Proposed Methodology for Developing Intersection Turning Movement Volumes Using Historical Counts and Big Data

MAY 2020

PREPARED BY:
Tyler Mickelson, EIT
Mike Schmitt, AICP CTP, PTP, RSP1
Chris Gregerson, P.E., T.E., PTOE, PTP
Olivia Potash, EIT
Yari Ascencio, EIT
Anthony Gallo, P.E.
Adam Gibson, P.E.

INTRODUCTION

The purpose of this white paper is to establish an acceptable methodology for the development of accurate, affordable, and defensible turning movement volumes despite the existing and potentially long-term inability to collect new traffic count data due to COVID-19. This paper proposes an analytical approach to determining methodology based on the availability of historical data. Where data is outdated, limited, or unavailable, this paper discusses how a big data model can be applied to derive volume estimates. A decision tree is provided to help identify the most appropriate method for developing existing conditions turning movement volumes.

Background

On January 22, 2020, the U.S. reported its first domestic case of Coronavirus (COVID-19)[1]. Widespread infection over the next two months caused state and local officials to enact various public health initiatives to curb the spread of COVID-19. On March 3, 2020, California Governor Newsom declared a state of emergency for California. In response, many school districts and universities began to close or shift to distance learning. In addition, large tech firms in the San Francisco Bay Area began having employees work from home. On March 16, 2020, multiple county public health officials in the greater Bay Area enacted the country's first shelter-in-place orders, prohibiting all non-essential travel. On March 19, 2020, Governor Newsom signed an executive order for a California-wide shelter-in-place[2]. Transportation, traffic, and congestion saw immediate effects.

The effects of the statewide–and eventually nationwide–shutdown of all non-essential travel disrupted the typical methodology for execution of traffic studies. Namely, traffic volume data collected after March 3, 2020, may no longer be representative of typical roadway conditions. Traffic trends reported by Google[3] and INRIX[4], show a nationwide reduction in traffic of 30 percent with reductions as high as 50 percent in the San Francisco Bay Area.

As the spread of COVID-19 declines and travel restrictions lift, an initial increase in traffic can be expected. It is likely, however, that traffic volumes post-COVID-19 will remain lower than volumes observed in January and February of 2020 for a variety of reasons. These reasons include the fact that many local agencies plan to take a staged approach to raising restrictions, significant unemployment and other economic considerations, and potential changes in travel behavior (e.g., new preference for telecommuting, continued nervousness over the outbreak, etc.). It may take an extended period of months, or even years, for traffic volumes to return to or exceed what was observed immediately prior to COVID-19. Like the Great Recession of 2008, the current disruption will make for challenging infrastructure planning as current traffic volumes continue to be lower than historical count data, leaving agencies to determine the conditions and timing for when count data is appropriate for use in preparing operations analyses and making important infrastructure investments.

Purpose and Need

There is a need in the transportation planning and traffic operations field to analyze existing conditions scenarios at intersections with accuracy. The ability to analyze the future effects of ongoing or planned developments on our transportation infrastructure now hinges on our ability to produce intersection turning movement volumes that represent a pre-COVID-19 scenario, which cannot be accomplished by collecting new counts.

[1] https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html
[2] https://covid19.ca.gov/img/Executive-Order-N-33-20.pdf
[3] https://www.google.com/covid19/mobility/
[4] https://inrix.com/blog/2020/03/covid19-us-traffic-volume-synopsis/

The purpose of this white paper is to establish an acceptable methodology for the development of accurate, affordable, and defensible turning movement volumes despite the current inability to collect new traffic count data because of COVID-19.

Organization

This paper is organized into three chapters:

1. Introduction: introduces the need for a new methodology for determining existing turning movement volumes at intersections and a background for why this is necessary.
2. Data Collection Methods: discusses the merits and limitations of various methods for deriving turning movement volumes.
3. Proposed Methodology: proposes the preferred methodology for deriving turning movement volumes at intersections based on available historical data using a decision tree.

DATA COLLECTION METHODS

The purpose of this chapter is to provide a brief background on the available data collection methods that can be used to obtain turning movement volumes and the inherent variability and possible sources of error from each one. Understanding the limitations of collected data, regardless of source, is a prerequisite to using that data in an analysis.

Traffic Counts

Historically, existing conditions intersection turning movement volumes for analysis in a traffic study have been counted on-site during peak hours on a single representative day of the year, which is assumed to represent the average or most common and predictable traffic conditions. This approach typically counts a fair-weather Tuesday, Wednesday, or Thursday on a week without any holidays or major events while nearby schools are in session. The exception for this approach would be projects where peak trip generation is expected to occur at another time, such as on a weekend or holiday, and generate a large share of vehicle trips on the nearby roadways at its peak period of trip generation. A ski resort, regional park, event venue, or sports stadium are examples of this exception where a weekend, late evening, specific event, or holiday count may be needed to supplement or replace a weekday count.

Industry best practice is to calibrate the collected turning movement volumes using daily, monthly, and seasonal adjustment factors based on nearby continuous count locations (when such locations are available). Monthly averages for daily traffic can fluctuate by 15 percent or more from the annual average daily traffic[5] (AADT). However, adjustments to reflect AADT conditions are rarely applied during the normal course of data collection, resulting in inherent differences as compared to average annual conditions. This source of error is important when considering the validity and use of supplementary traffic count data such as big data.

The strength of this traditional count method is that an actual count of every vehicle (and often bicycle and pedestrian) is conducted on-site to establish an up-to-date snapshot in time of traffic conditions. Because these counts are conducted on-site, other qualitative observations can be made at the intersection regarding operations, delay, or queueing. Also, data can be collected at 5-minute or 15-minute granularities to inform proper calibration of traffic simulation models, or so that analyses can be performed on the peak 15-minute period. When available, this method of determining existing conditions intersection counts is the typical standard of practice.

Big Data

The widespread accessibility of geolocation services included in modern electronics like cell phones and vehicles allows for continuous sampling of a transportation network's users. While adhering to strict privacy laws that preserve the anonymity of individual users, this data can be extremely useful for transportation planners in understanding travel trends across space and time. While there may be many big data products useful for estimating traffic volumes, the methodology in this paper was developed based on StreetLight as a big data solution. This paper is not intended to exclude any other similar products. The main limitation of big data is that it does not sample 100 percent of the target population because not every vehicle on the road as geolocation services available or enabled for query. For example, an intersection turning movement count on a Thursday morning from 8:00 AM – 9:00 AM may show 100 vehicles making a northbound right turn. Of the 100 vehicles making that turn, it is likely that some do not have location services in their car, and likely that some drivers do not have location services on their cell phones. Of the vehicles that have location services in a vehicle or cell phone, depending on the permissions of the individual user in allowing use of their location data, even fewer vehicles may be counted. Therefore, big data may show only 15 counts making that

[5] https://dot.ca.gov/programs/traffic-operations/census

northbound right turn at that time. The error of big data volume estimates comes from the conversion of a big data “count” to total vehicle volume on a roadway.

There are many approaches to adjusting and modeling big data to derive traffic volumes at an intersection. The essence of the process is to develop control points where the population (volume) size is known, then determine ratios between sample size and population size, and finally to apply these ratios to other data. There are many geographic models and methodologies for performing this kind of conversion, but regardless of method, the accuracy of control point data and the assumptions built into applying this ratio to other locations are major factors for introducing error between estimated and “actual” average turning movement volumes.

StreetLight Data

Every month, StreetLight collects and indexes anonymized location records. These records come from smart phones and embedded vehicle GPS devices. Using other contextual sources including census, parcel, and digital roadway network data, StreetLight refines the raw data to estimate complete transportation metrics. These metrics are further calibrated and validated using permanent traffic counters and embedded roadway sensors. Because StreetLight data is a continuous sample of existing traffic conditions, averages and estimates are not limited to peak hours on one typical count day. Instead, StreetLight aggregates average traffic patterns over an entire week, month, or year to develop an accurate AADT volume estimate at each study intersection.

A Word of Caution

When comparing results from traffic counts and big data models, the comparison of these two data sources may show even greater (or less) variability than is inherent in either model alone. Traffic count data shows volumes for a single day of the year and is assumed to represent average traffic conditions at an intersection for the entire year (although, as discussed previously, there is inherent error in this approach). A big data model meanwhile grows an average of small samples from the entire year and presents an estimate of average traffic conditions at an intersection. At the end of the day, both methods are estimates that have inherent error. A comparison of intersection volumes derived using both methods may show only one percent variation, but they may still be 10 (and nine) percent from the actual annual average. Similarly, a comparison could show a 20 percent variation; however, each source may only vary from the actual annual average volumes by 10 percent. The potential for the compounding of errors is an important consideration when evaluating big data’s validity, especially given the common bias of transportation professionals to represent single day traffic count data as an accurate reflection of annual averages.

PROPOSED METHODOLOGY

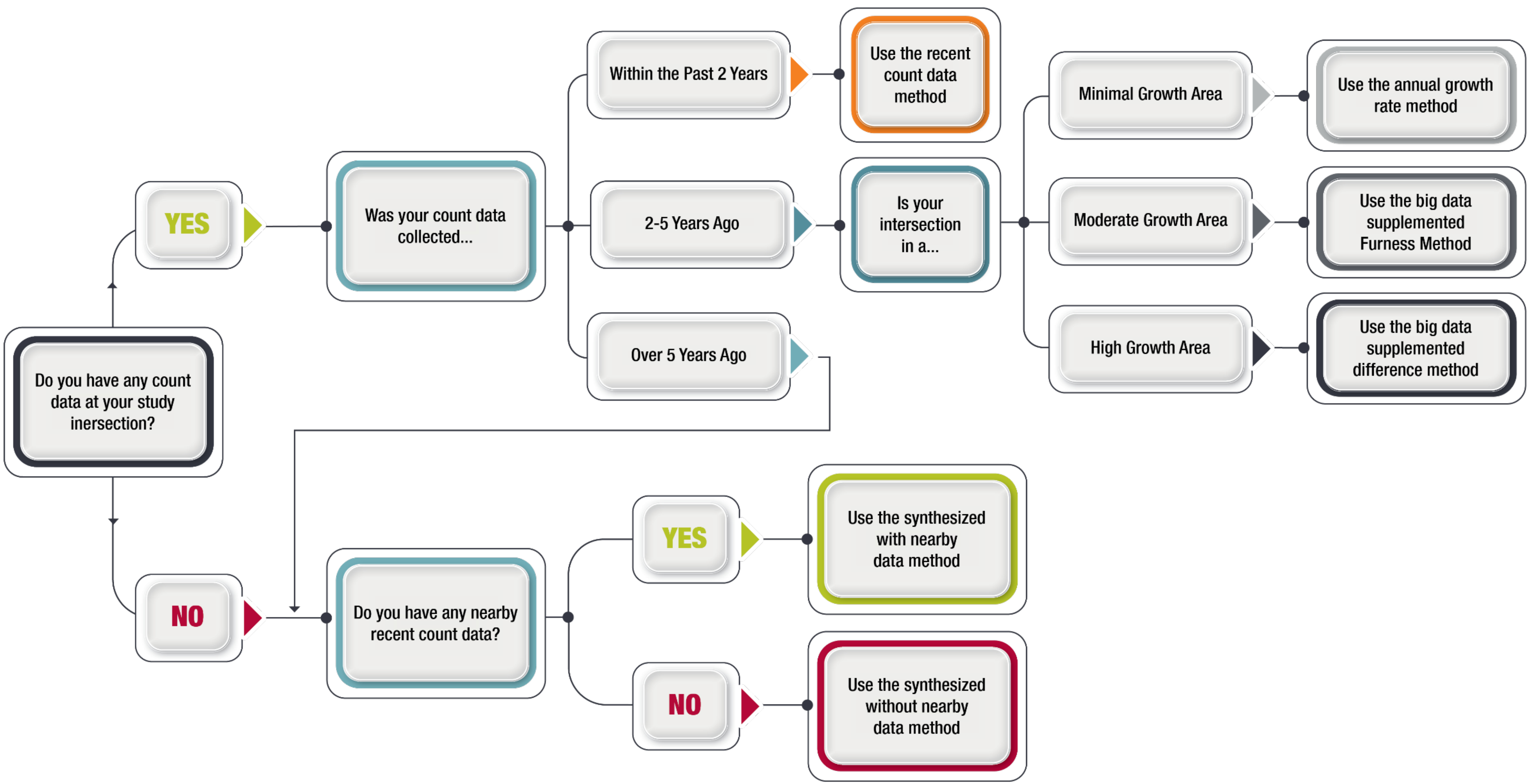
This chapter proposes methodologies for developing intersection turning movement volumes at a study intersection. The decision to use one method over another will depend on what data is available for the study intersection and how this data may need to be supplemented by a big data model. Exhibit 1 shows a decision tree to help identify the appropriate method for an analysis. A description of each method presented in the decision tree is also provided in this chapter.

The scenarios and methods described in this chapter are listed in order of preference. If a big data tool provides the capability to utilize calibration data, this is recommended to be undertaken to the extent possible. Generating specific hourly turning movement volumes "from scratch" (commonly referred to as "synthesizing") from a model is inherently difficult and the confidence in the model outputs directly correlates to the confidence in the data provided to the model for calibration.

Before undertaking the approach to synthesize traffic count data from big data, other sources of data should be reviewed. Depending on the circumstances, it may be preferable to use older traffic count data than would have been permissible prior to the COVID-19 outbreak. Possible sources of turning movement counts may include:

- Previous Studies – many intersections, especially in higher growth areas, have been studied recently and counted multiple times. Publicly available reports may already have counts at this location and be available at no additional cost.
- Local Jurisdiction Data – cities, counties, and state DOTs often collect turning movement counts for monitoring traffic conditions within their jurisdictional boundary. Master plans, TIFs, retiming studies, corridor studies, and other operations studies include count data. Requests to the governing jurisdictions should be made for any available data. This may include recent traffic studies completed and submitted to the jurisdiction by other firms. Because the data they provide can enhance the accuracy and quality of the study, most jurisdictions will not hesitate to provide this data.
- Local Count Firms – Many local count firms are offering to provide historical count data at reduced costs. If data cannot be obtained after exhausting previous studies and local jurisdiction resources, subconsultants and other industry partners should be considered.

Exhibit 1: Decision tree identifying appropriate methodology for derivation of existing conditions study intersection volume



Recent Count Data Method

For locations with recently collected intersection turning movement counts, it remains the best practice to use this data "as is" in a study. Many jurisdictions will allow the use of previously collected data if it's not older than two-years. However, this approach may vary, and as with any traffic study, confirmation from the governing jurisdiction for how recent a count must be to be used without adjustment should be obtained prior to using the count. The acceptable data time frame will likely be a range from one to three years.

If the accuracy of a recent count is called into question because of a change in circumstance, such as a new transportation improvement or the introduction of a major land use nearby between the time of the count and the present, there are several ways to adjust the count data including:

- In collaboration with the jurisdiction, identify all projects that have been completed and opened between the time of the count and the present where their studies added traffic to the study intersection. (Again, this is for recent counts taken within the past two-years. This list should be relatively short; otherwise, this method may become cost prohibitive). Add the opening day projected traffic added by these projects to the recent count using typical analyses methods.
- Identify growth using big data by comparing the volumes at the time of the count to the volumes in present conditions. If there is no growth shown, then this validates the use of the count "as is." If growth is shown, say one percent growth over 18 months, then this can be used to grow the recent count to an existing conditions volume.
- An agreed upon minimum annual growth rate can be applied to the count, conservatively increasing the counts to an acceptable existing conditions volume.
- In cases where travel patterns have changed in response to a new transportation facility, it is possible to use travel demand model data. This option, however, should be carefully considered, given that it would typically be preferential to use big data under these circumstances.

Peak hour factors, truck percentages, pedestrian, bicycle, and time of the peak hour should be taken from the recent count unless judgement suggests these have significantly changed (e.g., a new land use, like a school, could change these parameters).

Older Count Data Methods

If the big data source is available for a period of time less than desired (i.e., it doesn't date back to the count data year), it is acceptable to use a shorter time frame to determine a growth rate if it is determined to be representative of the conditions. Depending on the quality of the data, it may be appropriate to use daily volumes instead of peak period volumes as the basis for the growth rate. For locations where intersection turning movement counts collected in the past five years are available, the changing context of the area should inform the selection of the methodology. In the context of trip generation, determine if the surrounding area experienced a minimal growth, moderate growth, or high growth between the time of the two- to five-year-old count and present conditions. The definitions of these three categories are described as:

- **Minimal Growth Area –** An intersection in a minimal growth area will maintain the same context (urban, suburban, rural) and same land uses between the time of the old count and present day. While specific parcels and ownership may have changed, the zoning, land use intensity, and overall traffic have not changed. Any traffic volume growth at the intersection can be attributed to minimal infill development, or increased pass through traffic. Examples of a no growth area include a completely built-out area of a city with no increase in land use intensity, a completed suburban community with no densification, or a rural agricultural setting with no change in land use.

- Moderate Growth Area – A moderate growth area will maintain the same context (urban, suburban, rural) with increasing land use intensity and development between the time of the old count and present day. While traffic volumes are increasing due to new development near the intersection, the overall zoning and traffic distribution stays steady between the time of the count and present day. A growth area has already established urban/suburban/rural land uses and zoning that do not change; however, the region is also not yet fully developed.
- High Growth Area – A high growth area will have experienced a change in context (i.e., rural to suburban or suburban to urban), often resulting from rezoning of land for completely new development. Due to rezoning, traffic patterns at the intersection may significantly change with respect to time of day and directional distribution. Examples of high growth areas are vacant land being developed into a new master planned community, the recent construction of a major transit hub in a suburban area with resultant mixed-use development and high-density multifamily housing, or an old zoned industrial park rezoned and developed into a commercial retail city node.

Minimal Growth Area – Annual Growth Rate Method

In a minimal growth area, the recommended method for growing an older turning movement count is to apply a historical annual growth rate. While the annual growth in minimal growth areas is typically less than one percent per year in many areas, an annual growth assumption of one percent is a reasonable, albeit conservative, short-term growth basis. Because zoning, land use intensity, and context are not changing in these areas, the traffic distribution captured in an old count may reasonably be assumed constant, and a nominal growth rate method is adequate to capture minor variations in land use intensities (a business hires more workers, etc.).

Peak hour factors, truck percentages, pedestrian, bicycle, and time of the peak hour can reasonably be based on this count data.

Moderate Growth Area – Big Data Supplemented Furness Method

Big data average daily volume estimates should be collected for each approach of the study intersection for 3-6 months around the date of the historical count. Big data average daily volume estimates should then be collected for the 3-6 months prior to existing conditions. Given the effects of COVID-19, this is recommended to include October 2019, January 2020, and February 2020. Because of the nature of big data, it is important to capture a large enough sample while still controlling for “typical” weekdays. For this reason, data should be limited to Tuesdays, Wednesdays, and Thursdays.

Using the count data, the historic big data volume, and the present big data volume, the NCHRP 765 iterative directional method, commonly referred to as the Furness Method, should be applied to the old count to iteratively grow traffic volumes to existing conditions. This process provides corrective factoring to the directional distribution of traffic at the intersection that may not have been captured otherwise in the application of an average annual growth rate.

It may be appropriate to provide minor manual adjustments to peak hour factors, truck percentages, pedestrian, bicycle, and time of the peak hour based on changes in travel patterns that result from the application of the Furness Method. In general, however, it is not anticipated that this would be necessary.

High Growth Area – Big Data Supplemented Difference Method

Big data average *peak hour* turning movement volume estimates should be collected for each movement of the study intersection for 3-6 months around the date of the historical count. Big data average *peak hour* volume estimates should then be collected for the 3-6 months prior to existing conditions. Given the effects of COVID-19, this is recommended to include October 2019, January 2020, and February 2020. Because of the nature of big data, it is important to capture a large enough sample while still controlling for "typical" weekdays. For this reason, data should be limited to Tuesdays, Wednesdays, and Thursdays.

Using the count data, the historic big data hourly turning movement volumes and the present hourly big data turning movement volumes, the difference method should be applied as follows:

- The difference between existing big data turning movement volumes and historic big data turning movement volumes should be calculated for each movement
- The difference in big data volumes should be added directly to the historic count

Peak hour factors, truck percentages, pedestrian, bicycle, and time of the peak hour should be reviewed for appropriateness and manually updated based on reasonable manual adjustments that are supported by the changes made to count data based on the method described above. It may also be appropriate to supplement this data with alternative sources depending on the magnitude of change.

Care needs to be used when applying this methodology to determine when the initial count data is of such poor quality, or the changes so significant, that a full synthesis of the traffic count data might be preferable.

Synthesized Traffic Count Data with Nearby Traffic Data Method

When no turning movement count data is available at the study intersection, or this data is more than five years old, but you have recent daily traffic volumes, then a calibrated big data model may be used to synthesize traffic count data.

Peak hour factors, truck percentages, pedestrian, bicycle, and time of the peak hour should be taken from the nearby calibration data when possible. As appropriate, other datasets or methodologies may be used to supplement this information.

Synthesized Traffic Count Data without Nearby Traffic Data Method

When traffic data is not available at or nearby the study area intersection, then the final recourse is to synthesize intersection turning movement volumes using big data without additional calibration data. The big data should be used to determine average hourly volume on a Tuesday, Wednesday, or Thursday for the most recent 3-6 months (pre-COVID-19). At a minimum this should include October 2019, January 2019, and February 2019.

Truck percentage, pedestrian, or bicycle data may need to be derived from local published rates or other collected data. State and Federal transportation agencies and other professional organizations may provide helpful resources on how to estimate this data when it is not available based on facility type, location, and other factors.