## Vehicle Operating Speed

 Vehicle Operial Roadwayson Urban Arterial


## Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

## TECHNICAL REPORT DOCUMENTATION PAGE

| 1. Report No. 2. G <br> TTI-01-04  | nmen | sion No. | 3. Recipient's Catalog No. |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 4. Title and Subtitle <br> Vehicle Operating Speed on Urban Arterial Roadways |  |  | 5. Report Date January 2019 |  |  |
|  |  |  | 6. Performing Organization Code: |  |  |
| 7. Author(s) <br> Kay Fitzpatrick Subasish Das |  |  | 8. Performing Organization Report No. Report TTI-01-04 |  |  |
| 9. Performing Organization Name an Safe-D National UTC <br> Texas A\&M Transportation Institute The Texas A\&M University System College Station, Texas 77843-3135 |  |  | 10. Work Unit No. |  |  |
|  |  |  | 11. Contract or Grant No. 69A3551747115/Project TTI-01-04 |  |  |
| 12. Sponsoring Agency Name and Address Office of the Secretary of Transportation (OST) U.S. Department of Transportation (US DOT) State of Texas |  |  | 13. Type of Report and Period Final Research Report |  |  |
|  |  |  | 14. Sponsoring Agency Code |  |  |
| 15. Supplementary Notes <br> This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation - Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program, and, in part, with general revenue funds from the State of Texas. |  |  |  |  |  |
| 16. Abstract <br> This research explored (1) the relationship between suburban vehicle operating speed and roadway characteristics, especially the presence of bicyclists and (2) whether crowdsourced speed data could be used to estimate the unconstrained speed for a location. Both vehicle volume per lane and bicycle volume were found to be influential in affecting average speed on lower speed urban arterial roadways. For $40.3 \mathrm{~km} / \mathrm{hr}(25 \mathrm{mph})$ sites, an increase of 19 vehicles per $15-\mathrm{min}$ period would decrease average speed by $1.6 \mathrm{~km} / \mathrm{hr}(1 \mathrm{mph})$, and an increase of more than 39 bicyclists per $15-\mathrm{min}$ period would decrease average vehicle speed by a similar amount. Because of the limited number of $15-\mathrm{min}$ periods with bicycle counts greater than 1 , the research team also developed a model using all available 15min periods with on-road speed data. Speed and volume data in $15-\mathrm{min}$ increments for 2 weeks at nine sites were obtained using on-road tubes and via a vendor of crowdsourced speed data. The difference between the tube data and the crowdsourced data was calculated and called TMCS as a representation of tube (T) minus (M) crowdsourced (CS). The geometric variables that had the greatest influence on TMCS were the number of signals and the number of driveways within a corridor. When only including non-congested periods, weekends (Saturday or Sunday) were associated with the smallest TMCS. |  |  |  |  |  |
| 17. Key Words <br> Vehicle operating speed, arterials, bicyclists, crowdsourced data |  | 18. Di <br> No res <br> public <br> well a <br> Nation <br> Librar <br> Center <br> Librar | bution St tions. Thi ough the <br> followin <br> Transporta <br> Volpe Nati <br> ederal Hig <br> nd the N | ent <br> ocument is availab <br> -D National UTC <br> epositories: VTech <br> Library, The Tra <br> l Transportation Sy <br> ay Administration <br> al Technical Repo | o the bsite, as orks, The portation tems esearch Library. |
| 19. Security Classif. (of this report) <br> Unclassified | 20. Security Classif. (of this page) <br> Unclassified |  |  | 21. No. of Pages 23 | $\begin{aligned} & \text { 22. Price } \\ & \$ 0 \end{aligned}$ |


#### Abstract

This research explored (1) the relationship between suburban vehicle operating speed and roadway characteristics, especially the presence of bicyclists and (2) whether crowdsourced speed data could be used to estimate the unconstrained speed for a location. Both vehicle volume per lane and bicycle volume were found to be influential in affecting average speed on lower speed urban arterial roadways. For $40.3 \mathrm{~km} / \mathrm{hr}$ ( 25 mph ) sites, an increase of 19 vehicles per 15-min period would decrease average speed by 1.6 $\mathrm{km} / \mathrm{hr}(1 \mathrm{mph})$, and an increase of more than 39 bicyclists per 15-min period would decrease average vehicle speed by a similar amount. Because of the limited number of 15-min periods with bicycle counts greater than 1, the research team also developed a model using all available 15-min periods with on-road speed data. Speed and volume data in 15-min increments for 2 weeks at nine sites were obtained using on-road tubes and via a vendor of crowdsourced speed data. The difference between the tube data and the crowdsourced data was calculated and called TMCS as a representation of tube ( $T$ ) minus (M) crowdsourced (CS). The geometric variables that had the greatest influence on TMCS were the number of signals and the number of driveways within a corridor. When only including non-congested periods, weekends (Saturday or Sunday) were associated with the smallest TMCS.

\section*{Acknowledgements}

This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation - Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program, and, in part, with general revenue funds from the State of Texas.

The authors appreciate the assistance provided by the students on this project. Manaswini Condor, Marie Connie Rodriquez, and Elizabeth Clark. Sherry Ryan of San Diego State University assisted the research team with background information on the bike-counting network in San Diego along with road characteristics. Marcus Brewer provided a technical review of this report. The authors appreciate the help provided by Sherry and Marcus.


## Table of Contents

INTRODUCTION ..... 1
BACKGROUND ..... 1
Speed on Arterial Streets ..... 1
Bicycle Presence ..... 2
Performance Measures ..... 2
Research Objectives ..... 4
DATA ..... 5
Site Selection ..... 5
San Diego Bicycle Data ..... 5
Pneumatic Road Tube Data ..... 5
Crowdsourced Data ..... 6
Geometric and Traffic Control Device Data ..... 6
Additional Variables ..... 7
Level of Service (LOS) ..... 7
Light Level ..... 8
Day of the Week ..... 8
Crowdsourced Segment Length ..... 8
PREPARATION OF THE DATABASES ..... 8
RESULTS: INFLUENCES ON VEHICLE OPERATING SPEED ..... 9
Exploratory Data Analysis ..... 9
Model Development ..... 12
Model Results ..... 13
Model Based on All Available On-Road Speed Data ..... 13
Models Focusing on When Bicycle Counts are Available ..... 13
RESULTS: CROWDSOURCED DATA ..... 15
Exploratory Data Analysis ..... 15
Operational Analysis ..... 16
CONCLUSIONS AND RECOMMENDATIONS ..... 18
Bike Count Data ..... 18
Crowdsourced Data ..... 19
Challenges and Future Research Needs ..... 19
ADDITIONAL PRODUCTS ..... 20
Education and Workforce Development Products ..... 20
Technology Transfer Products ..... 20
Data Products ..... 20
REFERENCES ..... 21

## List of Figures

Figure 1. Relationship between three key measures for different time of day ..... 10
Figure 2. Violin and box plot distribution of three key measures per location ID. ..... 11
Figure 3. Variable importance plots ..... 13
Figure 4. Box plot for tube and crowdsourced speeds at each Site-Dir. ..... 15
Figure 5. Speed difference variation per Site-Dir for each day ..... 16
Figure 6. Decision tree for speed difference using only geometric variables. ..... 17
Figure 7. Decision tree for speed difference using temporal variables when only considering those15 -min periods when conditions are LOS A, B, or C.18

## List of Tables

Table 1. On-Site Data Collection................................................................................................... 6
Table 2. Description of Key Variables* ......................................................................................... 7
Table 3. Model Outputs for 15-min Periods with On-Road Speed Data...................................... 14
Table 4. Model Outputs for $15-\mathrm{min}$ Periods during Daylight Conditions and More Than One
Bicyclist on $40.3 \mathrm{~km} / \mathrm{hr}(25 \mathrm{mph})$ Roadways 14

## Introduction

Due to the mounting evidence of its health benefits and association with a reduction in air pollution, bicycling as a transportation mode has been become increasingly popular in recent years, resulting in the addition of more on-street bicycle lanes and separated bicycle facilities [1, 2]. McKenzie reported that the number of U.S. bike commuters increased from about 488,000 in 2000 to about 786,000 in 2008-2012, representing the largest percentage increase among all travel modes [3]. This increase in bicycle traffic requires a greater emphasis on bicyclist safety, as bicyclists are more vulnerable than vehicle drivers. The "safety in numbers" hypothesis supports the idea that vehicle speeds may be affected by an increase in bicycle traffic. The hypothesis is that, as more people cycle, the roadway becomes safer for cyclists because more cyclists are more visible to motorists, which may, in turn, make drivers cautious about keeping sufficient lateral clearance and reducing speed [4]. Identifying the amount of the speed change due to the presence of bicyclists could provide a better understanding of the potential for changes in bicyclist-vehicle crashes, including variations in crash severity levels. The primary goal of this research was to identify key contributing factors that influence vehicle operating speed in the presence of bicyclists on urban arterials.

The collection of on-road speed data provides the opportunity to compare that data with crowdsourced data. There are a number of technologies, such as Bluetooth or GPS probe vehicles, that can be used to collect travel time and speed data. Private companies are collecting and selling this data for several applications. Due to the expense of collecting on-site travel performance measures, such as delay or spot speeds, and with the growing availability of crowdsourced data, we investigated whether there is a reasonable protocol for using crowdsourced data as a representation of the operating speed at a select location or along a corridor.

## Background

## Speed on Arterial Streets

Based on information in the literature [5], several arterial street segment characteristics are known or are suspected to affect drivers' speed choice. In general, the following relationships were found:

- Operating speeds decrease as access density increases.
- Operating speeds decrease as the frequency of roadside objects increases.
- Operating speeds are higher with higher posted speed limits.
- Operating speeds are lower on horizontal curves with small radii or a larger deflection angle.

The research team obtained these roadway segment characteristics along with others for the evaluation.

## Bicycle Presence

A limited number of studies exists on vehicle operations in the presence of bicyclists. The majority of studies have focused on either the safe lateral distance from or the comfort levels of bicyclists. Li et al. reported that vehicle flow is an important factor affecting bicycle travelers' perception of comfort on on-street bike lanes [6]. Chen et al. reported that installation of bicycle lanes did not lead to an increase in crashes due to the reduced vehicle speeds because of an increased awareness of bicyclists [7]. Parkin and Meyers collected quantitative data regarding the passing distances between vehicles and bicyclists [8]. This study suggested that where cycle lanes are present, drivers may be driving within the confines of their own designated lane and pay less attention to adjacent bicyclists. Love et al. developed a linear regression model relating vehicle passing distances to quantitative variables like lane width, bicycle infrastructure, cyclist and street identity [9]. Using data from six cities in Ottawa with shared routes and with and without bike lanes, Apasnore et al. formulated a linear regression model [10]. The study found that motor vehicle speed and volume, lane width, number of through lanes, bicycle speed, bicycle position from adjacent curb edge line, adjacent curb parking, and grade slope are significant factors that affect lateral spacing between bicycles and vehicles. A search of current literature suggested no qualification on vehicle operating speed in the presence of bicyclists.

## Performance Measures

Measuring a roadway's performance provides the opportunity to check the quality of a facility's operations. It enables a comparison of the current or anticipated operations to the standards or goals set for that facility. Checks can be performed on a routine basis in order to appraise whether the facility is still performing as desired, or can be part of the facility's design process. Performance measures identified to evaluate the quality of a road include traffic flow, travel time, congestion, travel time reliability, delay, and speed.

In terms of identifying a posted speed limit for a facility, the Manual on Uniform Traffic Control Devices (MUTCD) [11] states, "When a speed limit within a speed zone is posted, it should be within 5 mph of the 85th-percentile speed of free-flowing traffic." Free-flow speeds represent the speed a driver selects when there are no interactions with other vehicles. For a roadway segment, several factors could influence free-flow speed, such as roadside development or road conditions. Ideally, the free-flow speed should be close to the posted speed limit to decrease speed variability and improve the street's performance.

The MUTCD also notes that "speed studies for signalized intersection approaches should be taken outside the influence area of the traffic control signal, which is generally considered to be approximately $1 / 2$ mile, to avoid obtaining skewed results for the $85^{\text {th }}$-percentile speed." Because of the challenges in identifying free-flow vehicles within an urban environment, especially when
signal spacing is less than a signal per mile, other speed measures could be considered as a performance measure.

National Cooperative Highway Research Program (NCHRP) Report 618 [12] and the Moving Ahead for Progress in the $21{ }^{\text {st }}$ Century Act (MAP-21) [13] have encouraged performance-based and result-oriented programs to attain national transportation goals by improving the efficiency of the surface transportation system. The MAP-21 program includes performance requirements to improve air quality, mobility, traffic management, and mitigate congestion by performance evaluations [14].

Common speed-based performance measures include free-flow speed, desired speed, $85^{\text {th }}$ percentile speed, average speed, unconstrained speed, and reference speed. In some cases, these speeds are measured in the field while in others the speed is predicted. The estimates of some predicted speeds use base free-flow speed. The Highway Capacity Manual (HCM) was first published in 1950. Major changes were introduced in HCM 2010 [15] related to level-of-service (LOS) calculation for arterials. Base free-flow speed for arterial street segments considers the effects of speed limit, access density, median type, curb presence, and signals.

Free-flow speed can be explained in many different ways depending on context and adaptability. For example, the HCM says free-flow speed is "the average speed of vehicles on a given segment, measured under low-volume conditions, when drivers are free to drive at their desired speed and are not constrained by the presence of other vehicles or downstream traffic control devices (i.e., traffic signals, roundabouts, or STOP signs" (15, page 4-5). A 2015 study in Delhi found the following factors to have a significant influence on free-flow speed for urban arterials: total vehicles; number of friction points (e.g., bus stops, pedestrian crossings), access points, intersections, and flyovers; and length of the section under consideration [16].

In general, free-flow speed on arterials occurs under low volume conditions when the driver's selection of a speed is unaffected by the presence of other vehicles on the segment with all greens in the trip. Theoretically, free-flow speed is defined as the speed at zero density and flow rate on a given study segment [17], which is hard to obtain in practical conditions. Sekhar et al. considered flow rates between 0 and 1000 passenger cars per hour per lane with vehicles having a lead headway of 8 seconds and lag headway of 5 seconds to be included in the free-flow speed estimation [18]. A method to determine a curve advisory speed suggested that free-flow vehicles are those with at least 3 seconds of space between leading and following vehicles [19].

The desired speed is defined as the speed, under free-flow conditions, that drivers choose to travel when not constrained by roadway design features [20].

Unconstrained speed can be defined as the comfortable speed that a driver would prefer on a given road segment when there is no vehicle interaction. Measuring unconstrained speed for an arterial is difficult due to the presence of signals and multiple driveways. Signal timing typically varies throughout the day to balance the demand on the different approaches and the demand during
different times of the day. Locations with intense peaks, such as schools or offices, can affect travel patterns and operations.

The Guide to Benchmarking Operations Performance Measures discusses definitions for unconstrained travel time as follows:

Unconstrained Travel Time represents a reasonable estimate of travel time in the absence of congestion during good weather conditions. Two different methods of determining unconstrained travel time may optionally be used as the basis for the appropriate performance measures. The first method is preferred: 1. 85th percentile travel time (corresponding to the 85 th percentile speed converted to an equivalent travel time) of traffic during off-peak periods. 2. Target travel time defined as the time it takes motorists to traverse a roadway section when they are traveling at speeds established by operations personnel as the desired speed for a given roadway under prevailing roadway and traffic conditions. Off-peak periods are defined as any time that traffic flow exhibits Level of Service C or better. [21, p. 8]

The 2015 Urban Mobility Report [22] used overnight speeds to approximate the free-flow speeds that were used as a comparison standard. Such an approach may work well for freeway systems but perhaps not for arterials. Other organizations have developed unique approaches to determine reference speeds. For instance, the Florida Department of Transportation (DOT) considers freeflow speed as the average speed where drivers are not controlled by traffic or roadway conditions; free flow speed is assumed to be the posted speed limit plus five [23].

Lomax et al. [24] renamed free-flow speed for arterials as uncongested speed and defined it as the average speed adopted by a driver on a road segment with no vehicle interaction along with given uniform roadway and traffic conditions. Overnight speeds from 10:00 p.m. to 5:00 a.m. were used to identify the free-flow speeds to be used as the comparison standard in evaluating congestion.

The Minnesota DOT [25] uses two reference speeds to evaluate the performance of arterials. One is light traffic speed, defined as the average of the highest two speeds during 14 daytime hours (6:00 a.m. to 8:00 p.m.), while the other is target speed, which is the light traffic speed adjusted by a signal density factor. Zhang and Chen [26] noted that reference speed does not work well for urban arterials with interrupted flows. They examined GPS-based speed data for statewide urban arterials in Kentucky and concluded that it is beneficial to use the daytime data, specifically, the $85^{\text {th }}$ percentile speed, as the reference speed.

## Research Objectives

The objectives of this research were to explore (1) the relationships between suburban vehicle operating speed and the presence of bicyclists and (2) whether crowdsourced speed data could be used to estimate the unconstrained speed for a location. These relationships may be based on the
geometric and traffic control devices characteristics for the area; therefore, those variables were obtained and used in the evaluations.

## Data

The data collection had four main components: downloading the bicycle counts, recording the geometric and traffic control device attributes of the selected sites, collecting the speed-volume data using on-road pneumatic tubes, and obtaining crowdsourced data for a similar time. The research team fused these components during the data reduction efforts.

## Site Selection

On-road tube speed data, crowdsourced data, and roadway characteristics were collected for several locations in the city of San Diego on urban street segments. The locations were selected based upon the availability of bike counter data.

## San Diego Bicycle Data

The San Diego Regional Bike and Pedestrian Counter Network is one of the largest pedestrian and bicycle counting programs in the U.S. It is a collaborative effort between San Diego Association of Governments, San Diego State University, and the County of San Diego Health and Human Services Agency. The network was initially funded by a grant from the Centers for Disease Control and Prevention. A total of 59 counters have been installed across 15 jurisdictions in the San Diego region.

The research team reviewed the locations in San Diego to identify the counters on roadway segments with on-road bicycle lanes. After reviewing the historical data for the on-road counter stations and the counters' battery status, the research team selected eight roads for inclusion in the study. The battery status indicated whether the counter would still be actively collecting data when the vehicle operating speed and volume data would be collected. The chosen sites had a bicycle counter operating 24 hours a day, seven days a week for on-road locations.

## Pneumatic Road Tube Data

On-road tube data collection is a common method for collecting short-term vehicular speed and volume data. A local vendor installed the road tubes at the selected sites and recorded speed and vehicle volume data for most of the sites for the same 2 -week period. Due to technical issues, the vendor collected the speed/volume data for two of the sites after the initial group of sites. Within a $15-\mathrm{min}$ period, vehicle volume was provided and the speed data were binned into groups of 5mph increments. Table 1 lists the data collection periods by site number and Site-Dir (a combination of the site with the direction of traffic). Note that Site 6 and Site 8 represent the same road with different time periods for the bicycle and vehicle data.

Table 1. On-Site Data Collection

| Site Number | Number of <br> Site-Dir | When | Duration |
| :---: | :---: | :---: | :---: |
| Site 1 to Site 5 (both directions) | 11 | $08 / 08 / 2017-$ | 2 weeks |
| Site 8 (only eastbound) |  | $08 / 21 / 2017$ | (total 14,784 15-min bins) |
| Site 6 (both directions) | 2 | $09 / 12 / 2017-$ | 1 week |
|  |  | $09 / 18 / 2017$ | (total 1,344 15-min bins) |
| Site 7 (both directions) | 2 | $11 / 21 / 2017-$ | 2 weeks |
|  |  | $12 / 04 / 2017$ | (total 2,688 15-mins bins) |

## Crowdsourced Data

Several vendors collect crowdsourced data for a wide network of roads throughout the U.S. These roads are split into segments with unique identification numbers. The research team identified the segments that included the on-road tube data and purchased a month of data for those segments. The vendor provided the speed data as a collection of historic speed, travel time data, corridor speed, and the confidence value for each $15-\mathrm{min}$ period. For instance, speed entry at 11:00 a.m. would result in vehicle time and travel speed from 11:00 a.m. to 11:15 a.m. on the given day. Each segment was directional, and segment lengths may have varied in a given direction at a site.

## Geometric and Traffic Control Device Data

Several factors, such as geometric characteristics, traffic control devices, parking activity, and so on, could influence operating speed. The research team collected geometric variables using aerial photographs and used the Google Earth street view function to obtain the posted speed limits. Table 2 summarizes the geometric and traffic control device variables that were influential in the analyses. Two roads (four sites) had $40.3 \mathrm{~km} / \mathrm{hr}(25 \mathrm{mph})$ posted speed limits, and the rest of the sites had $72.5 \mathrm{~km} / \mathrm{hr}(45 \mathrm{mph})$ posted speed limits. The number of through lanes ranged from two to six. The number of signalized intersections within 1.6 km ( 1 mile ) of the counter location ranged from zero to three signals. Driveway densities were above 20 driveways per mile for three sites (Sites 2, 6, and 8). Out of eight sites, five had bus stops. The influence distance represents the distance between the on-road tube counter location and a feature that could influence speed at the counter, typically a signalized intersection. Most of the sites had a raised median, with only Sites 6 and 8 having a two-way left-turn lane. The typical average lane width was $3.4 \mathrm{~m}(11 \mathrm{ft})$ with one road having $4.0 \mathrm{~m}(13 \mathrm{ft})$ average lane widths. All sites had a curb and gutter, level vertical alignment, two-way traffic, and sidewalks on both sides of the street. Additionally, all the selected sites had a bike lane with a pavement marking stripe separating them on at least one side of the street.

To aid in interpreting the modeling results, several variables were adjusted so that the minimum value would be zero. A base condition allows the ability to see the effects of an additional 0.31 meter ( 1 foot) of lane width. Base conditions were also used for other variables, such as the number of signals or the number of driveways.

Table 2. Description of Key Variables*

| Site-Dir | PSL km/h <br> $(\mathbf{m p h})$ | CSL <br> $\mathbf{m}(\mathbf{f t})$ | $\mathbf{L n}$ | Sig | DWS | DWO | ID <br> $\mathbf{m}(\mathbf{f t})$ | $\mathbf{M W}$ <br> $\mathbf{m}(\mathbf{f t})$ | BS | PW <br> $\mathbf{m}(\mathbf{f t})$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SD01-NB | $40.3(25)$ | $225(739)$ | 2 | 1 | 8 | 7 | $67.1(220)$ | $3.1(10)$ | No | $2.4(8)$ |
| SD01-SB | $40.3(25)$ | $225(739)$ | 2 | 1 | 8 | 7 | $39.7(130)$ | $3.1(10)$ | No | $4.3(14)$ |
| SD02-NB | $40.3(25)$ | $1031(3379)$ | 4 | 5 | 20 | 7 | $115.9(380)$ | $1.8(6)$ | Yes | $2.4(8)$ |
| SD02-SB | $40.3(25)$ | $789(2587)$ | 4 | 5 | 7 | 20 | $115.9(380)$ | $1.8(6)$ | Yes | $2.4(8)$ |
| SD03-EB | $72.5(45)$ | $644(2112)$ | 4 | 4 | 1 | 1 | $129.6(425)$ | $4.6(15)$ | Yes | $0(0)$ |
| SD03- <br> WB | $72.5(45)$ | $644(2112)$ | 4 | 4 | 1 | 1 | $103.7(340)$ | $4.6(15)$ | Yes | $0(0)$ |
| SD04-EB | $72.5(45)$ | $644(2112)$ | 6 | 5 | 2 | 3 | $91.5(300)$ | $1.8(6)$ | No | $0(0)$ |
| SD04- <br> WB | $72.5(45)$ | $644(2112)$ | 6 | 5 | 3 | 2 | $115.9(380)$ | $1.8(6)$ | No | $0(0)$ |
| SD05-EB | $72.5(45)$ | $531(1742)$ | 6 | 5 | 2 | 3 | $119(390)$ | $1.8(6)$ | No | $0(0)$ |
| SD05- <br> WB | $72.5(45)$ | $531(1742)$ | 6 | 5 | 3 | 2 | $91.5(300)$ | $1.8(6)$ | No | $0(0)$ |
| SD06-EB | $72.5(45)$ | $1498(4910)$ | 4 | 3 | 22 | 33 | $161.7(530)$ | $3.1(10)$ | Yes | $0(0)$ |
| SD06- <br> WB | $72.5(45)$ | $1498(4910)$ | 4 | 3 | 33 | 22 | $192.2(630)$ | $3.1(10)$ | Yes | $0(0)$ |
| SD07-NB | $72.5(45)$ | NI | 4 | 3 | 8 | 4 | $244(800)$ | $6.1(20)$ | Yes | $0(0)$ |
| SD07-SB | $72.5(45)$ | NI | 4 | 3 | 4 | 8 | $244(800)$ | $6.1(20)$ | Yes | $3.1(10)$ |
| SD08-EB | $72.5(45)$ | $1498(4910)$ | 4 | 3 | 22 | 33 | $161.7(530)$ | $3.1(10)$ | Yes | $0(0)$ |

*where:
BS = Bus stop present (yes or no).
CSL = Segment length for crowdsourced data ( $\mathrm{NI}=$ not included as crowdsource data not available for evaluation), m (ft).
DWO $=$ Number of driveways/unsignalized intersections $0.8 \mathrm{~km}(0.5 \mathrm{mi})$ either side of counter along opposite direction of travel.
DWS $=$ Number of driveways/unsignalized intersections $0.8 \mathrm{~km}(0.5 \mathrm{mi})$ either side of counter along in same direction of travel.
$\mathrm{ID}=$ Distance between counter location and feature that could be influencing the speed measured at the on-road tube counter such as signalized intersection or roundabout, $m(f t)$.
$\mathrm{Ln}=$ Number of through lanes in both directions.
MT = Median type ( $\mathrm{R}=$ raised, $\mathrm{T}=$ two-way left-turn lane).
MW = Typical or average median width for the segment, $\mathrm{m}(\mathrm{ft})$.
PSL = Posted speed limit, km/hr (mph).
$\mathrm{PW}=\mathrm{On}$-street parking width, $\mathrm{m}(\mathrm{ft})$.
Sig = Number of signalized intersections $0.8 \mathrm{~km}(0.5 \mathrm{mi})$ either side of counter, including any signals at the beginning or end of the segment.
Site-Dir $=$ site number along with travel direction for the vehicles measured at the tubes.

## Additional Variables

Based on the literature review along with initial exploratory analyses, several potential factors were identified that could affect the use of crowdsourced data as an estimate of arterial operating speed. These factors are discussed below.

## Level of Service (LOS)

In an urban environment, especially with the presence of signals and the amount of traffic present, vehicles are not able to travel at free-flow speeds for most of the day. To explore whether
crowdsourced speed data are closer to a spot speed during certain traffic conditions, the research team calculated the LOS value for each $15-\mathrm{min}$ period using the tube data.

The research team used the HCM 2010 methodology to predict the base free-flow speed. Base free-flow speed includes an assumed speed constant that is a function of the posted speed limit. It also includes adjustments for median type, cross section (curb or no curb), access density, and segment length. For segments with signalized intersections, the base free-flow speed was adjusted following the directions included in the 2010 HCM [15]. The measured speed for a $15-\mathrm{min}$ period as a percentage of the base free-flow speed was compared to the LOS criteria included in the HCM 2010, and a LOS was assigned to each $15-\mathrm{min}$ period. For example, if the travel speed represented $67-85 \%$ of the free-flow speed, then the $15-\mathrm{min}$ period was operating at LOS B (reasonably free flow).

## Light Level

Previous studies have demonstrated that natural light conditions affect speed [27]; therefore, the light level (day, night, dusk, and dawn) for each $15-\mathrm{min}$ period was identified. The research team collected sunrise and sunset times for each day represented in the dataset and defined dawn as 30 minutes before-and-after sunrise and dusk as 30 minutes before-and-after sunset. Depending on the month and day, dawn times ranged between 6:00 and 7:00 a.m., and dusk was between 6:00 and 8:00 p.m.

## Day of the Week

Preliminary evaluations indicated that speeds varied by the day of the week, and perhaps by whether the day was a weekday or a weekend. The research team examined whether the day of the week could be grouped into three types: weekdays (Monday to Thursday), Friday, and weekend (Saturday and Sunday). The speeds exhibited similar behavior on most of the weekdays with a different trend on Fridays for most of the sites. Saturday and Sunday generally had lower speeds compared to weekdays.

## Crowdsourced Segment Length

The crowdsourced speeds were determined using the travel time from one sensor to the next. The distance between these sensors could influence the speed measurement, primarily depending upon roadway characteristics, such as the number of signals or driveways. Table 2 shows the length of each segment.

## Preparation of the Databases

As discussed in the previous section, the data were fused together to obtain two databases used in the analyses. For the analysis of the relationships between roadway characteristics and operating speed, the collected bicycle data were merged with road tube data using site identification number and time stamps. The road tube data (vehicle speed and volume data) were merged with bicycle data using two separate temporal matching criteria:

- Priority 1: Exact $15-$ min binned temporal matching of the bicycle data and road tube data.
- Priority 2: Due to the unavailability of 2017 bicycle count data at SD01-NB, SD01-SB, and SD02-SB, similar days of the week of the same months of 2016 were matched for 15min bins.

The matched bicycle count data/vehicle speed and count data were fused with the roadway geometry and traffic control device data using site identification number. Out of 18,816 potential $15-\mathrm{min}$ bins, 113 records did not have speed data entries because they occurred late at night or very early in the morning when no vehicles were present within the $15-\mathrm{min}$ period. These records were removed from the final analysis. This dataset contains $18,70315-\mathrm{min}$ binned speed, traffic volume, and bicycle volume data, along with related geometrics/traffic control device information.

For the analysis of how crowdsourced data could be used to estimate uncongested speed, the research team matched the on-road tube data and the crowdsourced data by site using date and start time for the time increment and the geometric and traffic control device data were matched using the site number. For the crowdsourced data analysis, data for Sites 1, 2, 4, 5, and 8 were available, providing a total of $12,09615-\mathrm{min}$ periods for evaluation. The research team then eliminated the $15-\mathrm{min}$ periods where no vehicles were recorded ( 68 records, most of which occurred at night or early morning on a Sunday), thus resulting in a total of 12,028 15-min periods.

## Results: Influences on Vehicle Operating Speed

## Exploratory Data Analysis

Figure 1 shows the relationships between three key measures (average vehicle speed, vehicle volume per lane, and bicycle volume) by time of day for different posted speed limits. Unsurprisingly, bicycle volumes were higher during the daytime. For roadways with $40.3 \mathrm{~km} / \mathrm{hr}$ ( 25 mph ) posted speed limits, bicyclist volume peaked from 9:00-11:00 a.m., with a smaller peak occurring between 6:00 and 7:00 p.m. A decrease and increase in average speeds can be seen mirroring the increase and decrease in bike volume (i.e. speeds begin dropping at 7:00 a.m., which is when bike volume increased, and speeds appear to increase at 7:00 p.m., which is about the same time as a noticeable drop in bicycle volumes took place). This pattern is not as apparent in the morning peak hours for the roadways with $72.5 \mathrm{~km} / \mathrm{hr}(45 \mathrm{mph})$ posted speed limits. For the evening peak (6:00-7:00 p.m.), lower average vehicle speeds were visible. It is important to note that vehicle operating speed can be associated with vehicle volumes per lane, with lower speeds present during high vehicle volumes. For roadways with $40.3 \mathrm{~km} / \mathrm{hr}(25 \mathrm{mph})$ posted speed limit, higher vehicle volumes were visible at evening peak hours. For both the morning peak and evening peak hours, vehicle volumes were visibly higher for roadways with $72.5 \mathrm{~km} / \mathrm{hr}$ ( 45 mph ) speed limits.


Figure 1. Relationship between three key measures for different time of day.
Figure 2 shows a series of violin and box plots, which illustrate additional details of the three key measures per location. Because $15-\mathrm{min}$ binned bicycle volumes contained a significant number of zero counts, the violin and box plots were generated for hourly volume (both bicycle and vehicle) counts instead of $15-\mathrm{min}$ binned counts. The violin plot shows the kernel density estimation of the underlying distribution of the data with the height of the violin reflecting the sample size for the given speed. Within the violin plot is the box-plot showing the median and interquartile range of the data. Two colors, along with grouping the data, were used in Figure 2 to provide an easier identification of the Site-Dir with 25 mph posted speed limits (blue-green color, or the 11 top SiteDir in the plots) and 45 mph posted speed limits (red color, or the bottom 4 Site-Dir in the plots).

The median speeds for all segments show lower average speeds than the posted speed limits, which is not surprising due to the higher volumes on urban arterial roadways. The roadway segments with $40.3 \mathrm{~km} / \mathrm{hr}(25 \mathrm{mph})$ posted speed limits show a higher median bicycle volume compared to the roadway segments with $72.5 \mathrm{~km} / \mathrm{hr}(45 \mathrm{mph})$ posted speed limits.

The research team also obtained National Oceanic and Atmospheric Administration hourly precipitation data to observe the effect of precipitation and found that there was no major influence on the data included in the dataset.


Figure 2. Violin and box plot distribution of three key measures per location ID.
Several evaluations were conducted to explore the speed-volume relationships. The first used all available $15-\mathrm{min}$ periods with on-road speed data when the vehicle count was not zero. This dataset had $18,70315-\mathrm{min}$ periods. As the majority of the timestamps contained zero bicycle counts, the next evaluation used $15-$ min periods when bicycle counts were available and reflected the two available posted speed limits. Additional filters for this evaluation were that only daytime data were included and bicycle volume per $15-\mathrm{min}$ binned was greater than one. This final dataset contained 2,435 $15-\mathrm{min}$ periods for analysis ( 1,622 for 25 mph roadways, and 813 for 45 mph roadways). After performing a preliminary analysis, it was found that effect of the number of signalized intersections was unpredictable and random and this variable was thus removed from the evaluations.

## Model Development

The linear regression model is one of the most common modeling techniques used in transportation engineering. This model is applied when there is a linear association between the response variable and exploratory variables. In a linear model, the exploratory variables are considered as fixedeffect variables. In the real world, many relationships are not systematic like linear model assumptions. By considering a random effect in the modeling framework, it is possible to characterize the distinctive variation due to the individual differences. Random effects are usually non-systematic and unpredictable, and have random influence on the data. Consideration of both fixed and random effects in the modeling framework is known as mixed effect modeling. Several linear mixed models were developed, some focusing on using all available speed data and others focusing on data where more than two bicyclists were present within the $15-\mathrm{min}$ period. Other attempted modeling techniques included splitting the data by the different posted speeds.

The importance of the variables was ranked using random forest algorithms. The random forest method is based on the bagging principle and random subspace method, which relies on constructing a collection of decision trees with random predictors. Variable importance ranking is measured by classification accuracy and Gini impurity. This importance measure shows how much the mean squared error or the "impurity" increases when the specified variable is randomly permuted. If prediction error does not change when permuting the variable then the importance measures will not be altered significantly, which in turn will change the mean squared error (MSE) of the variable only slightly (low values). This implies that the specified variable is not important. On the contrary, if the MSE significantly decreases during the permutation of the variable then the variable is deemed important.

By keeping 15-min binned average speed as the dependent variable, the random forest algorithm was applied in developing the variable importance measures. Figure 3 illustrates the dot chart of variable importance as measured by the algorithm. The more a variable contributed, the more significant it was. A threshold of $5 \%$ MSE was applied in determining the final list of variables. The variable importance measure shows that the vehicle volume per lane (VVolPerLn) was the most influential variable for all three cases examined. Other variables considered included time of day (day/night/dawn/dusk), posted speed limit, distance between counter location and feature that could be influencing the on-road speed measurement adjusted to base condition (ID_BC), number of driveways along the same side of the road as the travel direction adjusted to base condition (DWS_BC), number of driveways along the opposite side of the road as the travel direction adjusted to base condition (DWO_BC), bicycle volume (BikeVolume), presence of on-street parking (Park), on-street parking width (PW), number of signals (SigInter), presence of bus stop (BusStop), presence of raised median (MED), median width adjusted to base condition (MW_BC), and number of through lanes (ThruLns).

a) all 15-minute periods with on-road speed data

b) 15-min periods with on-road daytime speed data, bicycle count greater than 1 , and $40.3 \mathrm{~km} / \mathrm{hr}$ [ 25 mph ) roadways

Figure 3. Variable importance plots.

## Model Results

## Model Based on All Available On-Road Speed Data

The top nine variables are shown in (Figure 3). These variables were considered for primary analysis to identify variables influencing operating speed on an arterial. Table 3 shows two models, with Model 2 removing the number of signals since that variable was not found to be significant in Model 1. The variables found to influence operating speed included the number of vehicles within the $15-\mathrm{min}$ period, the distance from the counter to a signalized intersection or a roundabout, the number of driveways (access density), and the posted speed limit. Nighttime speeds were not significantly different from daytime speeds. While dawn and dusk speeds were significantly different from daytime speed, the amount of difference was small and could be considered not of practical difference.

## Models Focusing on When Bicycle Counts are Available

The next set of models used data when more than one bicyclist was present within a $15-\mathrm{min}$ period and the light condition was daytime. The research team attempted to develop models using both posted speed limits and then by each posted speed limit. The attempts using the $72.5 \mathrm{~km} / \mathrm{hr}$ ( 45 mph ) speed limit data resulted in models where bicycle volume was not significant. This is not surprising given that the range of bicycle volume available for the higher-speed roads is limited; for most of the sites, fewer than five bicyclists were observed in an hour (see Figure 2).

The model using data for the $40.3 \mathrm{~km} / \mathrm{hr}(25 \mathrm{mph})$ roadways demonstrates that vehicle volumes per lane and bicycle volume are statistically significant and are negatively associated with vehicle operating speed (see Table 4). An increase of 19 motor vehicles per 15 -min binned results in a decrease in average speed of $1.6 \mathrm{~km} / \mathrm{hr}(1.0 \mathrm{mph})$. For bicyclists to have a similar impact on operating speed, the model indicates that more than 39 bicyclists per $15-\mathrm{min}$ period would be
needed. For this dataset, a maximum of 41 bicyclists was observed in a 15 -min period; therefore, it was not possible to determine whether a larger number of bicyclists would have a greater influence on operating speed within this study. An increase of $30.5 \mathrm{~m}(100-\mathrm{ft})$ influence distance will increase average speed by $5.0 \mathrm{~km} / \mathrm{hr}(3.1 \mathrm{mph})$. Presence of a bus stop was associated with a speed decrease of $6.4 \mathrm{~km} / \mathrm{hr}(4.0 \mathrm{mph})$. Driveway density was not significant for this model.

Table 3. Model Outputs for 15-min Periods with On-Road Speed Data

| Fixed Effect Variable | Model 1 <br> Estimate | Model <br> $\mathbf{1 ~ S t d .}$ <br> Err. | Model <br> $\mathbf{1}$ t-stat | Model <br> $\mathbf{1}$ p- <br> value | Model 2 <br> Estimate | Model <br> $\mathbf{2}$ Std. <br> Err. | Model <br> $\mathbf{2 ~ t - s t a t ~}$ | Model <br> $\mathbf{2 ~ p - ~}$ <br> value |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Intercept | 19.7490 | 3.2862 | 6.0097 | 0.0000 | 21.0073 | 1.6141 | 13.0146 | 0.0000 |
| Vehicle volume per lane | -0.0477 | 0.0007 | -68.076 | 0.0000 | -0.0477 | 0.0007 | -68.077 | 0.0000 |
| Influence distance-base <br> condition (ft) | 0.0225 | 0.0005 | 42.5667 | 0.0000 | 0.0225 | 0.0005 | 42.5696 | 0.0000 |
| Driveway density same <br> direction-base condition | -0.1646 | 0.0375 | -4.3897 | 0.0000 | -0.1484 | 0.0057 | -25.933 | 0.0000 |
| Driveway density opposite <br> direction-base condition | -0.0844 | 0.0373 | -2.2622 | 0.0237 | -0.0682 | 0.0042 | -16.255 | 0.0000 |
| Number of signalized | 0.5020 | 1.1496 | 0.4367 | 0.6623 | -- | -- | -- | -- |
| intersections |  | 0.0574 | 0.0587 | 0.9781 | 0.3280 | 0.0573 | 0.0587 | 0.9772 |
| Night | 0.4621 | 0.1035 | 4.4662 | 0.0000 | 0.4621 | 0.1035 | 4.4658 | 0.3285 |
| Dawn | -0.2281 | 0.0959 | -2.3791 | 0.0174 | -0.2281 | 0.0959 | -2.3792 | 0.0174 |
| Dusk | 6.8558 | 0.8194 | 8.3670 | 0.0000 | 7.2117 | 0.0772 | 93.4674 | 0.0000 |

Notes:

- Random effect variable:
- Median width-base condition: Model 1 intercept $=9.86$ (3.14), Model 2 intercept $=10.33(3.21)$.
- Number of through lanes: Model 1 intercept $=0.05$ ( 0.01 ), Model 2 intercept $=0.04(0.01)$.
- Model 1 Residual $=6.58$ (2.57), Model 2 Residual $=6.58$ (2.57).
- Model 1 AIC $=88,374.9$, BIC $=88,476.7$, Log likelihood $=-44,174.4$, Deviance $=88,348.9$.
- Model 2 AIC $=88,373.1$, BIC $=88,467.1$, Log likelihood $=-44,174.5$, Deviance $=88,348.1$.

Table 4. Model Outputs for 15-min Periods during Daylight Conditions and More Than One Bicyclist on 40.3 km/hr (25 mph) Roadways

| Fixed Effect Variable | Estimate | Std. Err. | t-stat | p-value |
| :---: | :---: | :---: | :---: | :---: |
| Intercept | 22.2879 | 0.2361 | 94.4009 | 0.0000 |
| Vehicle volume per lane | -0.0523 | 0.0015 | -34.5227 | 0.0000 |
| Bicycle volume | -0.0254 | 0.0101 | -2.5099 | 0.0121 |
| Influence distance-base condition (ft) | 0.0311 | 0.0016 | 19.4409 | 0.0000 |
| BusStop=Yes | -4.0076 | 0.3376 | -11.8725 | 0.0000 |

Notes:

- Random effect variable: number of signalized intersections, intercept 0.23 (0.01), residual 2.98 (1.73).
- Model AIC $=6390.7$, BIC $=6433.8, \log$ likelihood $=-3,187.3$, Deviance $=6,374.7$.


## Results: Crowdsourced Data

## Exploratory Data Analysis

The preliminary analysis involved visual inspection of patterns. A review of the distribution of average $15-\mathrm{min}$ speed by vehicle volume shows that the expected pattern of average speed was lower for higher vehicle volumes. The speed range for each Site-Dir is shown in Figure 4, which illustrates the difference in speeds based upon how the speed data were collected (i.e., using tubes or crowdsourced).

To explore the relationship between tube speed data and crowdsourced speed data, the research team calculated a new variable, TMCS, which is the difference between tube and crowdsourced speeds. Figure 5 shows the variation of TMCS based on the day of the week. In most cases, TMCS was positive, indicating that the speeds reported in the tube speed data were greater than those reported in the crowdsourced data.


白TubeSpeed追CrowdSourcedSpeed

Figure 4. Box plot for tube and crowdsourced speeds at each Site-Dir.


Figure 5. Speed difference variation per Site-Dir for each day.

## Operational Analysis

The evaluation used a decision tree approach to determine which variables were most associated with small differences between tube speed data and crowdsourced speed data (i.e., TMCS). A decision tree has the appearance of a flow chart, where each node represents the test on the variable (e.g., number of signals $\leq 4$ ) and the branches represent the outcome of that test. In the number of signals example, the branches would show the mean TMCS value for "yes" (the number of signals is $\leq 4$ ) or "no" (the number of signals is $>4$ ).

The research team adopted a rule fit method containing ranges of the factors, which used "rule generation" and "rule pruning" algorithms to obtain hidden rules, for predicting the importance of selected variables from a pool. The rule-based analysis can handle large dataset with a mixture of variables, ranging from numeric to categorical and resulting in easy interpretation. Rule-based decision trees were derived for the database by considering all parameters as intuitive parameters, thus eliminating less significant variables.

The research team used different subsets of the available variables to explore the characteristics that might affect the results. This work provided the opportunity to identify whether a subset of variables might be considered appropriate for estimating arterial performance. Several approaches were explored as part of the research; this report documents the two approaches with the most interesting findings. The first approach focused on roadway characteristics. The variables considered included number of signals, number of through lanes, the distance between the tubes and nearest signalized or roundabout intersection, median width, number of driveways, posted speed limit, and segment length. The other approach focused on uncongested periods-LOS A, B, and C. The Highway Capacity Manual [15] defines these uncongested conditions as follows: LOS A is free flow of traffic, LOS B is reasonably free flow of traffic, and LOS C is stable flow or at near free flow.

Figure 6 shows the results when only considering geometric variables. The analysis revealed that the most influential variable was the number of signals. Smaller differences between the crowdsourced speeds and the on-road tube speeds were present when there were few signals within the corridor; in this dataset, fewer than four signals within the corridor had better matches. Signals are associated with large disruptions in travel and can introduce delay within the corridor; therefore, it is logical that the number of signals needs to be considered when using crowdsourced data as a representative speed for a specific location. The next influential geometric variable was the number of driveways. Similar to signals, driveways can introduce conflicts within the travel stream. A vehicle slowing to turn right at a driveway can increase the travel time for other vehicles. Further, while a vehicle turning out of a driveway should wait for an adequate gap, experience has shown that drivers may turn onto the major roadway expecting drivers on that roadway to slow to avoid a crash. The more driveways, the greater the likelihood that delay will be added to the time to travel from one end of the corridor to the other, potentially resulting in a greater difference between crowdsourced speed data and speed data collected at a specific location.


Figure 6. Decision tree for speed difference using only geometric variables.
The decision tree shown in Figure 7 focuses on temporal variables, such as the day of week and the light level, with the added restriction of only including the 15 -min periods when conditions were at LOS A, B, or C (as determined from the tube data). This analysis addressed the question of which conditions could be used to approximate the speed during uncongested conditions. The analysis found the smallest difference between the crowdsourced speed and tube speed to be on Saturday and Sunday.


Figure 7. Decision tree for speed difference using temporal variables when only considering those 15-min periods when conditions are $\operatorname{LOS} A, B$, or $C$.

## Conclusions and Recommendations

## Bike Count Data

As part of this effort, the research team explored the relationship between vehicle operating speeds with urban arterial street characteristics and bicycle count data using linear mixed-effect modeling. Both vehicle volume per lane and bicycle volume were found to influence average speed on lower speed urban arterial roadways. For $40.3 \mathrm{~km} / \mathrm{hr}(25 \mathrm{mph})$ sites, an increase of 19 motor vehicles per $15-\mathrm{min}$ bin decreased average speed by $1.6 \mathrm{~km} / \mathrm{hr}(1.0 \mathrm{mph})$. For bicyclists to have a similar effect on operating speed, the model indicates that more than 39 bicyclists per 15-min period would be needed. For this dataset, a maximum of 41 bicyclists were observed in a $15-\mathrm{min}$ period; therefore, this study was unable to determine whether larger numbers of bicyclists might have a greater influence on operating speed.

Because of the limited number of $15-$ min periods with bicycle counts greater than 1, the research team also developed a model using all available $15-\mathrm{min}$ periods with on-road speed data. The variables found to influence operating speed included vehicle volume per lane, distance between counter location and feature that could affect speed (typically a signalized intersection), median width, number of driveways and number of signals $0.8 \mathrm{~km}(0.5 \mathrm{mi})$ on either side of the counter location, posted speed limit, and light condition (day, night, dawn, dusk).

## Crowdsourced Data

The objective of evaluating the crowdsourced data was to identify whether there were specific conditions when crowdsourced speeds on an arterial could appropriately reflect the speed at a specific location. Crowdsourced speeds reflect a corridor speed and may need adjustments to be able to adequately reflect the speed at a specific location within the corridor. Speed and volume data in $15-\mathrm{min}$ increments for 2 weeks at nine sites were obtained using on-road tubes and via a vendor of crowdsourced speed data. These data streams were fused using the site number, date, and time for each of the $15-\mathrm{min}$ periods. The difference between the tube data and the crowdsourced data was calculated and called TMCS. The research team conducted the analysis using decision trees to identify those variables most associated with the smallest values of TMCS.

The geometric variables that had the greatest influence on TMCS were the number of signals and the number of driveways within a corridor. Those corridors with a smaller number of signals (<4) and a smaller number of driveways ( $<13$ ) were associated with smaller values of TMCS. When focusing on variables associated with a specific $15-\mathrm{min}$ period, congested periods (LOS D or E; approaching unstable flow or unstable flow, respectively) were associated with the smallest TMCS values. When only including non-congested periods (i.e., LOS A, B, or C), weekends (Saturday or Sunday) were associated with the smallest TMCS.

## Challenges and Future Research Needs

The success of this research effort relied on several components, including the availability of the needed data along with the ability to fuse the data streams. While bicyclists were counted in several locations, many were on trails, which was not the focus of this study. When the counters were on an arterial street, they were frequently located near a signalized intersection or a roundabout, which affects vehicle operating speed. Sites were selected to maximize the number of bicyclists; however, a smaller range of bicycle volume was available than desired, especially for higher-speed streets. Another challenge was that even though the selected sites had sufficient battery life (indicating that future bicycle counts are possible), three sites did not have bicycle data to match to the onroad speed measurements conducted after the sites were selected. Previous years' bicycle counts had to be considered at those sites.

With the growing number of cities adding a bicycle count program, future studies may be able to assemble a larger sample size with a wider variety of arterial roadway features. Another approach could be to include additional measures (bicycle speed and lateral distances between vehicles and bicyclists) to determine the robust quantification of average vehicle speeds in presence of the bicyclists.

With respect to the crowdsourced study, the crowdsourced data included several empty cells. It also did not include vehicle volume for the given time period, which is a significant limitation for this type of research. Future studies with a larger sample size of sites and tube data with individual
vehicle speeds could provide better estimates of the relationship between spot speeds and corridor speeds as provided by crowdsourced data.

## Additional Products

The Safe-D website with information about this research is at: https://www.vtti.vt.edu/utc/safe-d/index.php/projects/influences-on-bicyclists-and-motor-vehicles-operating-speed-within-acorridor/.

## Education and Workforce Development Products

These efforts included the following:

- Provided the Texas A\&M University Civil Engineering professors the results from this research for incorporation into their courses, as appropriate.
- Presented the findings from the research at the Institute of Transportation Engineers Joint Meeting of the Western District and the Texas District in June 2018.
- Three students assisted with this research: Manaswini Condor (graduate student, but did not use topic for engineering paper), Marie Connie Rodriquez (undergraduate student), and Elizabeth Clark (summer intern).


## Technology Transfer Products

The following papers were generated:
Das, S., K. Fitzpatrick, M. C. (2018) "Effects of Bicyclists on Vehicle Operating Speed: A Study on urban Arterial Roadways," ITE Joint Meeting of the Western District and the Texas District.

Fitzpatrick, K., S. Das. (2019 anticipated) "Using Crowdsourced Data to Estimate Operating Speed on Suburban Arterials," Submitted for consideration for the publication in the IATSS Research journal.

## Data Products

While similar data were used in both analyses, there were cases where data were only available for certain conditions; therefore, the research team created two databases that matched the two objectives.

- The Bike-Veh AnG7alysis database [https://doi.org/10.15787/VTT1/KCNRS7] is available via the Safe-D Collection on the VTTI Dataverse. It includes vehicle data, bicycle count data, temporal variables associated with the $15-\mathrm{min}$ period such as time (hour), and site characteristics for 15 sites.
- The Crowd Sourced Analysis database [https://doi.org/10.15787/VTT1/WXR1I0] is available via the Safe-D Collection on the VTTI Dataverse. It includes vehicle, temporal, and site characteristics data for nine sites.


## References

1 United States Department of Transportation. The National Walking and Bicycling Study: 15-Year Status Report. U.S. Department of Transportation, Federal Highway Administration, Washington, DC, 2010.

2 National Association of City Transportation Officials. Cities for Cycling. National Association of City Transportation Officials, Washington, DC, 2010.

3 McKenzie, B. Modes Less Traveled-Bicycling and Walking to Work in the United States: 2008-2012. U.S. Department of Commerce, 2014.

4 Elvik, R. The non-linearity of risk and the promotion of environmentally sustainable transport. Accid. Anal. Prev. 41, 849-855, 2009.

5 Fitzpatrick, K., S. Das, T. Gates, K. Dixon, E.S. Park, R. Avelar. Guidance for the Setting of Speed Limits. NCHRP 17-76 Draft Report, Washington, D.C. 2018.

6 Li, Z., W. Wang, P. Liu, and D. Ragland. Physical environments influencing bicyclists’ perception of comfort on separated and on-street bicycle facilities. Transportation Research Part D: Transport and Environment, 17(3), 256-261, 2012.

7 Chen, L., C. Chen, R. Srinivasan, E. McKnight, R. Ewing, M. Roe. Evaluating the safety effects of bicycle lanes in New York City. Am. J. Public Health 6, 1120-1127, 2012.

8 Parkin, J., C. Meyers, The effect of cycle lanes on the proximity between motor traffic and cycle traffic, Accident Analysis \& Prevention, 42(1),159-165, 2010.

9 Love, C., A. Breaud, S. Burns, J. Margulies, M. Romano, R. Lawrence. Is the three-foot bicycle passing law working in Baltimore, Maryland? Accident Analysis \& Prevention, 48, 451456, 2012.

10 Apasnore, P., K. Ismail, A. Kassim, Bicycle-vehicle interactions at mid-sections of mixed traffic streets: Examining passing distance and bicycle comfort perception. Accid. Ana. Prev., 106, 141-148, 2017.

11 Federal Highway Administration. Manual on Uniform Traffic Control Devices for Streets and Highways http://mutcd.fhwa.dot.gov. 2009.

12 Cambridge Systematics, Dowling Associates, System Metrics Group, Inc., and Texas Transportation Institute. NCHRP Report 618. Washington, DC. 2008.

13 Federal Highway Administration. "Moving Ahead for Progress in the 21st Century (MAP21), Performance Management," Washington, DC: US Department of Transportation. 2012.

14 Federal Highway Administration. "Moving Ahead for Progress in the 21st Century (MAP21). Congestion Mitigation and Air Quality Improvement Program (CMAQ)," Washington DC: US Department of Transportation. 2012.

15 Highway Capacity Manual 2010, Transportation Research Board.
16 Rao, A. M., and K. R. Rao. "Free Speed Modeling for Urban Arterials - A Case Study on Delhi," Periodica Polytechnica Transportation Engineering, 43(3), 111-119. 2015.

17 May, A. Traffic Flow Fundamentals. Prentice Hall, Englewood Cliffs, NJ. 1990.
18 Sekhar, C. R., J. Nataraju, S. Velmurugan, P. Kumar, K. Sitaramanjaneyulu. "Free-flow speed Analysis of Two Lane Inter Urban Highways," Transportation Research Procedia, 11th Transportation Planning and Implementation Methodologies for Developing Countries, TPMDC, 17, 664-673. 2016.

19 Bonneson, J., M. Pratt, J. Miles, and P. Carlson. Horizontal Curve Signing Handbook, Report No. FHWA/TX-07/0-5439-P1. Available at: http://tti.tamu.edu/documents/0-5439-P1.pdf. 2007.

20 Fitzpatrick, K., L. Elefteriadou, D. Harwood, J. Collins, J. McFadden, I. Anderson, R. Krammes, N. Irizarry, K. Parma, K. Bauer and K. Passetti. Speed Prediction for Two Lane Rural Highways. Federal Highway Administration Report FHWA-RD-99-171, Washington, D.C: US Department of Transportation. 2000.

21 Tarnoff, P. J., S. E. Young, J. Crunkleton, N. Nezamuddin. Guide to Benchmarking Operations Performance Measures. NCHRP Preliminary Draft Final Report. Available at: http://www.catt.umd.edu/sites/default/files/documents/final_report_compiled_v26.pdf. 2008.

22 Lomax, T., et al "Appendix A: Methodology for the 2015 Urban Mobility Scorecard". Texas A\&M Transportation Institute. Available at: https://tti.tamu.edu/documents/mobility-scorecard-2015-appx-a.pdf. Accessed on: July 21, 2018. 2015.

Florida Department of Transportation, Transportation Statistic Office. "Multimodal Mobility Performance Measures Source Book," Florida Department of Transportation. Tallahassee, Florida. 2015.

24 Lomax, T., S. Turner, B. Eisele, D. Schrank, L. Geng, and B. Shollar. "Refining the RealTimed Urban Mobility Report," UTCM 11-06-73. 2012.

25 Turner, S., D. Schrank, and L. Geng. "Twin City Metropolitan and MN Statewide IRC Arterial Travel Time Analysis: Use of Private Sector Traffic Speed Data". Minnesota Department of Transportation. 2011.

26 Wang, C, and J. B. Huegy. "Determining the Free-Flow Speeds in a Regional Travel Demand Model Based on the Highway Capacity Manual," Presented at Transportation Research Board Annual Meeting. 2014.

27 Fitzpatrick, K., K. Dixon, and R. Avelar. "Evaluating Operational Implications of Reduced Lane and Shoulder Widths on Freeways," ASCE J. Transp. Eng. 142(11), DOI 10.1061/(ASCE)TE.1943-5436.0000884, 04016052. 2016.

