%) |
|-----------------------------|-----------------------|---------------------|----------------------|
| TPG1                        | 5.34                  | 4.87                | 15.45                |
| TPG2                        | 5.79                  | 5.23                | 13.69                |
| TPG3                        | 7.71                  | 6.54                | 13.53                |
| TPG4                        | 4.86                  | 4.75                | 14.91                |
| TPG5                        | 7.34                  | 4.34                | 11.81                |
| TPG6                        | 9.93                  | 7.51                | 12.81                |
| TPG7                        | 10.31                 | 8.92                | 16.10                |
| All assigned sites          | 6.40                  | 5.57                | 14.28                |
| No assigned TPG             | –                     | 10.21               | 19.56                |

Table 5. Sample Calculation of the Data-Driven Assignment Method

| Count | Start date-time | 48 h volume | AADT estimated using factors from each TPG |
|-------|----------------|-------------|-------------------------------------------|
|       |                |             | TPG 1 | TPG 2 | TPG 3 | TPG 4 | TPG 5 | TPG 6 | TPG 7 |
| Count 1 | September 4, 4:00 p.m. | 447 | 197.8 | 204.2 | 219.3 | 203.4 | 218.4 | 223.2 | 215.7 |
| Count 2 | August 28, 7:00 a.m.    | 420 | 175.8 | 171.1 | 149.7 | 176.3 | 157.3 | 154.2 | 161.4 |
| Absolute mean deviation    |               |             | 11.0  | 16.5  | 34.8  | 13.5  | 30.6  | 34.5  | 27.2  |

Note: AADT = annual average daily traffic; TPG = traffic pattern group.

Table 6. Comparison of the Data-Driven (DDA) and Algorithmic Assignment Methods

| Percent error | DDA (%) | Algorithm (%) |
|---------------|---------|---------------|
| Minimum       | −42.25  | −44.30        |
| Maximum       | 68.42   | 71.58         |
| 2.5%          | −26.60  | −29.57        |
| 25.0%         | −9.96   | −11.96        |
| 50.0%         | −0.80   | −2.82         |
| 75.0%         | 0.74    | 2.79          |
| 97.5%         | 19.17   | 23.35         |
| Mean          | −3.46   | −3.46         |
| Standard deviation | 11.33 | 13.44         |

| Absolute percent error | DDA (%) | Algorithm (%) |
|------------------------|---------|---------------|
| Minimum                | 0.00    | 0.00          |
| Maximum                | 68.42   | 71.58         |
| 2.5%                   | 0.05    | 0.24          |
| 25.0%                  | 0.77    | 2.81          |
| 50.0%                  | 4.97    | 8.16          |
| 75.0%                  | 12.05   | 15.31         |
| 97.5%                  | 30.42   | 32.84         |
| Mean                   | 7.86    | 10.32         |
| Standard deviation     | 8.86    | 9.27          |
Discussion

The case study began by establishing a benchmark for typical (total) error ranges using the conventional methods for assigning and factoring short duration count data to estimate AADT. The results from Step 1 showed that these values vary, depending on the characteristics of the site and its assigned TPG. On average, the simulated 48 h short duration count data produced AADT estimates with an absolute error of 6.40%. These errors were roughly normally distributed with a standard deviation of 8.5%. Thus, to answer the first research question, an agency could expect a total average absolute error of 6.4% when estimating AADT from a single 48 h count. Moreover, based on the distribution of errors, 95% of these estimates would be within 17% of the true AADT at the corresponding site. These results corroborate earlier findings by Milligan et al. (22), which estimated the MAPE to be 6.7% using a similar dataset, but show a higher precision than the 95% confidence limits presented by Krile and Schroeder (21), which were −22.47% to 25.09%. The errors identified here encapsulate each of the sources described in Figure 1 and represent the total expected error associated with using short duration count data to estimate AADT.

The analysis in Step 2 disaggregated the results by TPG to isolate the assignment error (i.e., the portion of error that is associated with the assignment step). Considering the pre-assigned sites, the average MAPEs using the best and worst potential assignments were 5.57% and 14.28%, respectively. This suggests that the range of potential errors in the assignment step is up to 8.71%. However, from Step 1, the MAPE for all assigned sites was 6.40% using the existing assignments. Thus, the existing assignment error, given the study data, was found to be less than 1% on average (i.e., 6.40% minus 5.57%). Notably, these sites were already assigned to TPGs based on their roadway, land use, and traffic characteristics; this depth of knowledge about the site’s

| Table 7. Frequency and Accuracy of Assignments to each Traffic Pattern Group (TPG) by Assignment Method at Previously Unassigned Sites |
|---------------------------------------------------------------|
| **Site-direction** | **True AADT** | **Algorithm** | **TPG 1** | **TPG 2** | **TPG 3** | **TPG 4** | **TPG 5** | **TPG 6** | **TPG 7** | **DA (MAPE (%))** | **Algorithm MAPE (%)** |
|---------------------------------------------------------------|
| 5 NB              | 186.2         | PG7           | 516      | 401      | 250      | 808      | 164      | 150      | 202      | 10.97             | 13.49            |
| 5 SB              | 175.9         | PG7           | 449      | 524      | 321      | 556      | 238      | 182      | 221      | 8.20              | 10.57            |
| 6 NB              | 112.3         | PG7           | 588      | 412      | 200      | 948      | 79       | 70       | 194      | 12.30             | 15.09            |
| 6 SB              | 100.9         | PG7           | 1158     | 160      | 16       | 939      | 59       | 35       | 124      | 28.54             | 30.42            |
| 7 EB              | 96.4          | PG7           | 911      | 267      | 70       | 1070     | 15       | 74       | 84       | 15.76             | 18.42            |
| 7 WB              | 93.9          | PG7           | 1215     | 149      | 157      | 912      | 16       | 9        | 33       | 19.45             | 22.20            |
| 17 EB             | 569.9         | PG6           | 55       | 181      | 1358     | 128      | 257      | 414      | 98       | 17.78             | 20.02            |
| 17 WB             | 576.6         | PG6           | 57       | 202      | 1143     | 157      | 320      | 409      | 203      | 10.39             | 12.63            |
| 18 EB             | 560.6         | PG6           | 202      | 249      | 930      | 238      | 232      | 485      | 155      | 6.51              | 8.60             |
| 18 WB             | 554.0         | PG6           | 198      | 291      | 861      | 190      | 336      | 380      | 235      | 7.11              | 9.48             |
| 19 EB             | 8669.2        | PG1           | 424      | 571      | 257      | 357      | 243      | 219      | 420      | 2.37              | 2.50             |
| 19 WB             | 8785.4        | PG1           | 420      | 596      | 133      | 491      | 144      | 147      | 560      | 2.27              | 2.00             |
| 22 EB             | 1650.9        | PG4           | 222      | 983      | 177      | 303      | 308      | 279      | 219      | 4.02              | 2.73             |
| 22 WB             | 1649.9        | PG4           | 276      | 954      | 61       | 603      | 211      | 102      | 284      | 2.76              | 2.56             |
| 23 NB             | 295.8         | PG2           | 92       | 186      | 815      | 138      | 380      | 621      | 259      | 3.85              | 10.70            |
| 23 SB             | 308.1         | PG2           | 111      | 177      | 635      | 86       | 480      | 730      | 272      | 5.43              | 13.02            |
| 26 NB             | 6524.3        | PG1           | 473      | 738      | 42       | 817      | 80       | 101      | 240      | 1.36              | 2.40%            |
| 26 SB             | 6507.8        | PG1           | 777      | 770      | 79       | 559      | 94       | 185      | 227      | 0.85              | 1.47             |
| 29 NB             | 225.1         | PG3           | 461      | 645      | 86       | 750      | 116      | 85       | 348      | 7.03              | 11.74            |
| 29 SB             | 224.5         | PG3           | 228      | 412      | 376      | 335      | 357      | 386      | 397      | 8.37              | 11.31            |
| 30 NB             | 438.1         | PG2           | 374      | 899      | 62       | 633      | 170      | 43       | 310      | 7.29              | 5.95             |
| 30 SB             | 435.3         | PG2           | 327      | 604      | 307      | 400      | 229      | 175      | 449      | 7.22              | 6.10             |
| 34 EB             | 125.8         | PG7           | 1098     | 407      | 118      | 814      | 5       | 3        | 46       | 8.66              | 12.33            |
| 34 WB             | 126.8         | PG7           | 849      | 526      | 2        | 959      | 35      | 17       | 103      | 8.79              | 12.03            |
| 44 EB             | 1560.0        | PG4           | 387      | 568      | 229      | 227      | 356      | 241      | 483      | 2.75              | 2.30             |
| 52 EB             | 610.7         | PG3           | 473      | 507      | 411      | 455      | 184      | 191      | 270      | 3.53              | 7.41             |
| 52 WB             | 615.2         | PG3           | 481      | 463      | 356      | 425      | 334      | 190      | 242      | 3.50              | 6.92             |
| 93 EB             | 101.1         | PG2           | 88       | 223      | 520      | 162      | 562      | 522      | 414      | 5.06              | 11.34            |
| 93 WB             | 89.0          | PG2           | 32       | 63       | 786      | 33       | 415      | 890      | 272      | 5.86              | 13.65            |

Note: AADT = annual average daily traffic; DDA = data-driven assignment; MAPE = mean absolute percent error; EB = eastbound; WB = westbound; NB = northbound; SB = southbound.
characteristics would not exist at typical short duration count sites.

To address this issue, the currently unassigned sites were included in the analysis in Step 2. The best and worst potential assignments produced average MAPEs of 10.21% and 19.56%, respectively, or a range of 9.35%. Thus, the unassigned sites showed a range in assignment errors similar to the pre-assigned sites. Based on this result, practitioners may conclude that the worst-case assignment would increase AADT errors by up to 9% of the true AADT. This addresses the second research question posed in this paper.

Step 3 considered only the previously unassigned sites. This removed the potential impact of autocorrelation between the sampled short duration count data and the temporal adjustment factors, both of which were created using continuous count data from the pre-assigned sites. The DDA method was used to assign each of these sites using the sampled short duration count data. This process was repeated for each combination of counts, producing a spectrum of assignments for each site. For comparison, the MHTIS assignment algorithm was also used to assign the same sites statically, based only on their roadway and land use characteristics. Only TPG 5 was unrepresented in the study set using this method. The MPE using both methods was −3.46% when aggregating results for all of the unassigned sites. This negative bias indicates that the group of unassigned sites have, on average, higher summer traffic volumes than the assigned TPGs. It is suspected that using samples from the full year would alleviate these biases, although this would deviate from the existing practice and guidance from the literature.

When considering absolute errors, the results using the DDA method were demonstrably more accurate than the assignment algorithm. The MAPE, aggregated for all sites, was reduced by nearly 2.5% (10.32% minus 7.86%) when using the DDA method. The results provide an answer to the third research question: the DDA method produces modest improvements in the overall accuracy of AADT estimates and could be used as a reasonable replacement to the existing algorithmic assignment method. However, the frequency with which the method assigned samples from the same site to the various pre-defined TPGs raises questions about the practicality of dynamically assigning short duration counts to TPGs in this way. Each site was assigned to each TPG at least once. In most cases, one or two TPGs emerge as the predominant assignment selections. While the results showed that, on average, dynamic assignments reduced the assignment error, in a real-life application, the repeated simulated counts from these sites would be unavailable and a pair of counts would yield a single TPG assignment using the DDA method for a particular year.

Overall, the results show that given the available data from Manitoba, Canada, the DDA method can improve the accuracy of AADT estimates using sampled short duration count data. This finding is subject to at least two limitations. First, the study was limited in scope to a single year. It is unclear whether the accuracy improvements found using the DDA method would be consistent in a multi-year study. Second, the study assumed that sampled short duration count data, taken from a continuous count site, emulates the real-life short duration count data collected as part of a conventional traffic monitoring program. Past studies have also employed this technique (5, 18, 19, 21), lending to its credibility. Moreover, the results from Step 3 of the study focus solely on the unassigned sites to minimize the impacts of this limitation on the overall analysis findings.

This paper identifies three avenues for future work. First, the DDA method may be applied to other jurisdictions and for study periods spanning more than one year. The method is meant to be generally applicable, so the results from multiple regions would strengthen this assertion. Second, there is a need to further consider the appropriateness of a dynamic assignment approach, as embodied by the DDA method in this study. The assignment step is predicated on the assumption that a short duration count site experiences traffic periodicities that resemble those at a group of continuous count sites (for which traffic characteristics are well understood). The results from the case study suggest that, perhaps, this assumption does not hold at sites with relatively unknown traffic characteristics, as is the case at short duration count sites. Further study is needed to corroborate this finding. Finally, should practitioners opt to incorporate the principle of dynamic assignment into their traffic monitoring programs, there is a need to evaluate potential trade-offs between the error reduction offered by such a change and any necessary modifications to the existing short duration count program. Specifically, questions remain about the frequency (i.e., number of counts conducted in a calendar year), duration, and count cycle (i.e., the number of years between counts at the same site) that would generate optimal assignments and ultimately minimize the error in AADT estimates from short duration counts.

Conclusion

This paper presents an approach to quantify the range of errors arising from the assignment of short duration counts to TPGs and proposes a novel data-driven method to mitigate these errors. In general, the assignment step is used to infer some connection between the temporal traffic characteristics at a group of continuous
counts and those expected at a site for which only short duration count data are available.

Based on data collected in Manitoba, Canada, the results in this paper indicate that the mean absolute error from 48 h short duration counts is 6.40% of the true AADT and that improper assignment can lead to a range in mean absolute errors of 9%. When applied to previously unassigned sites, the DDA method reduced mean absolute errors from 10.32%, using a conventional assignment method, to 7.86%.

In addition to these results, the paper contributes to traffic monitoring practice by proposing a novel approach to the assignment problem. Conventional practice statically assigns short duration count sites to TPGs, inherently assuming that observed similarities in roadway and land use characteristics correlate with unobserved similarities in traffic periodicities. The results in this paper suggest that any pair of 48 h counts provides sufficient data about a site’s temporal traffic characteristics to yield more accurate AADT estimates than could be produced using the conventional approach based on this assumed correlation. In theory, an agency could apply these findings to improve its expected average accuracy in AADT estimates produced using short duration count data. However, the improved accuracy comes at the expense of consistency and context in TPG assignments. The question, then, is whether consistency and context are valuable or if these are ingrained in the current state-of-the-practice unnecessarily. Further study is needed to explore this question.

Acknowledgments

The authors gratefully acknowledge the financial contributions of Manitoba Infrastructure. The financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC) is gratefully acknowledged.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: G. Grande, J.D. Regehr, P. Paramita; data collection: G. Grande, P. Paramita; analysis and interpretation of results: G. Grande, J.D. Regehr; draft manuscript preparation: G. Grande, P. Paramita, J.D. Regehr. 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) received the following financial support for the research, authorship, and/or publication of this article: Manitoba Infrastructure and the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

1. Albright, D. History of Estimating and Evaluating Annual Traffic Volume Statistics. Transportation Research Record: Journal of the Transportation Research Board, 1991. 1305: 103–107.
2. Tsapakis, I. FY2021 NCHRP Problem Statement. American Association of State Highway and Transportation Officials Special Committee on Research and Innovation, Washington, D.C., 2019.
3. 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, 2016. 2593: 103–109.
4. Grande, G., S. Wood, A. Ominski, and J. D. Regehr. Evaluating Annual Average Daily Traffic Calculation Methods with Continuous Truck Traffic Data. Transportation Research Record: Journal of the Transportation Research Board, 2017. 2644: 30–38.
5. Sharma, S. C., and R. R. Allipuram. Duration and Frequency of Seasonal Traffic Counts. Journal of Transportation Engineering, Vol. 119, No. 3, 1993, pp. 344–359.
6. Olfter, M., K. Klassen-Townsend, P. Paramita, J. Duke, M. G. Dolama, G. Grande, and A. Scaletta. Traffic on Manitoba Highways 2018. University of Manitoba and Manitoba Infrastructure, Winnipeg, Canada, 2019.
7. 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.
8. Li, M.-T., F. Zhao, and L.-F. Chow. Assignment of Seasonal Factor Categories to Urban Coverage Count Stations Using a Fuzzy Decision Tree. Journal of Transportation Engineering, Vol. 132, No. 8, 2006, pp. 654–662.
9. Federal Highway Administration. Traffic Monitoring Guide. United States Department of Transportation, Washington, D.C., 2016.
10. Regehr, J. D., R. Poapst, G. Rempel, J. Montufar, and M. Hallenback. Transportation Association of Canada Traffic Monitoring Practices Guide for Canadian Provinces and Municipalities. Transportation Association of Canada, Ottawa, Ontario, Canada, 2017.
11. Drusch, R. L. Estimating Annual Average Daily Traffic from Short-Term Traffic Counts. Highway Research Record: Journal of the Highway Research Board, 1966. 118: 85–95.
12. Reimer, M., and J. D. Regehr. Hybrid Approach for Clustering Vehicle Classification Data to Support Regional Implementation of the Mechanistic-Empirical Pavement Design Guide. Transportation Research Record: Journal of the Transportation Research Board, 2013. 2339: 112–119.
13. Regehr, J. D., J. Montufar, and H. Hernández-Vega. Traffic Pattern Groups Based on Hourly Traffic Variations in Urban Areas. Journal of Transportation of the Institute of Transportation Engineers, Vol. 7, No. 1, 2015, pp. 1–16.
14. Nordback, K., W. E. Marshall, B. N. Janson, and E. Stolz. Estimating Annual Average Daily Bicyclists: Error and Accuracy. *Transportation Research Record: Journal of the Transportation Research Board*, 2013. 2339: 90–97.

15. Minge, E., C. Falero, G. Lindsey, M. Petesch, and T. Vorvick. *Bicycle and Pedestrian Data Collection Manual*. Minnesota Department of Transportation, Saint Paul, MN, 2017.

16. Jackson, K. N., E. Stolz, and C. Cunningham. Nonmotorized Site Selection Methods for Continuous and Short-Duration Volume Counting. *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2527: 49–57.

17. Gadda, S., K. M. Kockelman, and A. Magoon. Estimates of AADT: Quantifying the Uncertainty. *Proc., World Conference on Transportation Research*, Berkeley, CA, 2007.

18. Sharma, S. C., B. M. Gulati, and S. N. Rizak. Statewide Traffic Volume Studies and Precision of AADT Estimates. *Journal of Transportation Engineering*, Vol. 122, No. 6, 1996, pp. 430–439.

19. Tsapakis, I., and W. H. Schneider. Use of Support Vector Machines to Assign Short-Term Counts to Seasonal Adjustment Factor Groups. *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2527: 8–17.

20. Tsapakis, I., W. H. Schneider, A. Bolbol, and A. Skarlatis. Discriminant Analysis for Assigning Short-Term Counts to Seasonal Adjustment Factor Groupings. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2256: 112–119.

21. Krile, R., and J. B. Schroeder. *Assessing Roadway Traffic Count Duration and Frequency Impacts on Annual Average Daily Traffic Estimation: Evaluating Special Event, Recreational Travel, and Holiday Traffic Variability*. FHWA, Office of Highway Policy Information, Washington, D.C., 2016, p. 42.

22. Milligan, C., J. Montufar, J. Regehr, and B. Ghanney. Road Safety Performance Measures and AADT Uncertainty from Short-Term Counts. *Accident Analysis and Prevention*, Vol. 97, 2016, pp. 186–196.

The results and conclusions presented are those of the authors and no official endorsement by Manitoba Infrastructure is intended or should be inferred.