Journal of Modern Mobility Systems

VOLUME 01 (2020)

Journal Home:

https://journals.gmu.edu/index.php/jmms

Document link

https://journals.gmu.edu/index.php/jmms/issue/view/230

Journal of Modern Mobility Systems (JMMS) publishes peer-reviewed research briefs (1700 words or less) with a targeted turnaround time of three months or less. To merit publication in JMMS, articles should highlight a significant new discovery or an innovative methodology that is of interest to the community of transportation practitioners and researchers at large in areas related to modern mobility systems.

Mason Journals
***by* Mason Publishing Group**
(A division of George Mason University Libraries)

This page was intentionally left blank.

Journal Home: https://journals.gmu.edu/index.php/jmms
Mason Journals ● Mason Publishing Group
Journal of Modern Mobility Systems

About the Journal

The field of transportation is experiencing revolutionary changes due to rapid innovations in technology and operations. For example, new vehicle technologies such as connected and autonomous vehicles are changing the dynamics of infrastructure planning, design, construction, operations and maintenance. Shared mobility services (ridesharing, micromobility, ride-hailing) are reshaping travel demand, travel behavior and the economics of transportation. Practitioners and researchers who are dealing with these rapid developments could benefit from quick access to the latest research. However, typical turnaround times for publishing full-length research articles in major journals, though improved from a decade ago, have not kept pace with recent rapid innovations. Journal of Modern Mobility Systems (JMMS) fills this gap by publishing peer-reviewed research briefs (1700 words or less[1]) in a timely manner. Turnaround time for JMMS articles is targeted to be 3 months or less. To merit publication in JMMS, the submitted articles should highlight a significant new discovery or an innovative methodology that is of interest to the community of transportation practitioners and researchers at large in areas related to modern mobility systems. The journal scope includes a broad range of research topics in transportation policy, planning, systems analysis, engineering, technological innovations and societal impact related to MMS.

If you want to submit a paper to this journal, MS Word Template for the paper can be downloaded from the journal website.

Journal details:

- ISSN: 2644-2388
- Publisher: Mason Publishing (George Mason University)
- Distribution: On-line only
- Publication Frequency: Immediately after acceptance. Archival is currently done annually.
- Indexing: Currently indexed by CrossRef and Google Scholar. Will be indexed in DOAJ and other premier indexing avenues as soon as the eligibility criteria are met.
- Article processing fee: Currently there is no fee to publish in JMMS

Sponsors[2]:

- Sid and Reva Dewberry Department of Civil, Environmental and Infrastructure Engineering (CEIE), George Mason University
- Association of Transportation Professionals of Indian Origin (ATPIO)

[1] The limit of 1700 words was adapted from the now discontinued extended abstract format used for TRB annual meeting proceedings. Length exceptions may be made for some papers with proper justification. In no case the length of the abstract and paper should exceed 200 and 2000 words, respectively.

[2] Sponsors provide financial support to bear the costs associated with publishing JMMS so that the authors and readers are not charged. We welcome all sponsors - individuals, not-for profit entities or corporations. Your sponsorship to the journal may be federal tax deductible. For sponsorship enquiries, please contact Mohan Venigalla, Chief Editor at mvenigal@gmu.edu.

Volume 01 (2020)
https://journals.gmu.edu/index.php/jmms/issue/view/230

TABLE OF CONTENTS

This page was intentionally left blank.

Editors' Note

Welcome to the Inaugural Issue of JMMS

The overarching goal of Journal of Modern Mobility Systems (JMMS) is to provide a high-quality venue to display time-critical research on a public venue that can positively impact society. Because technology and its impact on transportation systems are evolving at a faster pace in the recent years, the primary objective of (JMMS) is to facilitate the publication of cutting-edge, peer-reviewed research works in a timely manner. Specifically, JMMS strives to obviate the need for researchers to await a six- to twelve-month turn around time for presentation at a major conference or publication in a leading journal. From our perspective, it is also important to assure that authors retain the rights to their work while parts of their research results are disseminated quickly through JMMS.

Given the usual difficulties with the launch of a new journal that aspires to be at the forefront of research, especially during a disruptive pandemic, it has been a bit of a slow start for JMMS. However, the JMMS activities gained momentum in the later part of the year with a marked rise in quality articles submitted for review. We are pleased to release Volume 01(2020) of JMMS with five high quality research briefs.

The need to have transportation research work recognized and placed in a public forum was never more apparent than in 2020. We launched JMMS in a year that will forever be known for the outbreak of the COVID-19 pandemic. As such, a core theme of this journal is to disseminate early the research findings related to the disruptive forces in transportation that are reshaping travel, travel modes and travel demand worldwide. Worldwide, the pandemic infected over 82 million people, caused more than 1.8 million deaths, and caused economic hardships to businesses and people. Most relevant to this journal, the pandemic has disrupted both traditional and disruptive forces in the world of transportation. For example, as remote work has become one of the effective mechanisms to combat the spread of the virus, airline travel, transit ridership, traffic volumes have been down significantly. People have found alternative ways to get around. E-commerce has exploded that resulted in a marked increase in delivery trips. The share of travel by bicycling mode has increased to the point that there are reports of bicycle shortages.

Given the dramatic short-term and potentially long-term changes in the transportation systems that are direct result of the COVID-19 pandemic, transportation researchers will be faced with numerous challenges in evaluating the systems and developing solutions. With these changes, new research is being conducted that needs to be made public through reliable journals quickly.

Thus, we are encouraging submissions to JMMS on research works that are related to the broad areas of transportation systems with focus on modern innovations on mobility. For articles that merit publication, there will be a maximum of a three month turn around until publication. For 2021, we are especially interested in works that clearly outline new and innovative transportation research associated with the impacts of COVID-19 and potentially with the post-pandemic dynamics of travel and transportation. We will also consider reviewing articles that were submitted to other leading avenues but may not have met the cut for publication in the associated journals (e.g., Transportation Research Board). If those articles fit the scope of JMMS and have received very good reviews in prior round, we strongly encourage you to submit them to JMMS for expedited review.

We sincerely thank the Dean of Volgenau College of Computing and Engineering, the Dean of Libraries, the Provost and the President of George Mason University for their support and encouragement in launching JMMS. We also thank our sponsor ATPIO (http://ATPIO.org) for helping us deliver a highly professional product.

We are anticipating and looking forward to a robust year ahead for JMMS.

Mohan Venigalla
Thomas Brennan

Co-Editors in Chief
Journal of Modern Mobility Systems

This page was intentionally left blank.

Journal Home: https://journals.gmu.edu/index.php/jmms
Mason Journals ● Mason Publishing Group
Journal of Modern Mobility Systems (JMMS), Vol. 01(2020), 101-109
Document: https://doi.org/10.13021/jmms.2020.2736

Characterization of the Coronavirus Pandemic on Signalized Intersections Using Probe Vehicle Data

Bryan Remache-Patino[a], Thomas Brennan[a, *]

[a] *The College of New Jersey, Armstrong Hall - RM 173, 2000 Pennington RD, Ewing Township, NJ 08628; *Corresponding Author*

Received: 6 June 2020 ● **Revised**: 23 June 2020 ● **Accepted**: 26 June 2020 ● **Published**: 29 June 2020

ABSTRACT

The Coronavirus (COVID-19) pandemic presents a unique opportunity to establish a baseline for studying transportation performance metrics before, during, and the eventual transition to normalcy using probe vehicle data. Probe vehicle speed data are already being used to evaluate traffic congestion characteristics, resiliency, and network response at local, corridor, and regional levels. A better understanding of changes in traffic characteristics, 24-hours a day, 7 days a week, can be realized through the analysis of spatially located, temporal speed data. This paper explores the use of probe vehicle data sets to establish the baseline traffic conditions under the unique conditions resulting from COVID-19. The preliminary research analysed about 500,000 speed records over a 21-week period at two intersections in Northern New Jersey to numerically and visually characterize the speed patterns through the COVID-19 progression. Although further research and statistical analysis is necessary to evaluate the data as it relates to the New Jersey State pandemic and emergency management policies, the preliminary results indicate school closures and the stay-at-home order have significantly impacted normal traffic and thus present a unique research opportunity to study baseline, non-congested conditions.

Keywords: roads & highways, traffic management, infrastructure planning, COVID-19, travel speed

1. INTRODUCTION

Under the Operations Performance Measurement Program [*1*] of the United States Department of Transportation (USDOT), federal, state, and local agencies have been increasingly using the National Performance Management Research Data Set (NPMRDS) to evaluate road performance as well as external factors that impact roadway performance [*2*]. Probe Vehicle Data (PVD), which is part of the NPMRDS, has been applied to the development of a number of visually intuitive, quantifiable performance measures [*3, 4, 5,* 6*, 7*]. Other studies have incorporated PVD into congestion performance indices used in national reports [*8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18*]. This research applies a baseline performance metric and visualization technique to quantify the Coronavirus (COVID-19) pandemic's impact on traffic operations and verify the decrease in speed variability near signalized intersections. The pandemic resulted in the closure of New Jersey State schools and offices on March 18, 2020 [*19*] followed soon after by a stay at home order starting March 21, 2020 [*20*]. The response to these closures was expected to decrease traffic congestion and speed variability, thus providing a unique opportunity to evaluate signal performance under non-congested conditions. Without traffic volumes at the study sites, the variations in speed provide the only means to characterize how traffic is moving through the system and can be used as a means to understand congestion conditions immediately following an unplanned, sustained emergency event.

Preliminary results defined by Rick Schuman of INRIX [*21*], indicate that personal travel in NJ is down by 36% between March 1 and May 29, the second highest decrease for all states, with a peak reduction of 60%.

This study entails the aggregation of approximately 495,208 speed records for analysis. The visual analysis defined in this paper provides a way to chronical evolving disruption and eventual return to normalcy in traffic patterns resulting from a sustained event. This research is conducted in preparation for a more in-depth state-wide analysis requiring the development of additional performance metrics evaluating billions of speed records against the baseline COVID-19 conditions.

2. DATA AND TEST BED

The evaluation of anonymous probe vehicle speed data requires a cross-reference between spatially defined Traffic Message Channels (TMC) and temporal speed datasets collected in one minute increments. The speed data associated with each TMC was provided by a commercial provider and available on the Regional Integrated Transportation Information System (RITIS) website [*22*]. Two study sites, US Route 9 and Schanck Road, and US Route 22 and Rock Avenue (Figure 1) were used to evaluate the main approach speed data over a 150-day study period starting January 1, 2020. In Figure 1, the TMC distances as well as the proximity to the study sites are shown. For this paper the time prior to the start of March 18, 2020 is referred to as Before-COVID (BC) and time after as During-COVID (DC).

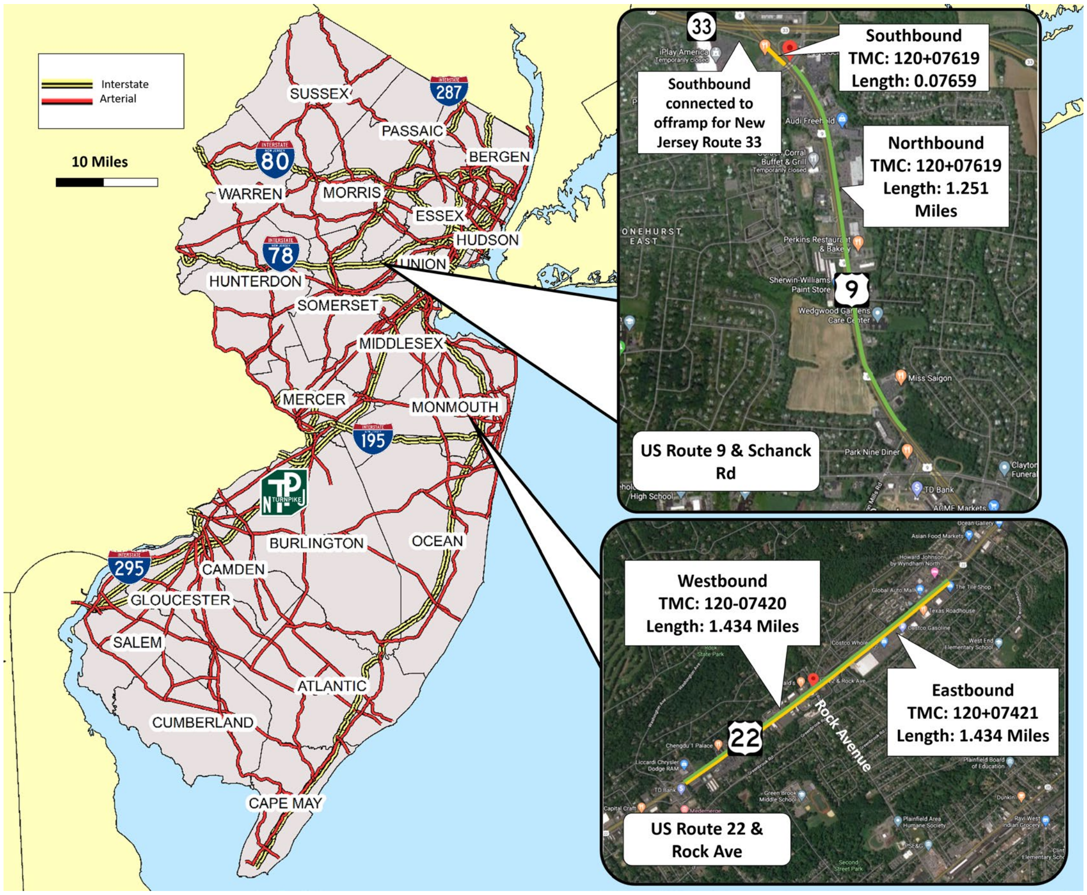


Figure 1. United States Route 9 and Route 22 study sites

3. MEASURES OF CONGESTION

Previous research using commercially available PVD had aggregated the data in 15-minute bins [*13*]. As part of the current research, three different bin sizes were compared to determine if a more granular data aggregation was necessary. The average speed for each TMC was determined by calculating the average speed of all available data in 15-, 10-, and 1-minute bins. The average speed (AvgSpeed) for each bin is calculated using the following:

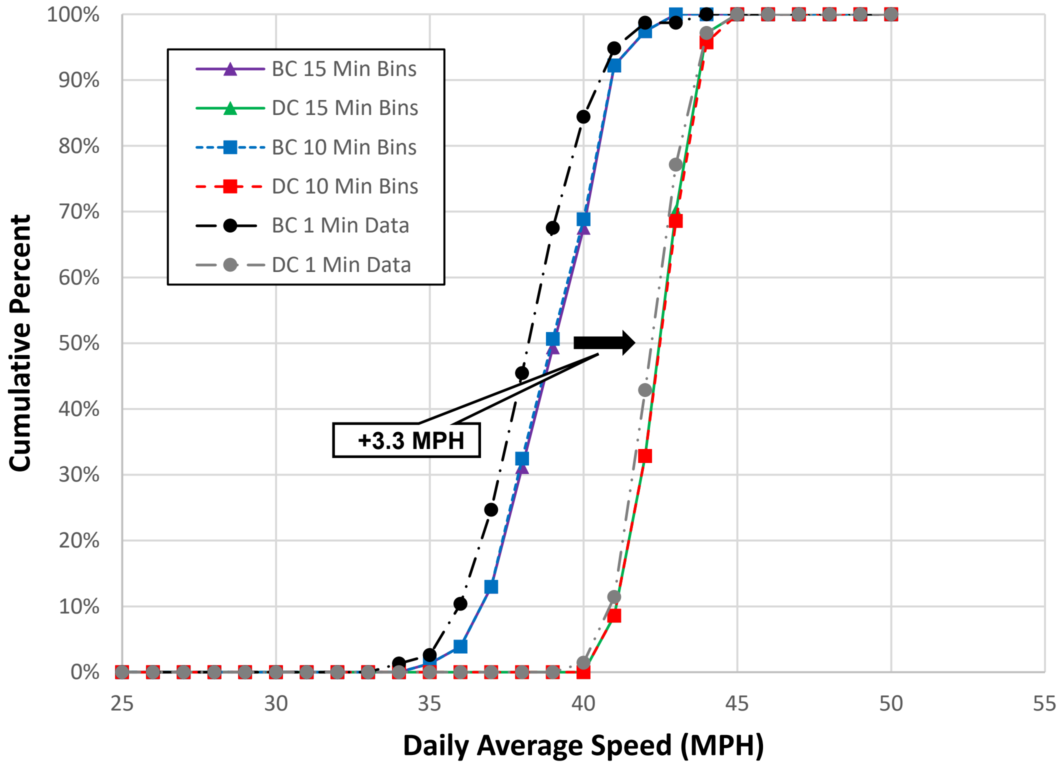
$$AvgSpeed_i = \frac{1}{n_j}\sum_{j \in F} v_{ij} \tag{1}$$

Where,

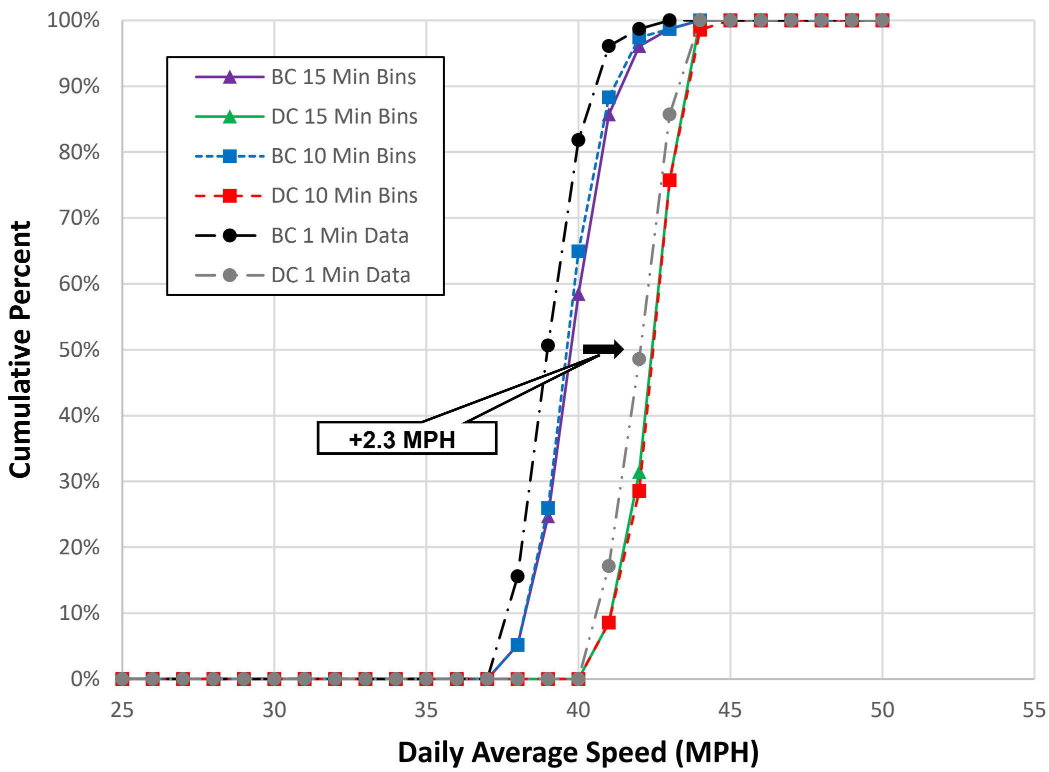
AvgSpeed is the average speed threshold for TMC i; v_{ij} are speed records for TMC i for the respective interval j; n_j is the total count of the binned intervals within study interval F, which is defined as all the bin periods (96 for 15-minute bins, 144 for 10-minute, and 1,440 for the raw 1-minute data) for each day of the study period. Only commercially available INRIX data with a high confidence score of 30 and a greater than 85% probability of reflecting current traffic conditions was used [*22*].

3.1 Comparison of Bin Sizes

A visual analysis of the cumulative distribution function (CDF) for all data collected is shown in Figure 2 and Figure 3. The figures illustrate the average BC and DC speeds for the 15-, 10-, and 1-minute bins. It is noted that no commercial data was available between March 28, 2020 through April 1, 2020. The CDF BC and DC median shifts for all the speed bins are reflected in the figures. Although all the shifts are relatively similar, the biggest shift occurred in Figure 3b, where southbound velocity increased by around 4 mph DC. It is also noted that the low average speed at this intersection is most likely due to the proximity to an off ramp from State Route 33. Although further research is needed to statistically evaluate different bin sizes, peak travel times, and the eventual return of the expected speeds, a simple analysis of the speed data appears to show that the three bin sizes provide similar results. In all cases there is an increase in speed, a result of COVID-19. A day-by-day breakdown of the data is shown in Figure 4 and Figure 5, where similar shifts in average speed can be seen for all direction in US-22 and US-9. Baseline speeds for these figures can be found in Table 1 and Table 2, where high variance can be seen in US-9 southbound. This may because of its short length, and connection to an offramp from NJ-33.

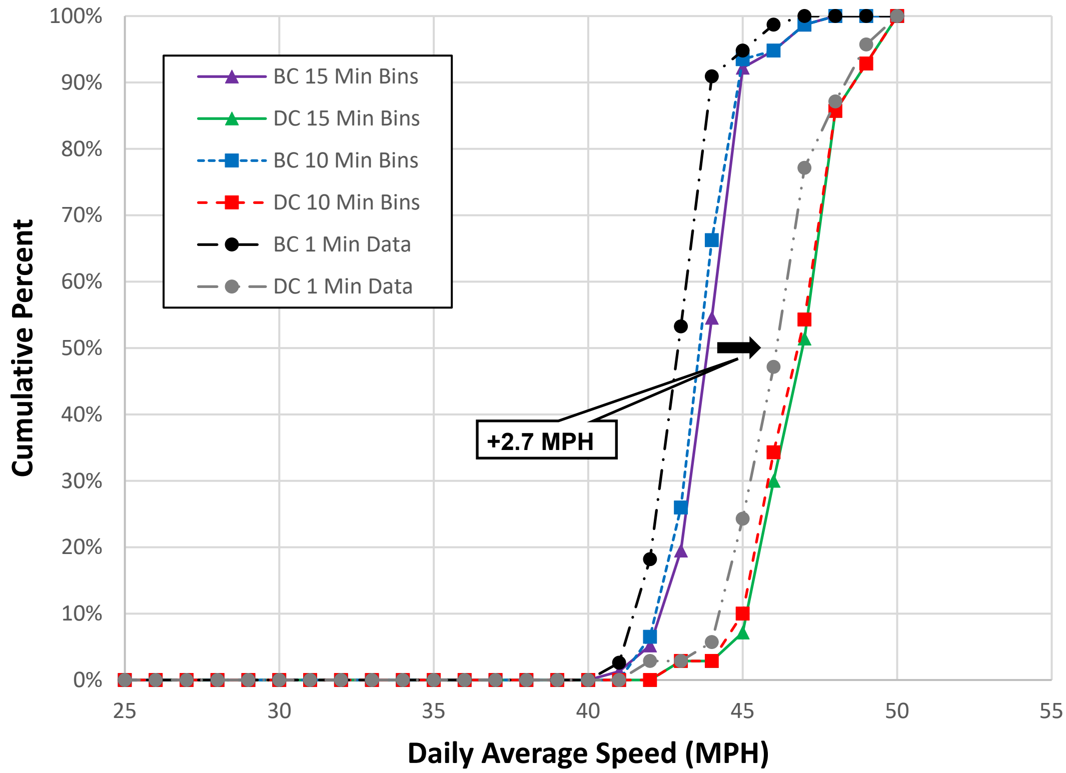


a. Eastbound

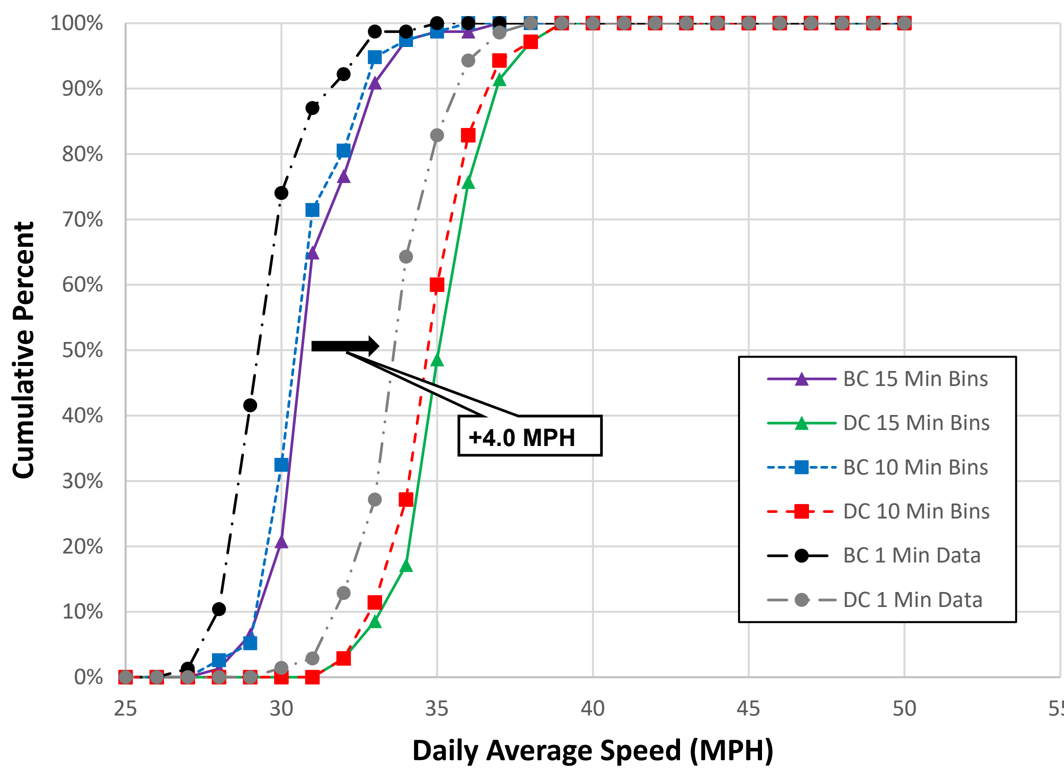


b. Westbound

Figure 2. US-22, Speeds for Varying Bin Sizes Before and After March 18, 2020 (Posted Speed 50 MPH)

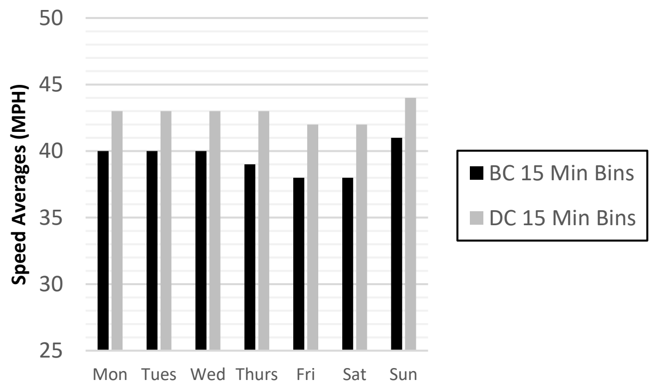


a. Northbound

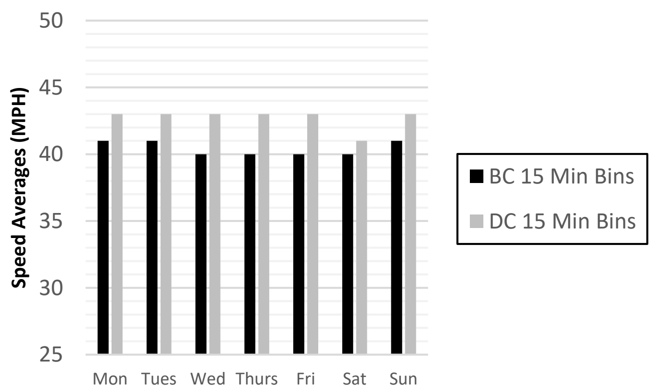


b. Southbound

Figure 3. US-9, Speeds for Varying Bin Sizes Before and After March 18, 2020 (Posted Speed 50 MPH)

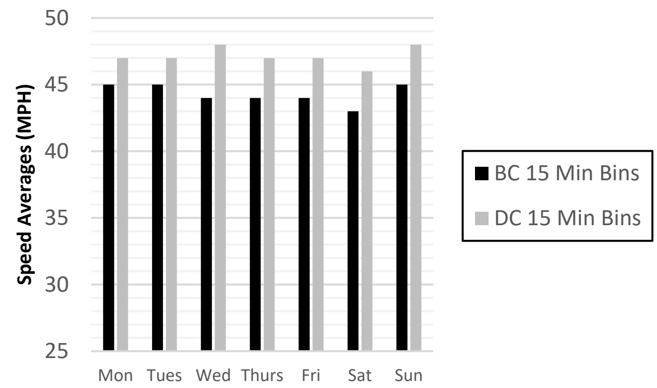


a. Eastbound

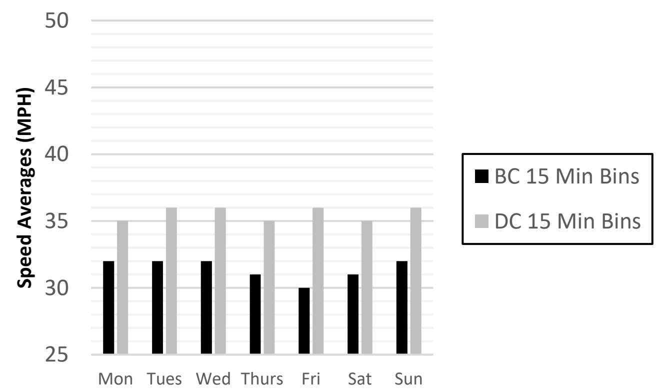


b. Westbound

Figure 4. Average Daily 15-Minute Speeds BC, DC for US Route 22 Intersection



a. Northbound



b. Southbound

Figure 5. Average Daily 15-Minute Speeds BC, DC for US Route 9 Intersection

Table 1. Before Covid Closings Weekday (Mon-Fri) Statistics Based on 1 Min. Data

Road/Direction	Avg(MPH)	Variance	St Dev.
US22 Eastbound	38.6	35.6	6.0
US22 Westbound	39.5	31.7	5.6
US9 Northbound	43.5	25.3	5.0
US9 Southbound	29.8	93.9	9.7

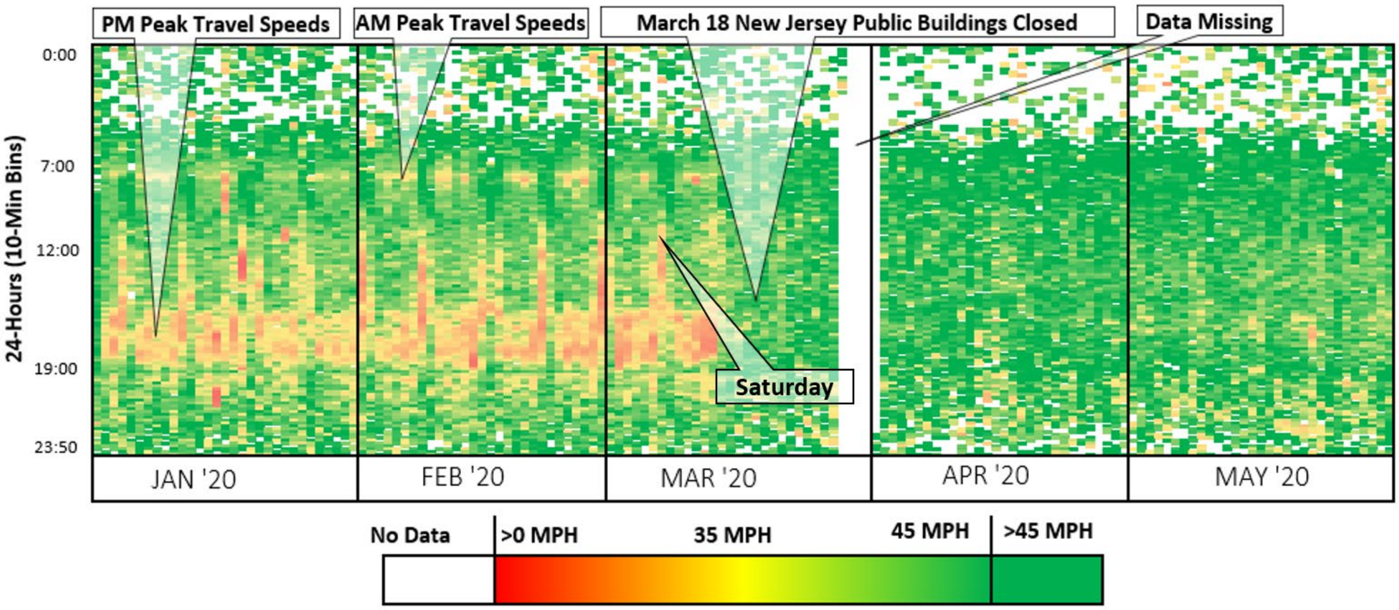
Table 2. During Covid Closings Weekday (Mon-Fri) Statistics Based on 1 Min. Data

Road/Direction	Avg(MPH)	Variance	St Dev.
US22 Eastbound	42.6	17.3	4.2
US22 Westbound	42.7	24.5	4.9
US9 Northbound	46.6	30.9	5.6
US9 Southbound	33.9	120.8	11.0

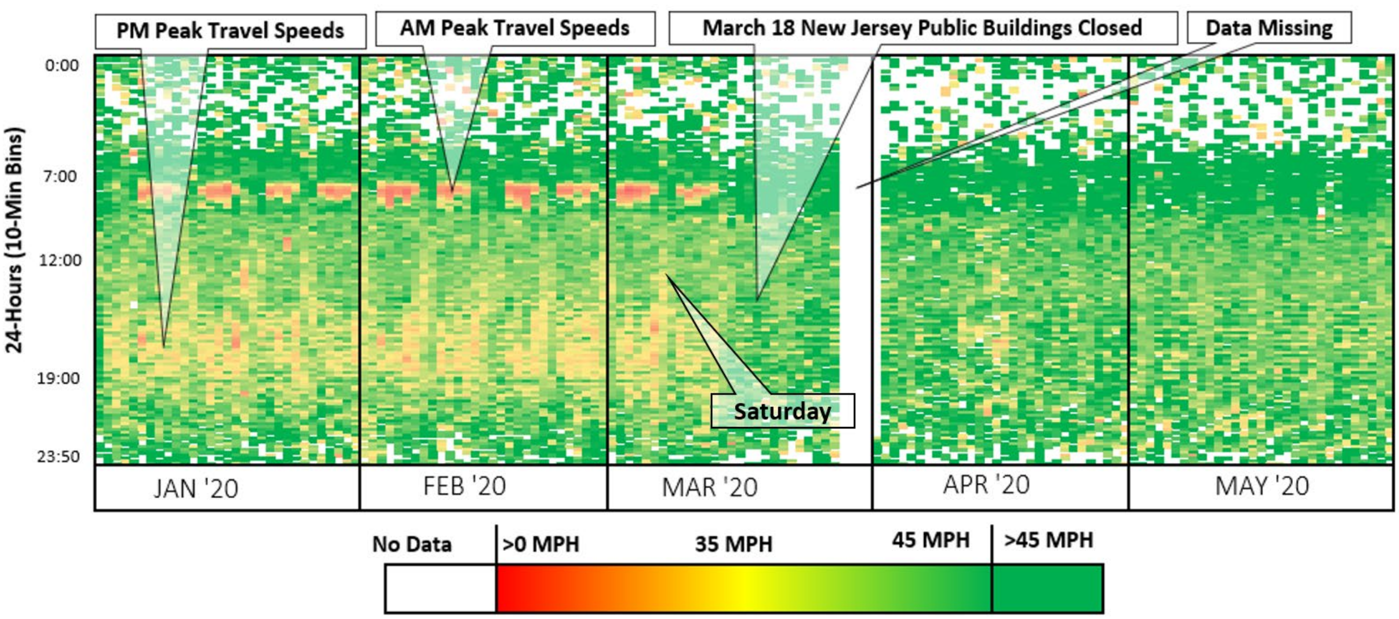
3.2 Average Speed Visualized

The US-22 study site was used to illustrate the juxtaposed average speeds per day in a heat map type format. For the eastbound and northbound approaches for US-22, both the 10-minute (Figure 6) and 15-minute (Figure 7) bins were used, where the y-axis is the respective bin period and the x-axis is the day of the year. Each cell reflects the AvgSpeed for their bin, which is used to characterize the speed near the intersection. These heat maps present a visualization of steady patterns, and deviations from those patterns. Traffic speed deviations emerge slightly earlier then March 18, beginning around Monday March 16. This may be due to the order going out March 16 [*19*]. In the figures, the transition from BC and DC is readily apparent. Also apparent are the BC AM and PM Peak times, weekend travel BC as well as the absence of both AM and PM Peak times DC. When comparing the two bin sizes, Figure 6 and Figure 7, it is apparent that there is not an immediately recognizable difference between the two.

The variance in speeds for BC and DC in Table 1 and Table 2, are visually represented in Figure 7 and Figure 8. For US22 the varience decreases as speeds become consistently faster, but with minimal peaks. US9 however sees its variances increase DC. This is apparent southbound, where peak travel times appear twice, once at 1300 and again at 2000, as compared to the more continuous flow BC. For US9 northbound the peak during the 1300 is more abrupt, as compared to BC when the congestion occurred throughout the day. It was expected that both intersections would have a decrease in varience, with a smoother flow. What was interesting about US9 is that the varience increased because the congestion that normally appears throughout the day became more confined to traditional peak times.

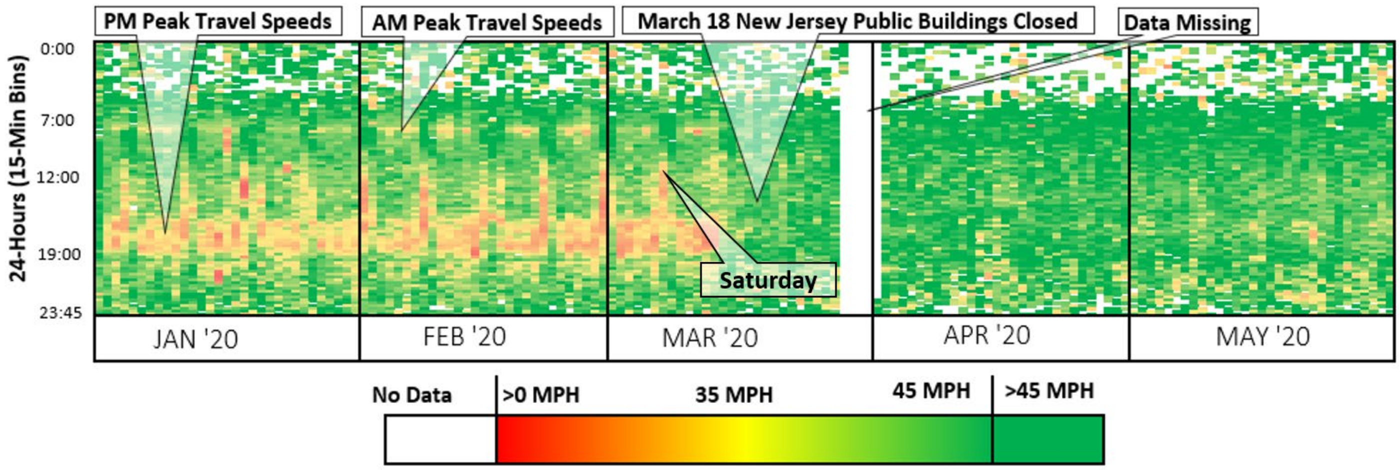


a. Eastbound

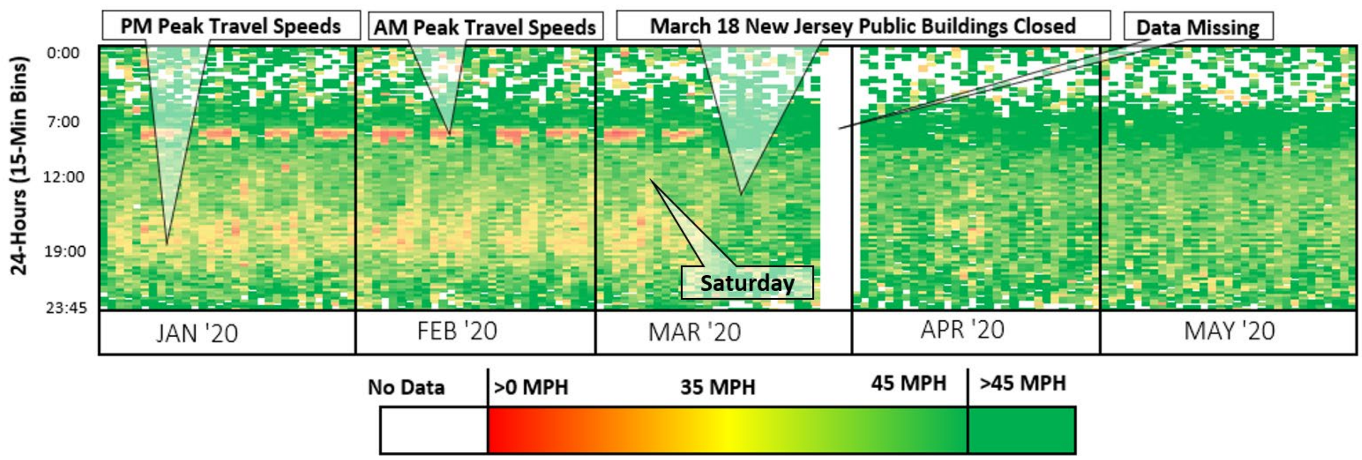


b. Westbound

Figure 6. 10-minute Bin Speed Distribution for US-22, 2020



a. Eastbound



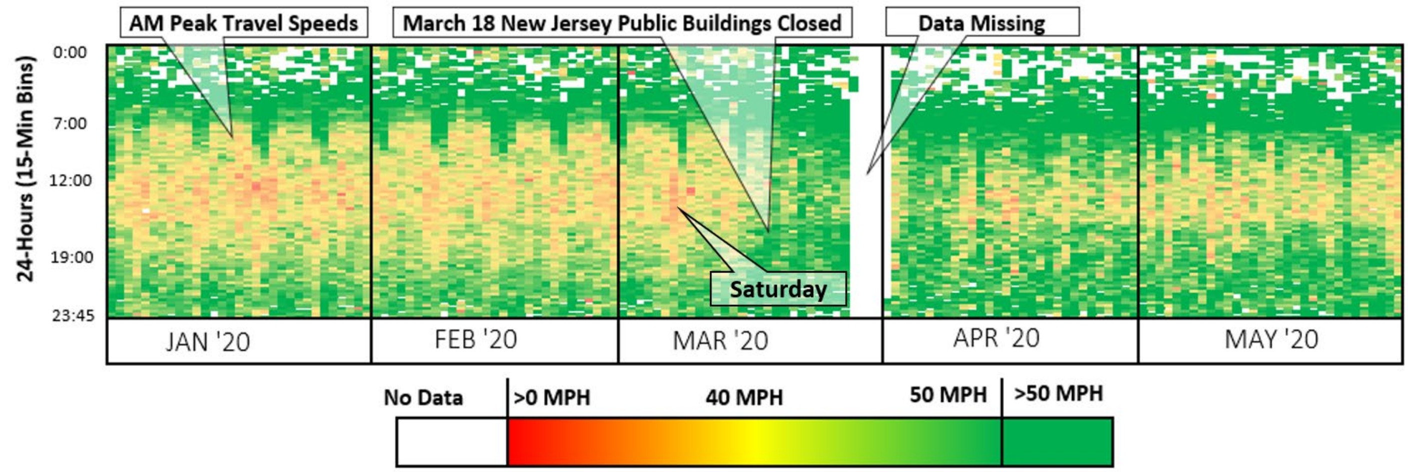
b. Westbound

Figure 7. 15-minute Bin Speed Distribution for US-22, 2020

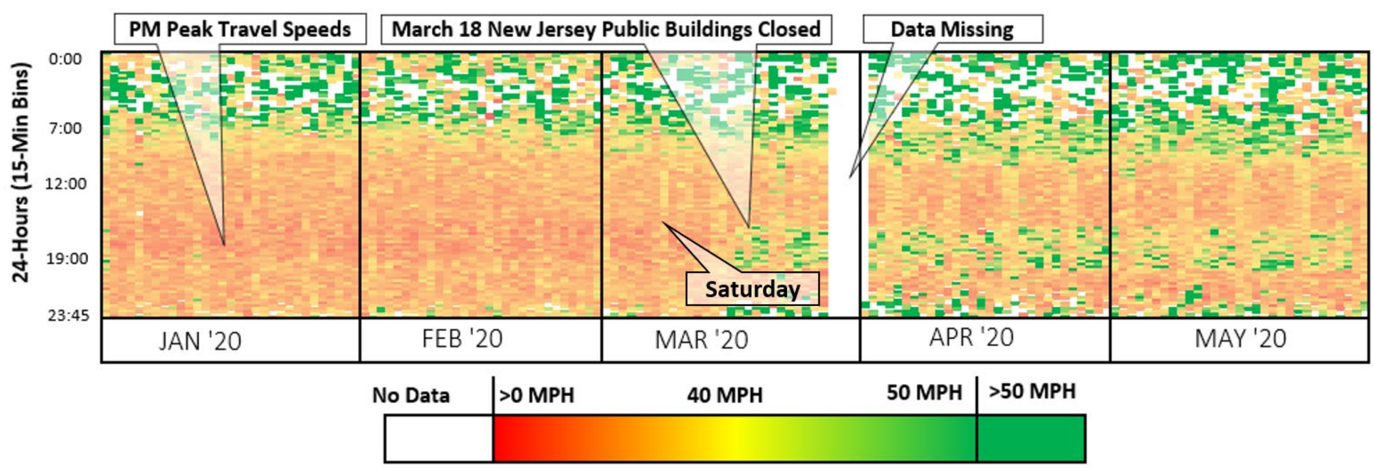
4. CONCLUSIONS

Due to the COVID-19 pandemic, a New Jersey executive order was made on March 16, 2020 to close schools starting March 18, 2020. This study was developed as a preliminary evaluation to determine how much of an impact this pandemic has on travel speeds, and speed variations through the use of visualizations and varying bin sizes. Ultimately this research will be expanded state-wide with more evolved performance metrics, but for this paper changes in average travel speed is measured at two intersections, US Route 22 and Rock Avenue, and US Route 9 and Schanck Road. The results of the analysis showed that there was definitive changes in travel speeds whose increase ranged from 2.3 to 4.0 MPH. A cumulative distribution frequency diagram demonstrated that the three different average speed bins (15-, 10-, and 1-minute) were relatively close, confirming that either the 15- or 10-minute bin would adequately characterize traffic speeds. Additional analysis indicated a decrease in speed variance for US Route 22, and an increase for US Route 9. It was expected that both intersections would have a decrease in variance, thus indicating smoother flow. For US 22, after reviewing the heatmap it became apparent that US 22 had defined peak travel times BC, which were gone DC. However, US 9 had a single Peak Travel time BC, but in DC it had two peak travel times, an indication that although traffic got less congested, the improvement of the traffic flow caused more variance in speed.

Based on these preliminary results, it appears that COVID-19 has had a significant impact on the motoring public of New Jersey. This was not unexpected. However, the unfortunate pandemic has presented a unique opportunity to study traffic patterns under smoother traffic flows under non-congested, at least partially non-congested conditions. The results support the need to extend this research across the region to better understand how the traffic system responded not only to the shut down, but to the eventual recovery. This data can also be used to better understand the traffic signalization and how retiming might improve traffic flow. Further research is underway to quantify the pandemic's disruption of state traffic patterns, tolling, and recovery process, as well as evaluating how to incorporate similar data revolving around COVID-19 [*23*].



a. Northbound



b. Southbound

Figure 8. 15-minute Bin Speed Distribution for US-9, 2020

5. ACKNOWLEDGMENTS

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein, and do not necessarily reflect the official views or policies of the sponsoring organizations. These contents do not constitute a standard, specification, or regulation. The speed data and segment information used in this report was obtained from INRIX Inc.

The authors confirm contribution to all parts of the paper. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Federal Highways Administration (FHWA) (2019). Operations Performance Measurement Program. Retrieved June 1, 2020 from https://ops.fhwa.dot.gov/perf_measurement/index.htm
2. Dadashova, B., P. Lasley, P. Koeneman, and S. Turner (2018), Approaches to Presenting External Factors with Operations Performance Measures, Report # FHWA-HOP-18-002
3. Brennan, T.M., Gurriell, R.A., Bechtel, A.J. and Venigalla, M.M., 2019. Visualizing and Evaluating Interdependent Regional Traffic Congestion and System Resiliency, a Case Study Using Big Data from Probe Vehicles. Journal of Big Data Analytics in Transportation, 1(1), pp.25-36. https://doi.org/10.1007/s42421-019-00002-y
4. Brennan Jr, T.M., Gurriell, R.A., Bechtel, A., & Venigalla, M.M. Performance Metrics for Visualizing Interdependent Regional Traffic Congestion Using Aggregated Probe Vehicle Data. Compendium of Papers, Transportation Research Board 97th Annual Meeting, January 2019.
5. Brennan Jr, T. M., Venigalla, M. M., Hyde, A., and LaRegina, A. (2018). Performance Measures for Characterizing Regional Congestion Using Aggregated Multi-Year Probe Vehicle Data Transportation Research Record: Journal of the Transportation Research Board (https://doi.org/10.1177/0361198118797190).

6. Brennan, T.M. and Venigalla, M., “Incorporating Speed Data to Analyze Evacuation Route Resiliency”, ITS World Congress; October 29-November 2, Montreal, Canada (2017) Paper AM-SP0892
7. Brennan Jr, T. M., Day, C. M., & Volovski, M. (2018b). Ranking Statistically Similar Highway Travel Corridors Using Speed Data Sets. Infrastructure Asset Management, 1-28.
8. California Department of Transportation, “Mobility Performance Report 2009”, February 2011.
9. Haghani, A., M. Hamdi, K.F. Sadabadi, “I-95 Corridor coalition vehicle probe project: validation of Inrix data July-September 2008-Final Report,” I-95 Corridor Coalition. January 2009.
10. Lomax, T., D. Shrank, B. Eiselee. 2011 Urban Mobility Report Powered by INRIX Traffic Data. College Station, Texas; The Texas A&M University System, Texas A& M Transportation Institute, 2011.
11. Lomax, T., D. Shrank, B. Eiselee. 2012 Annual Urban Mobility Report. College Station, Texas; The Texas A&M University System, Texas A& M Transportation Institute, 2012.
12. Maryland Department of Transportation, “2012 Maryland State Highway Mobility Report,” July 2012.
13. Remias S, Brennan T, Day C, Summers H, Cox E, Horton D, Bullock D. 2012 Indiana Mobility Report - Full Version. West Lafayette, Indiana: Purdue University, Indiana Mobility Reports; 2013. Report No.: ISBN: 978-1-62260-257-5.(2014)
14. Remias S, Brennan T, Grimmer G, Cox E, Horton D, Bullock D. 2011 Indiana Interstate Mobility Report - Full Version. West Lafayette, Indiana: Purdue University, Indiana Mobility Reports; 2012. Report No.: ISBN: 978-1-62260-209-4.
15. Remias, Stephen, M., T.M. Brennan, C.M. Day, H.T. Summers, D.K. Horton, E.D. Cox and D.M. Bullock, “Spatially Referenced Probe Data Performance Measures for Infrastructure Investment Decision Makers,” Submitted to Transportation Research Board, August 1, 2013, Paper No. 14-1062.
16. Schrank, D., Eisele, B., Lomax, T., & Bak, J. (2015). 2015 urban mobility scorecard.
17. Washington State Department of Transportation, “The 2012 Congestion Report,” August 2012.
18. Brennan, T. M., Remias, S. M., and Mariani, L. J. (2015). “Performance measures to characterize corridor travel time delay based on probe vehicle data.” Transportation Research Board 2526: 39-50)
19. Racioppi, Dustin. “Murphy Orders All NJ Schools to Close Starting Wednesday amid Coronavirus Outbreak.” North Jersey, 16 Mar. 2020, www.northjersey.com/story/news/new-jersey/2020/03/16/nj-coronavirus-schools-closed-across-state-announces-gov-murphy/5057646002/.
20. Schuman, Melanie. “New Jersey Governor Issues ‘Stay at Home’ Order.” CNN, 21 Mar. 2020, www.cnn.com/world/live-news/coronavirus-outbreak-03-21-20-intl-hnk/h_5749efe403dae8d2b7a9916b34db3e6e.
21. Schuman, Rick. “INRIX U.S. National Traffic Volume Synopsis Issue #11 (May 23 – May 29, 2020).” INRIX, 1 June 2020, inrix.com/blog/2020/06/covid19-us-traffic-volume-synopsis-11/.
22. www.ritis.org; lased accessed June 1, 2020
23. https://data.covid.umd.edu/; lased accessed June 1, 2020

Journal Home: https://journals.gmu.edu/index.php/jmms
Mason Journals ● Mason Publishing Group
Journal of Modern Mobility Systems (JMMS), Vol. 01(2020), 110-118
Document: https://doi.org/10.13021/jmms.2020.2894

Motivations and Mode-choice Behavior of Micromobility Users in Washington, DC

Siddartha Rayaprolu[a,*], Mohan Venigalla[a]

[a] *Sid and Reva Dewberry Department of Civil, Environmental and Infrastructure Engineering, George Mason University, Fairfax, VA 22030-4400; *Corresponding Author*

Received: 25 Sept. 2020 ● **Revised**: 23 Nov. 2020 ● **Accepted**: 24 Nov. 2020 ● **Published**: 27 Nov. 2020

Sponsor: https://ATPIO.org

ABSTRACT

The COVID-19 pandemic has reduced travel in general and disrupted travel patterns across the United States. The transit and ridehailing service ridership are particularly severely impacted. After an initial dip, shared micromobility services, including bikeshare, e-scooters, and e-bikeshare, have gained popularity as social distancing promoters with fewer points of contact. The findings of this article are based on the first phase of a two-phase mixed-mode survey of users and non-users of micromobility in Washington DC (n=440) in the Summer of 2019. While the phase-2 of the study is impacted by COVID-19 prevalence, results from the phase-1 are expected to serve as a critical baseline for post-pandemic travel behavior analysis and policy design. Findings indicate that each micromobility mode caters to different trip purposes and trip lengths of riders. While pleasure and time are identified as the biggest motivator for users, safety and pricing remain the most prominent barriers to users and non-users. Women and ethnic groups prefer to stay unimodal. Young and low-income users tend to be multimodal in their micromobility usage.

Keywords: Micromobility, Capital Bikeshare, E-scooters, E-bikeshare, Dockless bikeshare, Intercept survey, logistic regression, COVID-19

1. INTRODUCTION

As evidenced by their rapid adoption in recent years, shared micromobility services have resonated with consumers and investors, pointing to the likelihood of even more rapid growth in the future. Despite their widespread deployment in several metropolitan areas, very little is understood about the profiles and preferences of e-scooter users vis-à-vis a more mature station-based bikeshare system. As COVID-19 disrupted the travel behavior of users, it is of great importance to have a baseline reference to compare with the post-pandemic mode-choice behavior.

Earlier studies on station-based bikesharing have documented noteworthy findings on user demographics, mode-choice preferences, and the spatial equity of service [*1*]–[*5*]. However, very little is known about the relatively recent dockless systems users and their interactions with other modes. There are limited user-surveys that effectively portrayed the differences in characteristics among different micromobility users to understand their mode-choice behavior patterns. Furthermore, there are no past studies that analyze the multimodal behavior of micromobility users. A detailed summary of the user-survey literature is presented in Table 1

This research aims at understanding the demographics, perceptions, and preferences of micromobility users – both in absolute and relative terms - through a mixed-mode survey of micromobility users in the Washington DC metro area. We approach this goal by emphasizing three research questions:

1. What makes a person choose one micromobility mode over the other?
2. Which set of micromobility users tends to be multimodal?
3. How do users perceive individual micromobility mode?

2. THE STUDY AREA CHARACTERISTICS AND METHODOLOGY

At the time of this study (July 2019), Washington DC hosted seven e-scooter operators, one station-based bikeshare and one dockless e-bikeshare programs. The city also hosted dockless bicycles between late 2017 to early 2019, which were later replaced by e-scooters.

2.1 Survey design

A two-page mixed-mode survey instrument was designed to capture various characteristics of micromobility users. Approved by the Institutional Review Board, the survey was tailored to capture four types of potential respondents.

1. Capital Bikeshare (CaBi) users that do not prefer to use dockless vehicles
2. Users that prefer both CaBi and dockless vehicles based on individual trip purpose
3. Old CaBi users that completely shifted to dockless systems
4. New dockless vehicle users that never tried any micromobility systems before.

Dockless vehicle users include users from e-scooter, e-bikeshare and past dockless bikeshare services.

Table 1. Summary of recent literature on different shared-micromobility services

Authors	Year	Study area	Methods	Findings
				Station-based bikesharing
Kaviti et al.[3]	2019	Washington, DC	Survey	The majority of registered CaBi users (82%) chose bikes for commuting purposes, while a majority of casual CaBi users (57%) use them for social /recreational /sight-seeing /touring purposes
Chen, M., et al. [6]	2018	Hangzhou, China	Survey	1. Station-based bikesharing (SBS) and Free-floating bikesharing (FBS) have similar user structure, but different factors influence use frequency 2. SBS's strength is to have good quality with low cost while FBS is more flexible and is free to use
Buehler, R. & Hamre., A[7]	2019	Washington, DC	Survey	3. Savings in travel time (73% of users) and cost (25% of users) are significant motivators of CaBi. 4. Joining CaBi to save money had a significant positive association with new trips
				Dockless bikesharing
Hirsch et al.[8]	2019	Seattle, WA	Survey	Most resident bikeshare users are disproportionately young and white men who already use bicycles
Chen, Z., et al. [9]	2020	Beijing, China	Survey	1. Dockless bikeshare systems are more popular among younger, higher educated, or median-income groups and appear to be gender-independent. 2. Having a pro-bicycle attitude helps in the mode-choice behavior but does not account well for usage frequency
				E-bikesharing
Dill, J., & Rose, G.[10]	2012	Portland, OR	Survey	E-bikes help overcome some of the demographic barriers in society. They also address concerns over health problems related to inactivity, pollution, and other public policy problems to which private vehicles contribute
Campbell, A.A. et al.[11]	2016	Beijing, China	Survey	1. The average trip length of e-bikeshare in china to be between 2.5 to 2.8 miles. 2. They tend to divert users away from both the sheltered and unsheltered modes, as users tend to be less sensitive to trip distance, poor air quality, and severe temperatures.
He at al.[12]	2019	Salt Lake City, UT	Survey	1. The presence of e-bike systems near denser public areas with higher economic and recreational activity has a positive relationship with their ridership. 2. An average user identifies to be a visitor with a trip length of at least 5 miles, regardless of the hilly terrain.
Heineke et al.[13]	2019	United States	Market research	In the US, there is a $200B to $300B market potential for short-distance trips (under 5 miles), and shared micromobility can capture conservatively about 8 to 15% of this market
				E-scooter sharing
Smith, S.C., & Schwieterman, JP.[14]	2018	Chicago, IL	Trip data analysis	1. Popular trip length is between 0.5 to 2.0 miles. E-scooters can increase trips from 47% to 75% in a parking-constrained environment. 2. E-scooters do not compete with transit for longer trips due to economic viability, and they make at least 16% of jobs more accessible within 30 min of ride time
Liu et al.[15]	2019	Indianapolis, IN	Trip data analysis	The popular trip length of e-scooters is between 0.5 to 2.0 miles
Clewlow, R.[16]	2019	United States	Survey	1. E-scooters attained better gender equality compared to the earlier studies in the station-based bikesharing system. 2. 70% of the survey respondents supported micromobility and considered e-scooters a much convenient form of transport than personal car ownership.
James et al.[17]	2019	Arlington, VA	Survey	E-scooter trips in Rosslyn replaced trips otherwise taken by Uber, Lyft, or a taxi (39%), foot (33%), bicycle (12%), bus (7%), or car (7%)
McKenzie, G.[18][19]	2019	Washington, DC	API Data Analysis	1. The trip length of e-scooters is less than 5 min, as opposed to CaBi members (15 min) and casual users (40 min). 2. Capital Bikeshare tends to be more commuter focused whereas LimeBike reflects more leisure or non-commute related activities

Table 1. Summary of recent literature on different shared-micromobility services

Orr et al.[20]	2019	Portland, OR	Pilot/ Survey	1. Most users were people of color (74%), < 35 years (71%), with incomes < $30,000 (66%). 2. In the absence of e-scooters, people made trips with a motor vehicle (34%), TNCs/Taxis (15%), and Personal car (9%). 3. Cannibalized pedestrians (37%) and personal bike riders (5%). Attracted non-bikeshare users (74%) and non-bicyclists (42%) 4. Average trip length of 1.15 mi. Where users preferred e-scooters for trip connections (71%) and social/recreation trips (29%)
Sanders, R.L., et al. [21]	2020	Tempe, AZ	Survey	1. E-scooters used more for transport than recreation 2. Non-white non-riders significantly more likely to intend to try e-scooters 3. E-scooters disproportionately replace walking and bicycling for all trip types. 4. Women are significantly more likely to cite safety-related barriers to e-scooter use.

2.2 Survey execution

The intercept survey was conducted at 12 locations with higher activity of micromobility users. This field selection step was made based on preliminary observations from historical trip data of CaBi and E-scooters to understand the origin-destination patterns of these micromobility modes. Adequate samples were collected from all eight wards of Washington, DC, to ensure proper geographic coverage.

Although the intercept survey was designed to capture the users from the above categories of 1 and 2, it had limited potential to capture the resident respondents from categories 3 and 4. Such lack of response is a result of the time and price sensitivity of the dockless users, who does not prefer to be interrupted. Earlier research supports the importance of mixed-mode surveys in reducing the non-response error, and improve the quality of the data collected [22], [23]. Therefore, a web-version of the survey was circulated among the universities, major employer locations in the region, Reddit, Twitter, and LinkedIn.

3. DATA ANALYSIS AND RESULTS

A total of 440 responses from users and non-users of micromobility systems were analyzed. Users and non-user responses were distinguished based on an inbuilt option of "never used one before", within the questionnaire. A total of 309 respondents (Paper-based: 171; Web-based: 138) were found to have used a micromobility service at least once. Steps were taken to combine the responses from the mixed-mode survey and validate the sample against earlier peer-reviewed user survey studies on CaBi and E-scootersharing are described below.

3.1 Data validation

Table 2 presents the results of Pearson's chi-square test that compares the intercept and web-based survey samples. The test statistic, in conjunction with Cramer's V statistic, provides the strength of association between the two survey samples, in order to combine the datasets for model building. Except for gender, the respondent distribution from two types of surveys is not significantly ($\alpha = 0.05$) different from each other. Goodness-of-fit evaluations of current CaBi users against past user-survey study [24] indicate that the current sample of casual users closely resembles the CaBi users in all aspects except gender (Table 3). E-scooter users from the current survey were compared to the sample distribution of Portland's e-scooter pilot study (Figure 1). The percentage distribution of gender and racial characteristics of the users between the two studies are similar. However, the income group classifications among low-income groups appear dissimilar. Both studies suggest a higher dominance of higher-income groups among users.

Table 2. Pearson's Chi-square test for goodness of fit: intercept vs web-based surveys

Category	Subcategory [a]	Survey method: Intercept (n=171)	Survey method: Web *(n=138)*	χ^2	df	p-value	Cramer's V	Inference (based on α = 0.05)
Age	21-29 yrs.	73	73	5.681	3	**0.128**	0.126	*The age of the respondent is independent of the type of survey. Causation can be drawn on aggregated data*
	30-39 yrs.	64	48					
	40-49 yrs.	19	9					
	50-59 yrs.	13	5					
Gender	Female	68	38	4.312	1	0.038	0.137	*There is a moderate relationship between the gender of the respondent and the type of survey.*
	Male	101	97					
Income	< $20k	25	16	11.353	6	**0.078**	0.193	*The income group of the respondent is independent of the type of survey. Causation can be drawn on aggregated data*
	$20k-$34k	11	6					
	$35k-$49k	14	8					
	$50k-$74k	36	17					
	$75k-$99k	34	32					
	$100k-$149k	32	29					
	>$150k	17	28					
Race / Ethnicity	Asian	6	14	8.196	4	**0.085**	0.165	*The race/ethnicity of the respondent is independent of the type of survey. Causation can be drawn on aggregated data*
	Black/African American	9	10					
	Hispanic/Latino/Spanish origin	19	12					
	White	126	85					
	Other	10	11					

[a] Subcategories with a sample size of fewer than 5 respondents were not included in the test due to the chi-square test's analytical limitations.

Table 3. Sample characteristics of the current survey compared to CaBi user survey in 2017

The goodness of fit tests: Validation of the sample distribution

	Capital Bikeshare Members					*Capital Bikeshare Casual users*				
	2017[a]	2019[b]	χ^2	df	p-value[c]	2017	2019	χ^2	df	p-value[c]
	(n = 317)	(n = 86)	Inference			(n = 305)	(n = 148)	Inference		
Gender			1.2309	1	0.267			9.3684	1	0.002
Male	212	52	The gender composition of the two samples is not different			162	98	The gender distribution of casual users from both samples are different		
Female	105	34				155	50			
Ethnicity			1.379	1	0.2404			2.5604	1	0.1096
Non-White	60	22	The ethnic composition of member respondents is not different			103	43	The ethnic composition of casual users from both the surveys may not be different		
White	244	58				179	105			
Income			4.5174	2	0.1045			2.1234	2	0.3459
Low: < $35,000	32	10	The income level of the member respondents from both the surveys is not different			76	33	The two samples are different		
Medium: $35,000 - $100,000	127	44				130	70			
High: > $100,000	159	32				111	43			

[a] Capital Bikeshare user survey data from 2017, conducted by Shruthi et al. (Kaviti, Venigalla, and Lucas, 2019)

[c] The corresponding p-values were computed through the Monte Carlo simulation of B-replicates. Thereby, the degrees of freedom of the approximate chi-squared distribution of the test statistic are "NA"

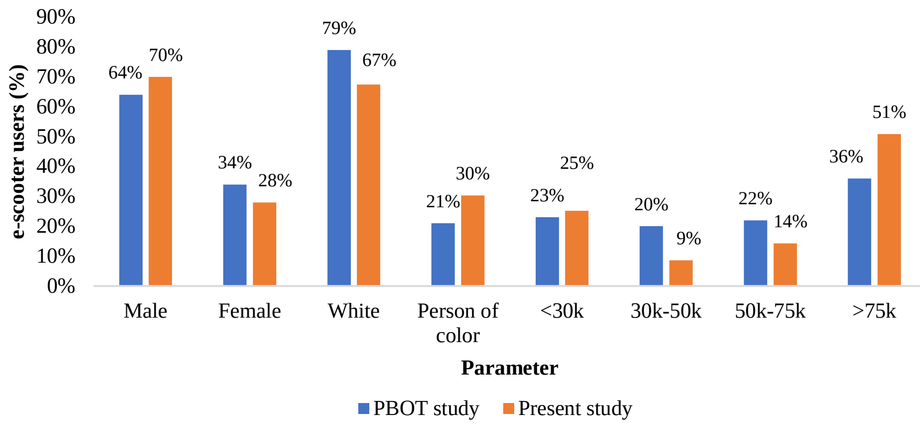


Figure 1. User characteristics of the current study v. Portland's pilot study

3.2 Logistic regression and odds ratio

The demographic characteristics of micromobility users (Figure 2) indicate perceivable differences between the users of multiple micromobility modes. Two logistic regression models (Table 4) were developed: One–to estimate the log-odds of the multimodal behavior of all the micromobility users; Two–to estimate the log-odds of the micromobility mode-choice outcomes of bikeshare users (CaBi and E-bikes) in comparison to E-scooter users. The logistic regression method estimates the odds or probability of response variable to take a particular value in response to a critical predictor value, usually while holding other predictors constant [25]–[27].

Multimodality refers to the tendency of a user to ride multiple transportation modes to reach their destination. The first regression model analyses the log-odds probability of a user to ride a single or multiple micromobility modes among all the four choices. Users that prefer to ride a single micromobility mode are classified as 'unimodal,' and the others are classified as 'multimodal'. The second model estimates the relative mode-choice preferences of bikeshare and e-scooters. Inferences from both the models together indicate the extent, serviceability and influence of each micromobility mode on a particular group of users.

The explanatory demographic variables include gender, age, income groups, ethnicity, car ownership, and usage frequency. The 'frequency of usage' variable is classified into two categories: Occasional (<1 ride per week) and Frequent (1 or more rides per week). Results from the Logistic regression of multimodal behavior of among micromobility users indicate that women (p-value=0.018) and people of color (p-value=0.052) are more likely to be unimodal, at higher significance levels. Lower-income groups (p-value=0.034) and younger users (p-value=0.032) are more likely to be multimodal. The comparative model indicates that younger users are more likely to choose e-scooters over CaBi (p-value=0.007) and E-bikeshare (0.036). Females (p-value=0.039) and medium-income households (p-value=0.053) are more likely to choose CaBi over e-scooters. There is no evidence of the significant influence of race, and personal car usage on the relative mode-choice.

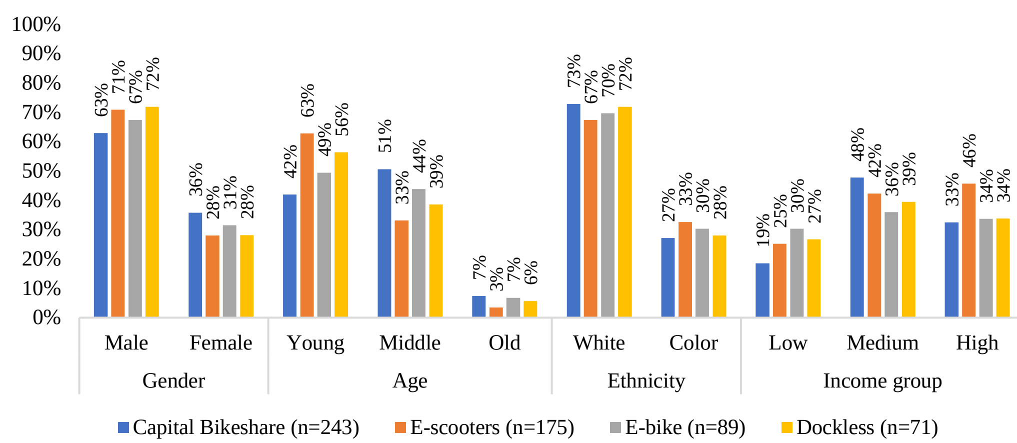


Figure 2. Demographic characteristics of micromobility users in Washington, DC

Table 4. Logistic regression models on user characteristics and mode-choice behavior

Logistic regression model 1: Multimodal characteristics of micromobility users

Mode (marginal %)	Parameter (marginal %)		β	SE	p-value	Exp(B)	95% CI	
	Intercept		-1.084	0.525	**0.039**		Lower	Upper
Multimodal User (45.2%)	Gender	Female (34.9%)	-0.623	0.263	**0.018**	0.536	0.321	0.897
	Age	Young (48.3%)	1.245	0.581	**0.032**	3.474	1.112	10.852
		Middle (44.5%)	0.986	0.558	0.07	2.679	0.898	7.993
	Income	Low (18.8%)	0.878	0.414	**0.034**	2.405	1.069	5.409
		Medium (47.3%)	0.107	0.295	0.716	1.113	0.624	1.985
	Race	Non-White (30.5%)	-0.53	0.272	**0.052**	0.589	0.345	1.004
Model fitting criteria	**(-2) Log-Likelihood**		**Chi-square**		**df**		**Sig.**	
	90.149		21.849		6		0.001	
Pearson's goodness-of-fit statistics			33.058		23		0.08	
a. The reference category is Unimodal user								
Reference parameters: Male, Old, High income, and White users of micromobility								

Logistic regression model 2: Mode-choice preferences of bikeshare users in reference to e-scooter users

Mode (marginal %)	Parameter (marginal %)		β	SE	p-value	Exp(B)	95% CI	
		Intercept	1.656	0.61	0.007		Lower	Upper
Capital Bikeshare (48%)	Gender	Female (33.1%)	0.481	0.233	**0.039**	1.618	1.025	2.555
	Age	Young (51.6%)	-1.593	0.589	**0.007**	0.203	0.064	0.645
		Middle (44.9%)	-0.607	0.555	0.274	0.545	0.184	1.617
	Income	Low (22.9%)	0.577	0.367	0.116	1.781	0.867	3.659
		Medium (44.9%)	0.54	0.279	**0.051**	1.716	0.99	2.967
	Race	Non-White (28.3%)	-0.208	0.24	0.386	0.812	0.508	1.3
	Car owner-ship	No (40.7%)	-0.024	-0.379	0.949	0.976	0.465	2.05
		Yes (48.2%)	-0.094	0.361	0.794	0.91	0.448	1.847
	Usage	Occasional (68.4%)	-0.945	0.242	**0.000**	0.389	0.242	0.625
		Intercept	-0.44	0.775	0.57			
E-bikeshare (17.5%)	Gender	Female (33.1%)	0.306	0.297	0.304	1.357	0.758	2.431
	Age	Young (51.6%)	-1.537	0.733	**0.036**	0.215	0.051	0.905
		Middle (44.9%)	-0.494	0.681	0.468	0.61	0.161	2.317
	Income	Low (22.9%)	0.701	0.454	0.123	2.015	0.828	4.907
		Medium (44.9%)	0.012	0.361	0.97	1.012	0.499	2.056
	Race	Non-White (28.3%)	-0.158	0.304	0.603	0.854	0.47	1.55
	Car owner-ship	No (40.7%)	0.386	0.49	0.43	1.472	0.564	3.841
		Yes (48.2%)	-0.138	0.467	0.767	0.871	0.349	2.174
	Usage	Occasional (68.4%)	0.594	0.369	0.108	1.81	0.879	3.73
Model fitting criteria	**(-2) Log-Likelihood**		**Chi-square**		**df**		**Sig.**	
	332.91		65.899		18		0	
Pearson's goodness-of-fit statistics			110.974		192		1	
a. The reference category is E-scooters								
Reference parameters: Male, Old, High income, White, Car owners that do not drive, and Frequent users of micromobility								

3.3 Shared and Micromobility Mode-choice and Trip length

A chi-square test of independence among the user preferences of CaBi, e-bikeshare, and e-scooter indicated that the mode-choice is predominantly dependent on trip purpose ($\chi 2$: 14.31, p-value: 0.02636). Figure 3 illustrates the mode-choice preferences of the private shared-mobility and micromobility modes through a stacked bar plot.

Average trip length is useful in understanding the role of a particular mode within a set of modes available to a user. It depends on several factors like time and price sensitivity of the customers and trip purpose. The odds ratio analysis (Table 5) suggests that e-bikeshare and CaBi are more popular for trips less than 5-min and trips between 15-30 min, respectively. While E-scooters are found to be popular for 5-15 min trips, this finding is less significant.

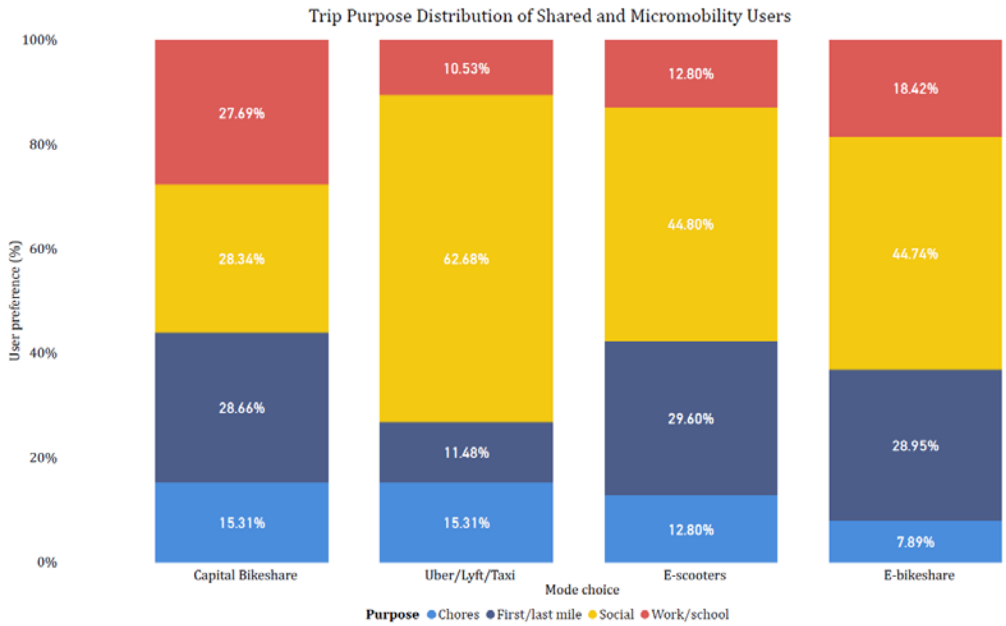


Figure 3 Mode-choice of micromobility users based on trip purpose

Table 5 Trip-length based odds ratio analysis of mode-choice

Trip length	Reference Service	Parameter	*Capital Bikeshare	Dockless bikeshare	E-bikeshare	Scootershare
			(n = 237)	*(n = 71)*	*(n = 77)*	*(n = 160)*
< 5 min	E-bikeshare	Odds ratio	4.34	0.783		1.6
		95% Conf. interval	2.212 < OR < 8.515	0.39 < OR < 1.573		0.8538 < OR < 2.998
5 - 15 min	Electric scooters	Odds ratio	1.191	1.418	1.239	
		95% Conf. interval	0.798 < OR < 2.615	0.7946 < OR < 2.531	0.7143 < OR < 2.149	
15 - 30 min	Capital Bikeshare	Odds ratio		2.062	1.913	2.514
		95% Conf. interval		1.166 < OR < 3.647	1.108 < OR < 3.301	1.624 < OR < 3.892

** Third generation bikeshare;*

Interpretation examples:

[1] The odds of choosing Dockless bikeshare for trips < 5 minutes are 5.5 times higher than that of Capital Bikeshare

[2] The odds of choosing Capital Bikeshare for trips 15-30-minute duration are 2.237 times higher than that of Dockless bikeshare

3.4 Shared and Micromobility Mode-choice and Trip length

Very little is known about the reasons behind the relative differences in consumer affinity for these shared micromobility systems. Users and non-users were asked to provide their opinion on potential reasons for using different modes 'more' or 'less' often. While users perceived these modes as fun and time-saving alternatives, safety and disinterest remain major barriers to their patronage among both users and non-users Figure 4. Among all micromobility modes, e-scooters were significantly considered to be unsafe.

Around 23% of the respondents considered e-scooters unsafe, but a majority of them considered them to be fun (59%) and time-saving (51%). This observation complements our earlier finding that e-scooters are more preferred for social and recreational trip purposes. Among all micromobility modes, both e-scooters and CaBi provide better-perceived incentives to the users than the other two modes included in the survey. However, CaBi appears to have added advantages of being perceived as more economical, fitness-promoting, time-saving, and easier to use than e-scooters.

Anecdotal references from the survey respondents indicate that most CaBi users prefer not to use other modes due to their existing long-term membership. However, some users expressed their interest in using the e-scooters for social or recreational purposes occasionally. This observation reinforces our earlier deductions from the odds ratio analysis that CaBi members are more likely to use e-scooters occasionally than regularly. Furthermore, several respondents indicated that the uncertainty in the dockless vehicles' availability at a specific location had reduced their interest in choosing those modes.

Question: What makes you ride this micromobility mode more often?				
Incentives	Capital Bikeshare (n=244)	Dockless Bikeshare (n=71)	E-bikeshare (n=89)	E-scooters (n=175)
Hassle free/ Easy to use	45%	25%	27%	39%
Economical	48%	21%	16%	23%
Time saving	45%	27%	39%	51%
Safe	11%	7%	6%	4%
Healthy	45%	20%	10%	7%
It's fun!	47%	30%	39%	59%
*Question: What makes you ride this micromobility mode less often?**				
Barriers	Capital Bikeshare (n=440)	Dockless Bikeshare (n=440)	E-bikeshare (n=440)	E-scooters (n=440)
Not interested/ Not viable	36%	45%	44%	34%
Expensive	8%	13%	16%	18%
Time consuming	5%	2%	2%	2%
Unsafe	10%	8%	9%	23%
Traffic/Pollution	6%	4%	4%	7%
Scale				
Most preferable			*Least preferable*	
100%	80%	60%	40%	20%
**Both users and non-users were asked to provide their opinion*				

Figure 4 Heat chart of the user and non-user perceptions on individual micromobility mode

4. DISCUSSION AND CONCLUSION

Logistic regression results suggest that lower-income groups and younger adults are more likely to ride multiple micromobility modes. As each micromobility mode caters to different trip purposes and trip lengths, multimodality indicates the consistency of user reliability on micromobility modes for most of their travel needs, without drifting away to high-carbon modes. For example, survey responses suggest that e-scooters attracted users from personal cars (36%) and Uber/Lyft/taxi services (22%). While younger adults have a higher likelihood of using e-scooters, women and medium-income groups preferred CaBi to e-scooters. Significant differences in trip lengths and trip purposes among different micromobility users indicate that each mode caters to the needs of specific groups of people. However, respondents were drawn away from these modes, primarily due to safety and budget concerns. Such an observation indicates the need for more protected bike lanes, parking infrastructure, and community outreach programs.

The research findings can serve as a basis for cities to deploy more detailed and large-scale surveys to understand the impact of community emergencies on regional and local transportation networks. However, as the user preferences and perceptions tend to vary with geographic region, caution must be exercised in extending the findings of this study to other regions. This study was conducted a few months before the COVID-19. As the pandemic is widely expected to change the travel demand by traditional and micromobility modes, this study can potentially serve as a valuable baseline for evaluating the variations in mode-choice behavior of Washington DC micromobility users in the post-pandemic environment.

5. REFERENCES

[1] K. B. Campbell and C. Brakewood, "Sharing riders: How bikesharing impacts bus ridership in New York City," *Transp. Res. Part A Policy Pract.*, vol. 100, pp. 264–282, Jun. 2017.

[2] S. A. Shaheen, S. Guzman, and H. Zhang, "Bikesharing in Europe, the Americas, and Asia," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2143, no. 1, pp. 159–167, Jan. 2010.

[3] S. Kaviti, M. M. Venigalla, and K. Lucas, "Travel behavior and price preferences of bikesharing members and casual users: A Capital Bikeshare perspective," *Travel Behav. Soc.*, vol. 15, pp. 133–145, Apr. 2019.

[4] M. Venigalla, S. Kaviti, and T. Brennan, "Impact of Bikesharing Pricing Policies on Usage and Revenue: An Evaluation Through Curation of Large Datasets from Revenue Transactions and Trips," *J. Big Data Anal. Transp.*, vol. 2, no. 1, pp. 1–16, Feb. 2020.

[5] E. Fishman, S. Washington, N. Haworth, and A. Mazzei, "Barriers to bikesharing: An analysis from Melbourne and Brisbane," *J. Transp. Geogr.*, 2014.

[6] M. Chen, D. Wang, Y. Sun, E. O. D. Waygood, and W. Yang, "A comparison of users' characteristics between station-based bikesharing system and free-floating bikesharing system: case study in Hangzhou, China," *Transportation (Amst).*, vol. 47, no. 2, pp. 689–704, Apr. 2020.

[7] R. Buehler and A. Hamre, "Business and Bikeshare User Perceptions of the Economic Benefits of Capital Bikeshare," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2520, no. 1, pp. 100–111, Jan. 2015.

[8] J. A. Hirsch, I. Stewart, S. Ziegler, B. Richter, and S. J. Mooney, "Residents in Seattle, WA Report Differential Use of Free-Floating Bikeshare by Age, Gender, Race, and Location," *Front. Built Environ.*, vol. 5, p. 17, Mar. 2019.
[9] Z. Chen, D. van Lierop, and D. Ettema, "Exploring Dockless Bikeshare Usage: A Case Study of Beijing, China," *Sustainability*, vol. 12, no. 3, p. 1238, Feb. 2020.
[10] J. Dill and G. Rose, "Electric Bikes and Transportation Policy," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2314, no. 1, pp. 1–6, Jan. 2012.
[11] A. A. Campbell, C. R. Cherry, M. S. Ryerson, and X. Yang, "Factors influencing the choice of shared bicycles and shared electric bikes in Beijing," *Transp. Res. Part C Emerg. Technol.*, vol. 67, pp. 399–414, Jun. 2016.
[12] Y. He, Z. Song, Z. Liu, and N. N. Sze, "Factors Influencing Electric Bike Share Ridership: Analysis of Park City, Utah," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2673, no. 5, pp. 12–22, May 2019.
[13] F. Heineke, K., Kloss, B., Scurtu, D. and Weig, "Micromobility's 15,000-mile checkup," *McKinsey & Company*, Jan-2019.
[14] C. S. Smith and J. P. Schwieterman, "E-Scooter Scenarios: Evaluating the Potential Mobility Benefits of Shared Dockless Scooters in Chicago," Dec. 2018.
[15] M. Liu, S. Seeder, H. L.-I. of T. Engineers, and U. 2019, "Analysis of E-Scooter Trips and Their Temporal Usage Patterns," *Inst. Transp. Eng. ITE J.*, vol. 89, no. 6, pp. 44–49, 2019.
[16] R. Clewlow, "The Micro-Mobility Revolution: The Introduction and Adoption of Electric Scooters in the United States," 2019.
[17] O. James, J. Swiderski, J. Hicks, D. Teoman, and R. Buehler, "Pedestrians and E-Scooters: An Initial Look at E-Scooter Parking and Perceptions by Riders and Non-Riders," *Sustainability*, vol. 11, no. 20, p. 5591, Oct. 2019.
[18] G. McKenzie, "Spatiotemporal comparative analysis of scooter-share and bike-share usage patterns in Washington, DC," *J. Transp. Geogr.*, vol. 78, pp. 19–28, Jun. 2019.
[19] G. McKenzie, "Docked vs. Dockless bike-sharing: Contrasting spatiotemporal patterns," in *Leibniz International Proceedings in Informatics, LIPIcs*, 2018, vol. 114.
[20] B. Orr, J. MacArthur, and J. Dill, "The Portland E-Scooter Experience," *TREC Friday Semin. Ser.*, Feb. 2019.
[21] R. L. Sanders, M. Branion-Calles, and T. A. Nelson, "To scoot or not to scoot: Findings from a recent survey about the benefits and barriers of using E-scooters for riders and non-riders," *Transp. Res. Part A Policy Pract.*, vol. 139, pp. 217–227, Sep. 2020.
[22] D. A. Dillman *et al.*, "Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet," *Soc. Sci. Res.*, vol. 38, no. 1, pp. 1–18, Mar. 2009.
[23] J. W. Sakshaug, A. Cernat, and T. E. Raghunathan, "Do Sequential Mixed-Mode Surveys Decrease Nonresponse Bias, Measurement Error Bias, and Total Bias? An Experimental Study," *J. Surv. Stat. Methodol.*, vol. 7, no. 4, pp. 545–571, Dec. 2019.
[24] S. Kaviti, M. Venigalla, and K. Lucas, "Portraying and Differentiating Profiles and Preferences of Casual Users and Registered Members of Capital Bikeshare," 2019.
[25] D. G. Kleinbaum, K. Dietz, M. Gail, M. Klein, and M. Klein, *Logistic regression*. Springer, 2002.
[26] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*, vol. 398. John Wiley & Sons, 2013.
[27] M. Szumilas, "Explaining odds ratios," *J. Can. Acad. Child Adolesc. Psychiatry*, vol. 19, no. 3, pp. 227–229, Aug. 2010.

Journal Home: https://journals.gmu.edu/index.php/jmms
Mason Journals • Mason Publishing Group
Journal of Modern Mobility Systems (JMMS), Vol. 01(2020), 119-124
Document: https://doi.org/10.13021/jmms.2020.2929

Applicability of Long Short-Term Memory Traffic Volume Imputation Model to Drive Connected Corridor Simulation

Abhilasha Saroj[a, *], Angshuman Guin[a], Michael Hunter[a]

[a] *School of Civil and Environmental Engineering, Georgia Institute of Technology, 788 Atlantic Drive NW, Atlanta, GA 30332; *Corresponding Author*

Received: 12 Dec. 2020 • **Revised**: 28 Dec. 2020 • **Accepted**: 29 Dec. 2020 • **Published**: 29 Dec. 2020

ABSTRACT

For effective implementation of connected corridor applications, it is imperative to study the characteristics of the high-resolution connected corridor data streams leveraged in smart city applications. In a previous effort, a smart city application – real-time corridor data-driven traffic simulation model, i.e., Digital Twin – is developed. Investigation of the corridor field volume data revealed the presence of data gaps. To address these gaps, deep Long Short-Term Memory (LSTM) Recurrent Neural Network univariate and multivariate volume imputation models are developed. In this paper, the impact of the developed model imputations on the digital twin generated travel times are investigated. Simulation runs are conducted for typical and atypical traffic, for three volume input cases: base (original volumes), univariate model imputations, and multivariate model imputations. For the given methodology it was seen that: 1) the travel times generated using multivariate imputations are the closest to that generated using base data, 2) the impact of imputations on travel times is focused on congested routes, and 3) the impact on travel time is minimal despite input volume overestimation on routes that have the capacity to accommodate higher volumes. These findings demonstrate the need to prioritize data streams based on the given application and underlying corridor conditions.

Keywords: smart cities, connected corridor, long short-term memory, real-time simulation, traffic volume imputation

1. INTRODUCTION

Smart cities across the world utilize smart corridor testbeds to explore technology implementations [*1*, *2*, *3*, *4*, *5*]. Often, a smart corridor is equipped with communications technologies [*3*], enabling the transfer of significant data between vehicles, the infrastructure, and corridor management centers. These data can be in different forms, such as connected vehicle data providing high resolution instantaneous vehicle specific data and signal phase and timing data, vehicle counts from in-road or roadside detectors, probe vehicle data such as that from INRIX [*6*], HERE [*7*], etc., to name a few. Smart corridor applications seek to convert these data into actionable information, to improve corridor performance. However, the presence of data gaps in the data streams can impair such efforts. Thus, it is imperative to develop data imputation methodologies as well as to understand the impact of such imputation on the application performance.

In a previous effort the authors developed a smart corridor application, a real-time data-driven traffic simulation model, i.e., Digital Twin, for the North Avenue Smart Corridor in Atlanta, Georgia [*8*, *9*]. The Digital Twin, driven using high frequency volume and signal data, is capable of dynamically providing corridor traffic and environmental performance measures [*8*, *9*]. However, investigation of the corridor data streams revealed the presence of data gaps. To address the volume data gaps, bi-directional Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNNs) univariate and multivariate imputation models were developed. Experiments were conducted to investigate the LSTM RNN model performance under typical and atypical day conditions [*10*]. Of specific interest was exploring if supplementing historic data from the given data stream with the most recent data from similar data streams (multivariate model) provides superior predictions over utilizing only historic data (univariate model). Experimental results indicated the potential for a multivariate LSTM RNN model to provide reasonable imputations on typical and atypical days [*10*].

In this paper the reasonableness of the data estimations is explored by identifying limitations and evaluating the appropriateness of the imputation models to drive the Digital Twin. A simulation experiment is conducted using the Digital Twin to investigate the impact of the univariate and multivariate model imputations on generated travel times for selected corridor routes.

2. LSTM RNN MODEL DEVELOPMENT

RNNs are a variant of Neural Networks, capable of utilizing the "memory" of previous event data to predict the next values in a sequence [*11*]. However, RNNs can suffer from the vanishing gradient problem in backpropagation implementation [*12*]. This may be a drawback when accounting for long-term dependency in a sequence is crucial to prediction accuracy. LSTM RNN [*13*] seeks to address this issue through the inclusion of a 'memory' cell component along with gates to regulate the memory cell value [*14*]. A variant of LSTM RNN is bidirectional LSTM RNN (BLSTM), where output mapping may learn from both past and future information [*15*].

In the previous study [*10*], deep bidirectional LSTM RNNs were used to develop univariate and multivariate volume time series prediction models for six selected detectors on three approaches on the North Avenue corridor, a 2.3-mile long actuated corridor, as shown in Figure 1. Each of these approaches has two lanes, referenced as L_1 and L_2. The multivariate models are trained using the historic data of the detector experiencing data loss as well as data from a corridor detector drawn from a cluster of detectors that have been identified to have a similar time series data pattern using cluster analysis. For a comprehensive literature review on time series similarity measures, traffic data imputation methodologies, and the LSTM RNN model development process, the reader is referred to Saroj [*10*].

3. EXPERIMENT DESIGN

A simulation experiment is designed to study the impact of the previously developed univariate and multivariate prediction models on simulation generated performance measures for a typical weekday, Monday, March 18th, 2019, and a weekday with atypical traffic conditions, Monday, May 27th, 2019, (Memorial Day). For each of these days the PM peak hours (3 PM to 6 PM) are simulated for three traffic volume sets, input at the three corridor approaches: 1) base traffic condition (original volume), 2) univariate model imputations, and 3) multivariate model imputations. A discussion of the base traffic volumes and signal timings for each experiment day may be found in [*9*]. The second and third volumes cases assume a three-hour data gap in the base traffic data at the three study approaches, utilizing the imputed volumes for these locations. Volumes during simulation run-time are imputed (i.e., predicted) as would occur in a real-time event, that is, the simulation model and algorithms are only fed data up to the equivalent wall clock time, i.e., the actual time in the field. Imputations are then based on the current wall clock (real-time) data and previous (historical) data. For each of the three data cases, for each of the two traffic days, ten replicate simulation trials are run to evaluate the impact on travel times on the nine corridor routes, i.e., the three studied side-street approaches and the six mainline routes (Figure 1).

4. RESULTS AND DISCUSSION

The developed univariate and multivariate models for each of the six detectors are used to predict volumes from 3 PM to 6 PM. Table 1 presents the performance error measures for the model predictions.

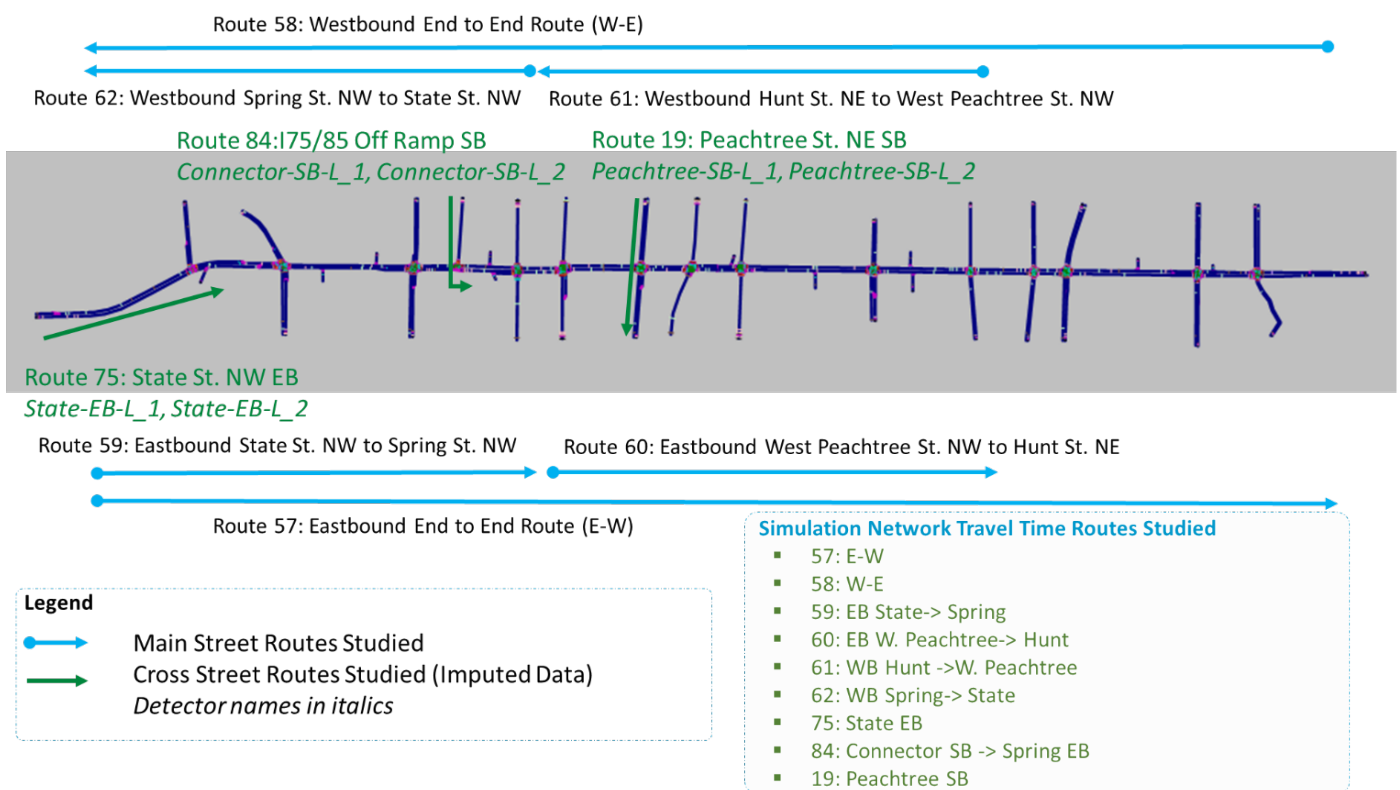


Figure 1. LSTM RNN model developed for three North Ave. corridor approach locations and the nine studied routes [*10*]

Table 1. Error measures for LSTM RNN model predictions for 3 PM to 6 PM.

Detector	Model Type	March 18th (Typical Day)			May 27th (Atypical Day)		
		MAE	RMSE	Std. Dev	MAE	RMSE	Std. Dev
State-EB-L_1	Univariate	5.2	6.3	6.3	20.6	22.7	9.6
State-EB-L_1	*Multivariate**	5.0	6.1	6.1	5.2	6.1	4.0
State-EB-L_2	*Univariate	4.8	6.0	6.0	32.0	33.1	8.4
State-EB-L_2	Multivariate**	5.4	7.24	7.2	16.4	18.2	7.8
Connector-SB-L_1	Univariate	32.1	38.3	21.5	40.5	42.9	16.1
Connector-SB-L_1	*Multivariate**	19.4	26.2	20.7	7.0	8.2	8.0
Connector-SB-L_2	*Univariate**	8.7	11.1	10.7	12.2	15.0	13.0
Connector-SB-L_2	Multivariate	10.0	12.8	12.8	23.1	25.1	9.8
Peachtree-SB-L_1	*Univariate**	6.4	8.6	7.4	9.2	11.1	8.1
Peachtree-SB-L_1	Multivariate	7.2	8.6	7.8	9.8	11.3	7.9
Peachtree-SB-L_2	*Univariate**	6.9	8.4	7.5	4.6	6.0	5.6
Peachtree-SB-L_2	Multivariate	8.3	10.6	8.0	12.5	13.3	5.1

Notes:

* An asterisk indicates lower error values among the two model types on typical day predictions

** Two asterisks indicate lower values among the two model types on atypical day predictions

MAE: mean absolute error, RMSE: root mean square error, Std. Dev: Standard Deviation of Errors

It is observed that the multivariate and univariate predictions tend to be similar on the typical day, with univariate errors often lower than that of multivariate, consistent with previous research findings [*10*]. On the atypical day, it is expected that the multivariate model will provide improved imputation values compared to the univariate model, as the historical data is not consistent with current conditions. This is seen to be true for most detectors. However, at Connector-SB-L_2 and Peachtree-SB-L_2 the univariate prediction errors are observed to be much lower than multivariate prediction errors. For example, for Connector-SB-L_2 Figure 2 shows the plots for observed traffic volumes from midnight to 6 PM for each day along with the univariate and multivariate model predictions from 3 PM to 6 PM.

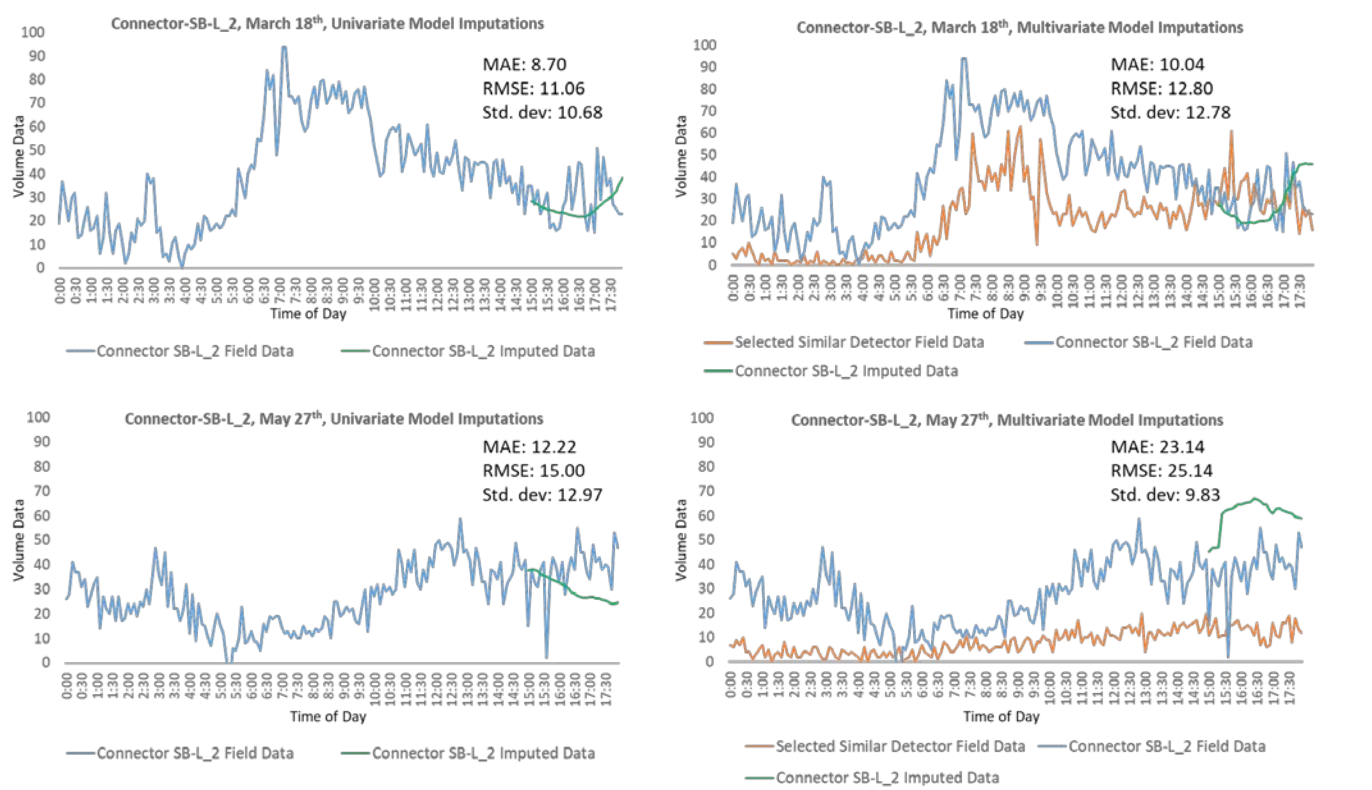


Figure 2. Model predictions for Connector-SB-L_2 starting at 3 PM.

As seen in Figure 2, one reason for the weaker performance of the multivariate model at Connector-SB-L_2 is a poor correlation between the traffic pattern on the matched detector and the given detector. The identification of detectors with similar patterns was undertaken utilizing data streams from multiple "typical" days [*10*]. This raises a possibility that detectors that are reasonably correlated under typical conditions may not be well correlated under atypical conditions. Thus, future improvements to the method may be achieved by identifing different matching detectors for different conditions.

In further exploring the simulation performance given the imputed data it is noted that the field volume data was available in six-minute bins. Therefore, the imputation was also set to generate six-minute binned data. Thus, in the simulation implementation the volume data is entered into the model in six-minute intervals, randomly distributed (shifted poisson distribution interarrival times) over the interval length. However, if the entry link is oversaturated (i.e., a vehicle queue extends to the link entrance point) the new vehicles will not be able to enter the network during their set interval, and will instead be held until space becomes available. Figure 3 shows the volumes that entered the simulation model at Peachtree St. SB for the ten univariate and multivariate replicate trials, for the typical and atypical day scenarios, as well as the imputed volume that sought to enter. A low difference between the imputed and processed entry volume per six-minute interval suggests under-saturated conditions. However, the variation in the March 18th volume entry counts for the multivariate imputations suggests that this approach operates near saturation state during the typical day PM peak period, with the slightly higher multivariate imputed volumes sufficient to create over-saturated conditions. The variation in entry volumes given the univariate imputed values is significantly less, as the imputed volumes are lower than those of the multivariate model (Figure 4). Another clear observation from Figure 4 is that both imputation approaches have a tendency to smooth the volumes relative to the field conditions.

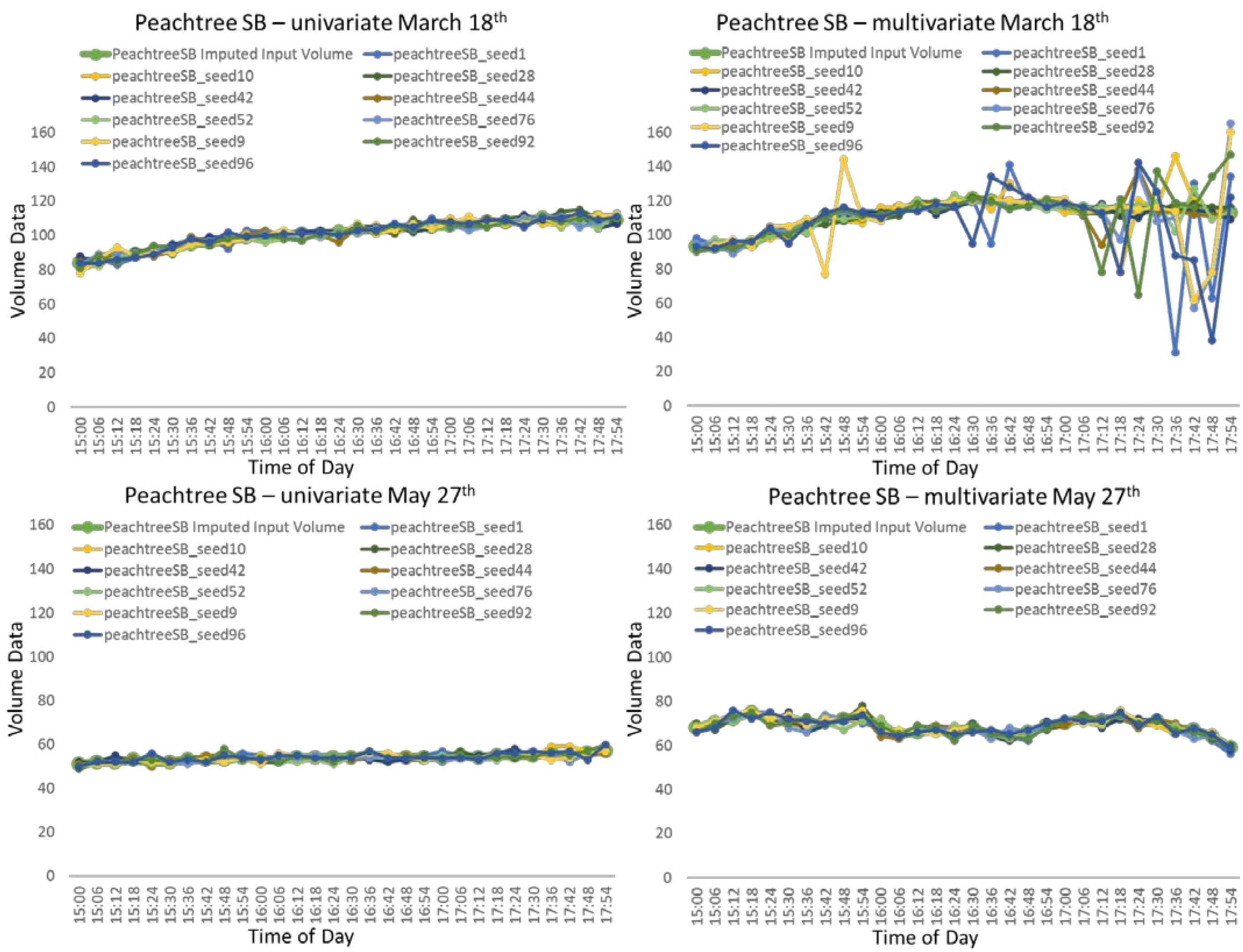


Figure 3**.** Imputed vs Entry volume (10 replicate seeds) for Peachtree St. SB approach volumes in the six-minute bins.

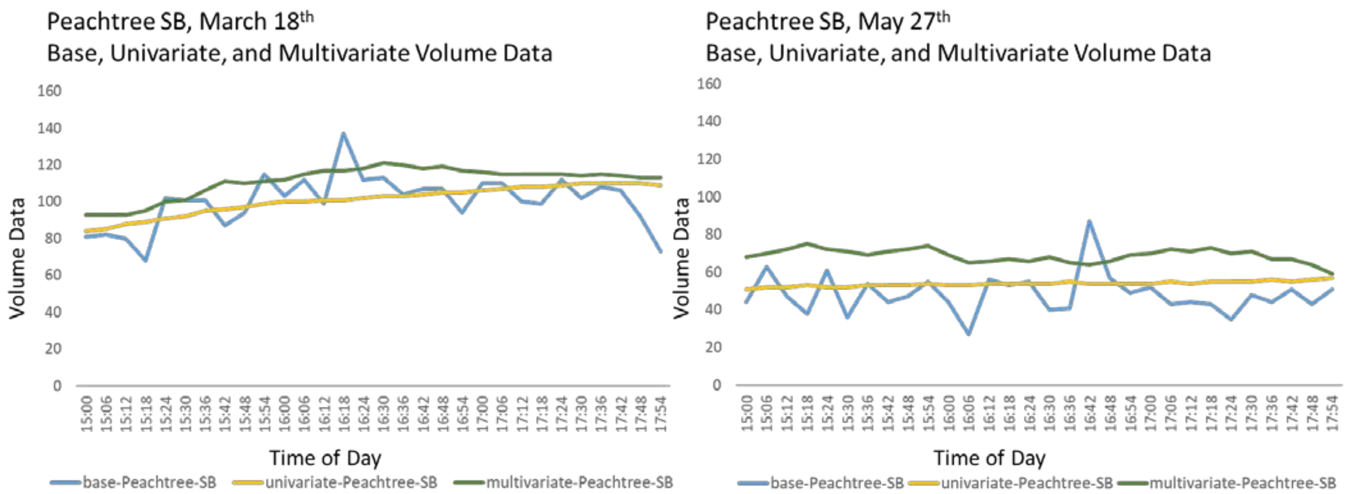
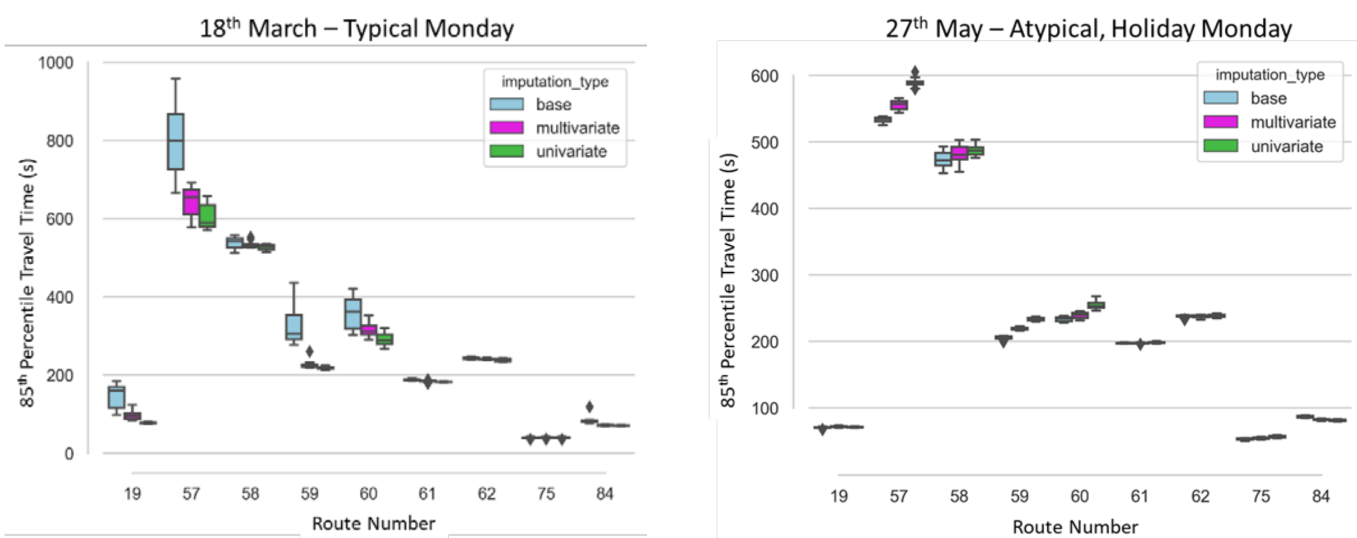


Figure 4. Approach volume for three cases for 3-6 PM at Peachtree St. SB.

4.1 Impact on Digital Twin Generated Travel Time

Figure 5 presents box plots of the 85th percentile travel times obtained from the replicate trials at the nine routes, for the three data input scenarios, for the typical and atypical days. It is observed that for the typical and atypical scenarios, travel times simulated using the multivariate imputations are generally closer to that of the base day than those simulated using the univariate imputations. For 8 of 9 routes under typical conditions multivariate provide closer results than univariate, reducing errors on average by 4%. Under atypical conditions, also for 8 of 9 routes, multivariate provides closer results, reducing errors on average by 3%.

The impact of underlying corridor demand, i.e., saturated vs under-saturated, can be seen on the simulated travel times. For example, lower travel time variation is seen on the atypical traffic day across cases, likely due to the lower holiday traffic. In addition, on Route 19, there are observable travel time differences for the three cases under typical conditions likely due to a saturated traffic state for the PM peak on the typical day versus the under-saturated holiday traffic. The low travel time differences for typical and atypical holiday traffic on Routes 75 and 84 are a result of under-saturated conditions on both days, even though for these routes the univariate imputation provides higher volume estimates than both the multivariate model and base data on the atypical day.



(a) Typical traffic

(b) Atypical traffic day

Figure 5. Boxplots of 85th percentile travel time at the nine study routes for (a) March 18th and (b) Monday, March 27th

Given the under-saturated conditions the overestimated volumes were not sufficiently erroneous to impact travel times. However, over-estimation of volumes is likely a factor that contributed to the increased travel time for the univariate model scenario compared to the base case on Routes of 57, 58, 59, and 60 (Figure 5b). Here, the volume estimation error combined with the underlying near-saturation conditions were sufficient to negatively influence the predicted travel times.

These observations clearly indicate that when developing smart applications, it is critical to identify those locations with the most potential to influence results. Key attributes of the applications, such as identification of a matching detector in the given example, should be assured as well as increased data control and data quality efforts at these locations.

5. CONCLUSIONS AND FUTURE WORK

This effort investigated the impact of the previously developed LSTM RNN multivariate and univariate model imputations on Digital Twin generated travel times. The results indicate that for the studied typical and atypical traffic, the multivariate imputations lead to simulated travel times that are closer to that of the base day. However, additional improvements in the multivariate method may be achieved by improved matching detector selection. Next, the importance of the underlying corridor conditions, i.e., saturation level, is observed. It is demonstrated that when developing smart applications both the imputation methodology and the local conditions must be considered. Specific to this effort, to improve the performance of the LSTM RNN models, future investigations may consider additional atypical training and test data as well as hyperparameter tuning.

6. ACKNOWLEDGMENTS

The information, data, or work presented herein was funded in part by the City of Atlanta (Research Project FC-9930- Smart Cities Traffic Congestion Mitigation Program) and in part by National Center of Sustainable Transportation (NCST). The authors thank the sponsors for their support of this research. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors. This paper does not constitute a standard, specification, or regulation.

REFERENCES

[1] USDOT. *Connected Vehicle Pilot Deployment Program*. n.d. [cited 2020 February 26]; Available from: https://www.its.dot.gov/pilots/.

[2] Frost, A. *Singapore to develop first 5G C-V2X research testbed on NTU campus*. 2019 October 22, 2019 [cited 2020 February 26]; Available from: https://www.traffictechnologytoday.com/news/connected-vehicles-infrastructure/singapore-to-develop-first-5g-c-v2x-research-testbed-on-ntu-campus.html.

[3] CenterForTransportationResearch. *City of Austin Connected Corridors*. 2018 [cited 2020 February 26]; Available from: https://ctr.utexas.edu/nmc/research-2/projects/current-and-ongoing-projects/city-of-austin-connected-corridors/.

[4] HighwaysEngland, *Signs of the future: new technology testbed on the A2 and M2 in Kent*. 2018, Highways England: https://www.gov.uk/government/news.

[5] *California CV Testbed*. n.d. [cited 2020 February 26]; Available from: http://www.caconnectedvehicletestbed.org/index.php/.

[6] INRIX. *INRIX IQ*. n.d. [cited 2020 27 December 2020]; Available from: http://docs.inrix.com/traffic/speed/.

[7] HERE. *HERE Technologies*. n.d. [cited 2020 27 December 2020]; Available from: https://www.here.com/platform/traffic-solutions.

[8] Saroj, A., S. Roy, A. Guin, M. Hunter, and R.M. Fujimoto, *Smart city real-time data-driven transportation simulation*, in *Proceedings of the 2018 Winter Simulation Conference*. 2018, IEEE Press: Gothenburg, Sweden. p. 857–868.

[9] Hunter, M., R. Guensler, A. Guin, A. Saroj, and S. Roy, *Smart Cities Atlanta - North Avenue*, in *City of Atlanta Research Project*. 2019: http://realtime.ce.gatech.edu/RenewAtlanta-GeorgiaTech-Final-Report.pdf. p. 82.

[10] Saroj, A., *Development of a real-time connected corridor data-driven digital twin and data imputation methods*, in *School of Civil and Environmental Engineering*. 2020, Georgia Institute of Technology: https://smartech.gatech.edu/handle/1853/63642.

[11] Kostadinov, S. *How Recurrent Neural Networks work*. 2017 December 2017 [cited 2020 February 22]; Available from: https://towardsdatascience.com/learn-how-recurrent-neural-networks-work-84e975feaaf7.

[12] See, A. *Vanishing gradients and fancy RNNs*. Natural Language Processing with Deep Learning 2019 2019 [cited 2020 July 14]; Available from: https://web.stanford.edu/class/cs224n/slides/cs224n-2019-lecture07-fancy-rnn.pdf.

[13] Hochreiter, S. and J. Schmidhuber, *Long Short-term Memory*. Neural computation, 1997. **9**: p. 1735-80. https://www.mitpressjournals.org/doi/abs/10.1162/neco.1997.9.8.1735

[14] Weber, N. *Why LSTMs stop your gradients from vanishing: A view from the backwards pass*. 2017 November 15 2017 [cited 2020 February 23]; Available from: https://weberna.github.io/blog/2017/11/15/LSTM-Vanishing-Gradients.html.

[15] Yao, Y. and Z. Huang. *Bi-directional LSTM Recurrent Neural Network for Chinese word segmentation*. in *Neural Information Processing*. 2016. Cham: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-319-46681-1_42

Journal Home: https://journals.gmu.edu/index.php/jmms
Mason Journals • Mason Publishing Group
Journal of Modern Mobility Systems (JMMS), Vol. 01(2020), 125-130
Document: https://doi.org/10.13021/jmms.2020.2931

Identifying Road Links and Variables Influencing the Applicability of Variable Speed Limits Using Supervised Machine Learning and Travel Time Data

Sarvani V. Duvvuri, M.S.,[a] Sonu Mathew, Ph.D.,[a] Raghuveer Gouribhatla, M.S.,[a] Srinivas S. Pulugurtha, Ph.D., P.E., F.ASCE[a,*]

[a] *The University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28223; *Corresponding Author, sspulugu@uncc.edu*

Received: 2 Dec. 2020 • **Revised** 29 Dec. 2020 • **Accepted**: 29 Dec. 2020 • **Published**: 29 Dec. 2020

ABSTRACT

With increasing congestion and associated challenges to manage the transportation network, intelligent transportation systems (ITS) have gained popularity due to their data-driven approach and application of advanced technologies. A variable speed limit (VSL) is a popular ITS-based solution which uses dynamic speed limit to promote harmonization along a corridor. However, not much was done in identifying road links and influencing variables for their applicability. Therefore, this paper focuses on examining road link-level data to identify road links and variables influencing the applicability of VSL signs. A multivariate cluster analysis was first used to identify potential road links susceptible to speed variation for the implementation of VSL. A supervised machine learning algorithm, forest-based classification and regression, was then used to model and examine the influence of average annual daily traffic (AADT), historical speed of the road link, and the speeds of upstream and downstream road links on the average speed of the corresponding road link. Modelling and validation were performed using data for Mecklenburg County, North Carolina, USA, for road links including all kinds of speed variation.

Keywords: Intelligent transportation systems, Variable speed limit, Supervised machine learning, Big data.

1 INTRODUCTION

The posted speed limits on roads are typically determined based on the road design, operating speed, geometry, and type of the facility [*1*]. The Federal Highway Administration (FHWA) describes traffic congestion as a direct measure of vehicle speeds [*2*], referred to as speed in this paper. A consistent and significant decrease in speeds on a road link indicates severe recurring congestion on the road link [*3*].

Dynamic message signs with variable speed limits (VSLs) is a widely explored intelligent transportation systems (ITS)-based solution to regulate the speeds on highly congested road segments, in work zone areas, during adverse weather conditions, or during incidents on a road [*4*]. The VSL control strategy also improves mobility and safety in adverse weather conditions [*5*].

The VSLs are estimated dynamically based on the traffic condition and optimized to improve the road capacity. Researchers in the past proposed various algorithms to compute the VSLs. They include simulation-based approach [*6*], cellular transmission models using bottleneck information [*7*], macroscopic simulation [*8*], algorithms like fuzzy logic with simulation-based validation [*9*], and model predictive control [*10*].

One of the most important aspects of the VSLs is the extent to which the speed limit is changed. A significant increase or decrease in the speed limits might raise a concern. Many researchers set thresholds while modelling the speed limit. Abdel-Aty et al. [*6*] used 5 mph increments for the road facilities while Hegyi et al. [*11*] considered a threshold of ± 6.2 mph to ensure safer stream performance. State agencies implementing the VSL signs used thresholds up to 7.5 mph (New Mexico), 30 mph (New Jersey), 10 mph (Washington State), or increments of 10 mph (Nevada) [*4*]. From a safety perspective, the maximum changes to the speed limit of a facility could be up to 10 mph [*6*].

The existing VSL signs use algorithms to generate the speeds needed for the corresponding time of the day and day of the week. Some of the simplest algorithms used include the display of speeds in increments of 5 mph based on the 85th percentile speeds [*12*]. Assigning the algorithm or technique to improve the traffic flow is one of the most common challenges due to its dynamic nature. Further, speeds of the upstream and downstream road links have an influence on a road link speed and should be accounted for in the VSL design process [*12*].

VSL may not be applicable to all the road links. It is important to analyse the patterns in travel times and examine the historical data when computing the speeds for VSL signs. Past research on the dynamic travel time predictions used pattern recognition using the probe data [*8*]. The VSLs from simulation models could be different from what may be observed using the field data. It is, therefore, important to identify the road links which are susceptible to higher variation in speeds using the field data.

Supervised machine learning is designed to forecast using a training dataset and is applicable to even model non-linear relationships. It has the potential to identify the road links with a significant variation in speeds from the posted speed limits and is considered appropriate for this type of "big data" application. Therefore, the objectives of this research are to compute the variability in speeds, identify vulnerable road links, and apply a supervised machine learning algorithm to examine the influence of selected explanatory variables on speed patterns.

2 STUDY AREA, DATA, AND RESEARCH METHOD

Mecklenburg County in the State of North Carolina, USA was considered as the study area for this research. The travel time data and their corresponding network data such as the annual average daily traffic (AADT) and functional class of the road were considered for analysis.

The travel time data was obtained from Regional Integrated Transportation Information System (RITIS) with support from the North Carolina Department of Transportation (NCDOT). The data consists of raw travel times with samples collected at a 1-minute interval for each road link identified by the traffic message channel (TMC) code. The raw travel time data for March of the year 2019 during the peak period was processed using Microsoft SQL Server. The 85th percentile speed and the average speed of considered road links were computed, and the variations were examined for the corresponding analysis hour. Furthermore, data associated with the corresponding upstream and downstream road links were also considered for the analysis.

The research method adopted is two-fold. Firstly, cluster analysis was performed to identify the groups of road links with speed variations by comparing the 85th percentile and average speeds. Secondly, the influence of selected explanatory variables on the average speed of a road link was examined using forest-based classification and regression.

The K-means clustering was used in this research. The algorithm establishes thresholds to minimize the heterogeneity in speeds. It identifies the initial seeds randomly based on the number of allocated clusters, while the other seeds are typically allocated by employing a random component [*13*].

Datasets for the forest-based classification and regression analysis comprised of all the road links, road links with low-speed variation, and road links with high-speed variation. These separate datasets were considered for modelling and analysing the importance of the selected explanatory variables.

The forest-based classification and regression algorithm trains the model data [*14*], estimates the dependent variable (the average speed in this research), and helps understand the speed patterns of roads using the selected explanatory variables. The mechanism of the forest-based classification and regression includes the usage of hundreds of randomly generated trees to predict the average speed. Hence, the result from each tree contributes to the overall accuracy of the model. For higher data points, the tree-based mechanisms are suggested [*14*].

The selection of explanatory variables for analysis and modeling plays a major role in the predictability and understanding their influence on the dependent variable. All the explanatory variables are typically selected to develop a model and assess their influence on the dependent variable in forest-based classification and regression [*15*]. Therefore, the correlation between the explanatory variables was not examined in this research.

The influence of the explanatory variables is computed based on the prediction accuracy using the training dataset. For example, each decision tree in the model uses a certain portion of data to train and generate the outcomes. The remaining data is used to compute the influence and importance of each explanatory variable in predicting the dependent variable, by estimating the decrease in the prediction accuracy [*15*]. In general, a higher value indicates a higher degree of the explanatory variable's importance in the model prediction.

Modelling was performed with 80% of the data and the remaining 20% of the data were used for the validation. The functional class of the road, AADT, historical average speed of the road link, downstream road link average speed, and upstream road link average speed are considered as the selected explanatory variables.

3 RESULTS AND DISCUSSION

Data for 563 road links in the study area were considered for analysis in this research. The study area and the road links are shown as Figure 1. Tables 1 and 2 summarize the descriptive statistics (minimum, median, mean, maximum, and standard deviation) and frequency distribution of the variables considered in this research, respectively.

3.1 Cluster Analysis Results

A total of six clusters were defined by using the optimal R^2 value. The box whisker plot (Figure 2) shows the clusters along with the variations associated with the 85th percentile speed and the average speed for the analysed road links. The low-speed variation comprised of clusters with variation ranging from -7.8 mph to 4.0 mph. The remaining clusters with large variation in the speed on negative side were categorised as "high-speed variation" dataset. The spatial distribution of road links based on the defined clusters are shown in Figure 3.

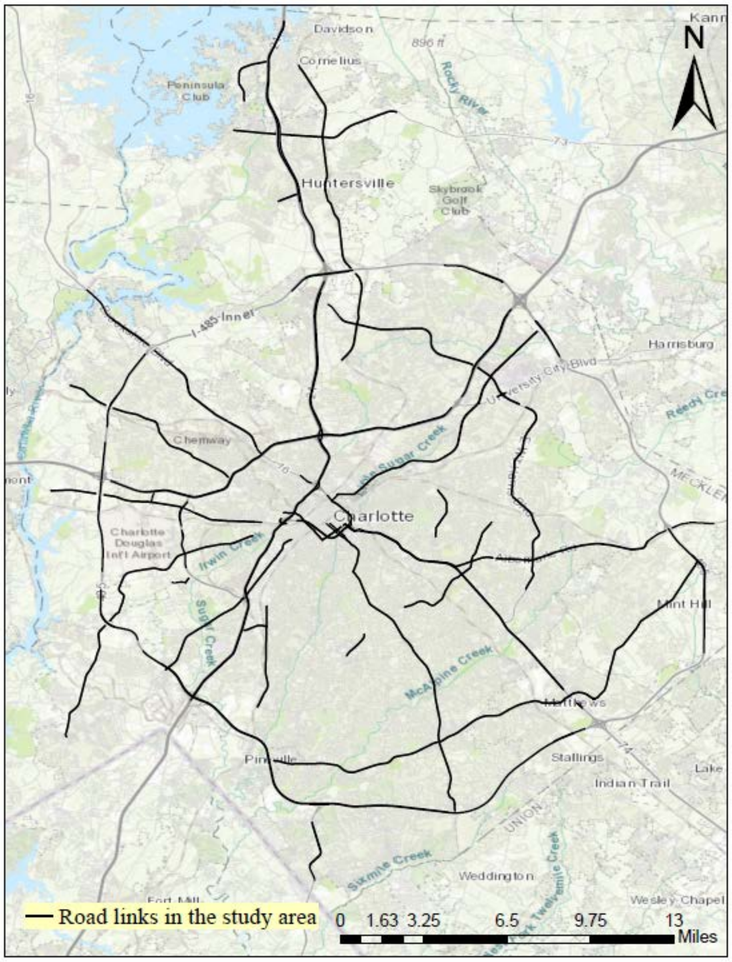


Figure 1. Study area

Table 1. Descriptive statistics of data

Variable	Min.	Median	Mean	Max.	Std. Dev.
Average speed	7.14	40.77	43.83	73.40	16.56
85th percentile speed	14.00	42.00	45.57	70.00	15.47
Difference between the 85th percentile and average speeds	-22.40	0.13	1.74	31.40	7.80
Historical average speed	5.80	40.57	43.84	73.40	16.33
Upstream average speed	7.47	40.00	43.61	73.73	16.63
Upstream reference speed	12.00	42.00	45.37	70.00	15.56
Downstream average speed	7.14	40.47	43.41	74.47	16.44
Downstream reference speed	10.00	42.00	45.33	70.00	15.37
AADT	3700	52000	73160	183000	50626

Table 2. Frequency distribution by facility type

Variable	Categories	Frequency	Percentage
Functional class	1: Interstate	242	42.98
	2: Principal Arterial - Other Freeways and Expressways	15	2.66
	3: Principal Arterial - Other	278	49.38
	4: Minor Arterial	27	4.80
	5: Major Collector	1	0.18
Number of through lanes (in both the travel directions)	2	50	8.88
	3	2	0.36
	4	255	45.29
	5	8	1.42
	6	129	22.91
	8	104	18.47
	10	11	1.95
	12	7	1.24

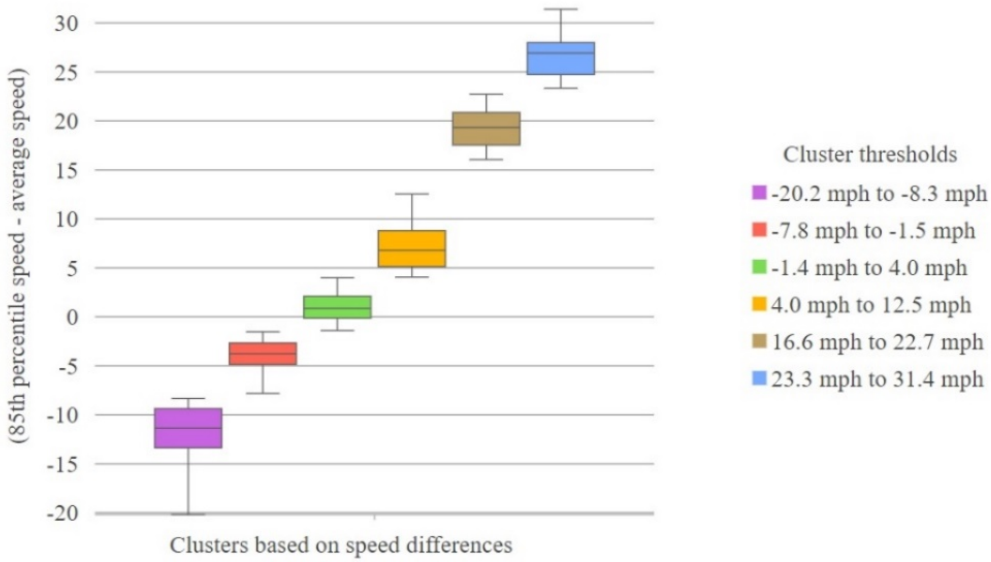


Figure 2. Multivariate cluster analysis results

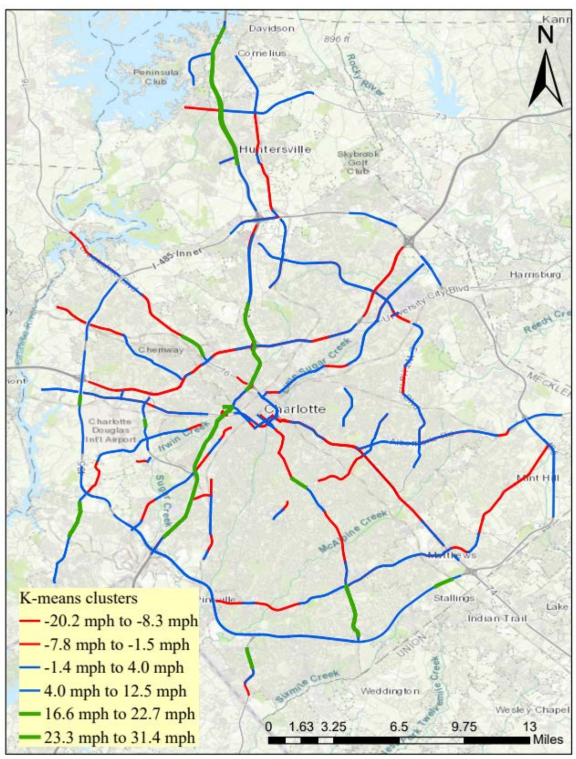


Figure 3. Spatial distribution of clusters in the study area

3.2 Classification and Regression Results

The results from the application of forest-based classification and regression using the three datasets are summarized in tables 3 and 4.

Table 3 shows the importance of the selected explanatory variables from the model results in terms of percentages. The historical average speed is the most important explanatory variable, followed by the upstream and downstream road link average speeds, in the model associated with all road links dataset. In the low-speed variation dataset-based model, the functional class followed by the historical average speed and AADT are the most important explanatory variables. However, the model results from the dataset with high-speed variation dataset indicate that all the selected explanatory variables are important, with the historical average speed of the road link being the most important explanatory variable.

Table 3. Explanatory variables and their importance in terms of percentages

Explanatory variable	Modelling dataset		
	All data	Low-speed variation	High-speed variation
Functional class	7.57	45.05	5.61
AADT	1.25	18.86	14.59
Historical average speed	47.46	22.29	40.49
Upstream average speed	26.48	9.85	20.57
Downstream average speed	17.25	3.96	18.74

Table 4. Predictability results

Parameter / Measure	Modelling dataset		
	All data	Low-speed variation	High-speed variation
R^2	0.94	0.97	0.92
Mean percentage error (%)	-2.96	-1.38	-11.96
Mean absolute percentage error (%)	11.64	6.11	20.33
Root mean square error (in mph)	4.88	3.04	6.77

The predictability results (Table 4) from the forest-based classification and regression indicate a high R^2 value (>0.90) for all the three models (which explains the variability in each dataset). The mean percentage error varied between -1.38% and -11.96%, while the mean absolute percentage error varied between 6.11% and 20.33%. The root mean square error varied between 3.04 mph and 6.77 mph. The errors are highest for the high-speed variation dataset, followed by all the road links dataset. This could be attributed to the low sample size and/or variations in the explanatory variables.

4 CONCLUSIONS

This research explores cluster analysis and the plausible application of machine learning algorithms like the forest-based classification and regression to analyse the speed patterns on road links and assess the applicability of VSLs for congestion mitigation and transportation network management. Travel time data and selected network characteristics for road links in Mecklenburg County were considered in this research. The multivariate cluster analysis was performed to identify groups of road links with varying speeds by comparing the 85th percentile and average speeds. Datasets with all road links as well as road links with low- and high-speed variation were considered to model using the forest-based classification and regression algorithm and examine the influence of the selected explanatory variables.

The functional class of a road and AADT are the most important explanatory variables in the models associated with low- and high-speed variation datasets. However, the functional class of a road and AADT are the least important explanatory variables in the model associated with all the road links dataset. The historical average speed of the road link and upstream road link average speed are the most important explanatory variables irrespective of the dataset considered for modeling in this research.

The R^2 values are high and errors are relatively low, indicating the predictability and potential applicability of supervised machine learning algorithms for determining VSLs. The relatively high errors for high-speed variation dataset indicate that other explanatory variables and more data should be used for analysis and modeling. Furthermore, thresholds for the applicability of VSLs by area type and functional class of a road should be explored in the future.

This research proposes and illustrates the working of a method for identifying vulnerable links and implementing VSLs using travel time data and supervised machine learnig. Researching the applicability of VSLs using larger travel time datasets for even more number of links with varying road and traffic characteristics, by day of the week and time of the day, merits further investigation.

5 ACKNOWLEDGMENTS

The paper is partially based on ideas and information collected for projects funded by the United States Department of Transportation - Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) University Transportation Centers Program (Grant # 69A3551747127) and NCDOT. The authors thank the NCDOT, the city of Charlotte Department of Transportation, and RITIS for providing access to the data needed for this research.

6 DISCLAIMER

This paper is disseminated in the interest of information exchange. The views, opinions, findings, and conclusions reflected in this paper are the responsibility of the authors only and do not represent the official policy or position of the University of North Carolina at Charlotte or other entity. The authors are responsible for the facts and the accuracy of the data presented herein. This paper does not constitute a standard, specification, or regulation.

REFERENCES

1. M. Papageorgiou, M. Ben-Akiva, J. Bottom, P. H. Bovy, S. P. Hoogendoorn, N. B. Hounsell, , ... and M. McDonald, "ITS and traffic management," *Handbooks in Operations Research and Management Science*, vol. 14, pp. 715-774.
2. Cambridge Systematics, "Traffic congestion and reliability: Trends and advanced strategies for congestion mitigation," Report No. FHWA-HOP-05-064, Federal Highway Administration, 2005, Retrieved from https://ops.fhwa.dot.gov/congestion_report/.

3. K. Fitzpatrick, B. Shamburger and D. Fambro, "Design speed, operating speed, and posted speed survey. *Transportation Research Record*, vol. 1523, no. 1, pp. 55-60, 1996.
4. M. Robinson, "Examples of variable speed limit applications," Speed Management Workshop at 79th Annual Meeting of Transportation Research Board, Washington, D.C., 2000.
5. M. Hadiuzzaman, T. Z. Qiu, and X. Y. Lu, "Variable speed limit control design for relieving congestion caused by active bottlenecks," *Journal of Transportation Engineering*, vol. 139, no. 4, pp. 358-370, 2013.
6. M. Abdel-Aty, J. Dilmore, and A. Dhindsa, "Evaluation of variable speed limits for real-time freeway safety improvement," *Accident Analysis & Prevention*, vol. 38, no. 2, pp. 335-345, 2006.
7. M. Hadiuzzaman, and T. Z. Qiu, "Cell transmission model based variable speed limit control for freeways," *Canadian Journal of Civil Engineering*, vol. 40, no. 1, pp. 46-56, 2013.
8. R. Yu and M. Abdel-Aty, "An optimal variable speed limits system to ameliorate traffic safety risk," *Transportation Research Part C: Emerging Technologies*, vol. *46*, pp. 235-246, 2014.
9. A. H. Ghods, A. R. Kian, and M. Tabibi, "A genetic-fuzzy control application to ramp metering and variable speed limit control," In 2007 IEEE International Conf. on Systems, Man and Cybernetics, pp. 1723-1728, Montreal, Que., Canada, 2008.
10. H. Chen, H. A. Rakha, and C. C. McGhee, "Dynamic travel time prediction using pattern recognition," In *20th World Congress on Intelligent Transportation Systems*, TU Delft, 2013.
11. A. Hegyi, B. De Schutter, and J. Hellendoorn, "Optimal Coordination of Variable Speed Limits to Suppress Shock Waves". *IEEE Transactions on Intelligent Transportation Systems*, vol. 6(1), pp. 102-112, 2005.
12. B. Katz, J. Ma, H. Rigdon, K. Sykes, Z. Huang, K. Raboy, and J. Chu, "Synthesis of Variable Speed Limit Signs," Report *No. FHWA-HOP-17-003*, Federal Highway Administration, 2017, Retrieved from https://ops.fhwa.dot.gov/publications/fhwahop17003/fhwahop17003.pdf.
13. ArcGIS Pro, Esri. Multivariate Clustering Analysis, Retrieved from https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/how-multivariate-clustering-works.htm.
14. ArcGIS Pro, Esri, Forest-based Classification and Regression, Spatial Statistics, Retrieved from https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/forestbasedclassificationregression.htm.
15. F. Degenhardt, S. Seifert, and S. Szymczak. "Evaluation of variable selection methods for random forests and omics data sets". Briefings in Bioinformatics, vol. 20(2), pp. 492-503, 2019.

Journal Home: https://journals.gmu.edu/index.php/jmms
Mason Journals ● Mason Publishing Group
Journal of Modern Mobility Systems (JMMS), Vol. 01(2020), 131-137
Document: https://doi.org/10.13021/jmms.2020.2709

Could Value-Based Pricing Improve Economic Sustainability of Bikeshare?

Mohan Venigalla[a,*] , Siddartha Rayaprolu[a], Thomas Brennan[b], Shruthi Kaviti[c]

[a] George Mason University, 1400 University Drive, Fairfax, VA 22030; [b] The College of New Jersey, Armstrong Hall - RM 173, 2000 Pennington RD, Ewing Township, NJ 08628; [c] Institute of Aeronautical Engineering (IARE), Hyderabad 500 043, India; [*]Corresponding author: mvenigal@gmu.edu

Received: 3 Feb. 2020 ● **Revised & Resubmitted** 28 Dec. 2020 ● **Accepted**: 30 Dec. 2020 ● **Published**: 31 Dec. 2020

ABSTRACT

We scrutinize the reactions of casual users of bikesharing services to fare menu, product pricing, and promotion. We hypothesize that by introducing value-based pricing into the fare-option mix, revenues can be increased and therefore enhance the economic sustainability of the bikesharing system. We conducted a controlled experimental survey of 157 current and potential bikeshare users across six cities in the United States. The survey registered the respondents' choice of fare options in two groups: one with a binary choice set (control group) and the other with an additional value-priced choice (experimental group). Evidence points to users' perception of value in bikeshare fare options would contribute to variations in revenues for the same ridership levels. Revenue projections and statistical tests showed that the introduction of value-based pricing options could lead to significant revenue increases. Furthermore, how the fare options are presented to the user would have an impact on users' reception to the value-based pricing options in the product mix. The study results could be useful for numerous bikeshare systems in re-examining their product mixes and how they are presented to the users on websites, mobile apps and kiosk locations.

Keywords: *decoy pricing, value-based pricing, behavioural economics, bikeshare, pricing, micromobility, shared mobility, revenue, ridership*

1 INTRODUCTION

Like transit fare, the cost of the ridership of a bikeshare trip plays a significant role in the mode choice behaviour of users and the system's economic sustainability. While subsidies are important, a healthy farebox recovery is the most essential ingredient for the economic sustainability of transportation services that are operated in the public interest (such as transit and bikesharing services). To this effect, bikeshare service providers routinely make changes to pricing structure and fare menus for all user-types.

When making changes to bikeshare pricing or introducing a new fare option, it is important to consider users' perceptions of the economic value of the new product on its own and in relation to the prices of other products in the 'product line' [*1*]. Very few studies addressed the infrastructure and pricing policy implications on general cycling usage [*3*]. Despite the importance of pricing to bikeshare patronage, only a limited number of studies focused on the impact of a well-defined pricing strategy on revenue and ridership [*1*,*2*,*5*-*8*].

1.1 Objective

In the marketing parlance, the essential elements of a marketing plan, namely: *product*, *price*, *place*, and *promotion* (known as 4Ps of the marketing mix), help develop marketing strategies and tactics [*9*]. We theorize that choices of bikeshare users, like that of consumers of any other commercial product, are influenced by perception of value and behavioural economics. In this research, we focus on service options (products), pricing, and presentation of public bikesharing systems.

We hypothesize that, by introducing value-based pricing options into the product mix, bikesharing revenues can be increased. We test this hypothesis by conducting a controlled experimental survey of 157 current and potential bikeshare users across six cities in the United States. We also examine the promotional aspects of bikeshare fare options by testing the revenue impacts of user choices when the same product menus are presented in different formats.

2 MOTIVATION

Bikeshare fare options and subscription plans for casual users (temporary users with no long-term commitment) and members (also known as subscribers) vary from system to system. They also change over time. The fare options for these two prominent user types represent the 'product' in the 4P-concept as applied to bikesharing. The market share of ridership and revenue for members and casual users varies across systems. For example, members account for 72% of ridership and 29% of the revenue at Capital Bikeshare (CaBi), while casual users account for only 28% ridership and yet 71% of its revenue [*8*]. Similarly, the revenue split between members and casual users for Citi Bike (NY) is 32.3% and 67.7% [*10*].

Consumer behaviour in transportation mode-choice was first modelled in the early 1970s [*11*,*12*]. Though consumer-pricing research shows that product(s) and pricing mix is an essential determinant of customer patronage and revenues, consumer-oriented research in pricing bikesharing services is rare. A non-scientific polling of three bikeshare providers in the USA indicated that decisions related to bikeshare product lines and pricing are often arbitrary, have minimal or no scientific basis, and based on local political considerations. At the time of conducting this research, scooter-sharing is in the nascent stages of deployment and its impact on bikesharing was not considered.

For this research, we define value-based pricing as the strategic process of pricing a product or service that offers economic value to consumers. The value may be absolute or relative to other products in the choice set, and it may be real or perceived. Because it considers the customer perspective, value-based pricing increases the likelihood of maximizing revenues from the same set of customers simply by altering their product-selection from the given product mix [*13*].

Table 1 presents bikeshare product lines and their prices (fare options) at the seven largest bikeshare systems in the United States. The websites of all these systems emphasize that the annual membership is the "best value" option for users. However, only two systems offer a 3-day pass (valid for 36-hours after purchase), and the monthly pass option is not available at two of the systems. Although casual users account for a significant share of revenues [*2*,*8*], none of these systems appear to be emphasizing a "best value" option for casual users.

These observations led us to asking the following questions:

1. Does the product-mix itself have an impact on which option bikeshare consumers (especially casual users) choose and, therefore, on system-wide revenues?
2. If product-mix has an impact, what would be a value-based pricing strategy and the optimal product mix for bikeshare services?

Kaviti et al. partially addressed these issues in their study on the impact of the launch of a single trip fare (STF) product for $2/trip on the revenue and ridership of CaBi at the jurisdiction level in the metro Washington DC area [*5*]. The study showed a significant increase in casual user ridership after the introduction of STF. In an analysis conducted at individual station-level, Venigalla et al. observed that the introduction of STF led to a significant increase in casual user ridership, coupled with a positive growth rate; and a significant decrease in revenue per ride with a negative growth rate [*1*]. The research presented in this paper builds on the studies by Kaviti et al. [*5*] and Venigalla et al. [*1*].

3 REACTIONS TO VALUE-BASED PRICING

A notable gap exists in literature with respect to understanding user behaviour towards bikeshare product pricing. Literature in consumer product pricing presents numerous examples of how to expose the relative value through such behavioural economics considerations as decoy pricing [*22-24*], value-based pricing, and menu-engineering [*25*,*26*]. For example, restaurants design their wine menu based on a widely known discovery that the second lowest-priced wine is usually the best seller on the wine list [*27*]. Ariely illustrated that by introducing a decoy option within the product mix might increase revenues [*28*]. Ariely's experiment (Table 2) divided the subscription options for The Economist magazine into two choice sets (ACS1 & ACS2). Two separate groups of 100 students at Massachusetts Institution of Technology were asked to select a subscription from one of the choice sets given to them.

As the results of the Ariely's experiment indicate, the hypothetical revenue from ACS2 is 43% higher than that of ACS1. In this experiment, the mere introduction of a decoy in ACS2 has unlocked the value in one of the two other options. Thus, it is conceivable that a carefully designed choice-set of fares will influence the behaviour and choices of bikeshare consumers towards increasing the revenues. We tested two versions of a controlled survey that is similar to Ariely's experiment. Both versions had the same questions on user demographics (gender and income), prior experience with bikeshare, and willingness to pay for a regular subscription in both versions. However, the choice-set (CS) of fare options is different for both versions.

To assess users' perception of the relative value of fare op, the following question and the associated information were included in the survey form as the lead to a choice set (CS).

If bikeshare is/were available in the city where you work OR in the city you are visiting for sightseeing for a weekend, which fare option would you choose? Assume you can pretty much go wherever you want using bikeshare, and the weather is not an issue. Also, there will be a usage fee for usage above 30 min.

Table 1 Product lines and prices at the seven largest bikeshare systems in the USA

	For Casual Users			Subscription Membership	
Public Bikeshare System	Single Trip Fare (STF)	24-hour pass (Daily Pass)	3-day pass (Multiday Pass)	30-day pass (Monthly Pass)	365-day pass (Annual Pass)
CitiBike[1,4] (New York, NY)	$3.00	$12.00	$24.00	NA	$169.00
Divvy[4] (Chicago, IL)	$3.00	$15.00	NA	NA	$99.00
Capital Bikeshare[4], CaBi (Washington, DC)	$2.00	$8.00	$17.00	$28.00	$85.00
Metro (Los Angeles, CA)	$1.75	$5.00	NA	$17.00	$150.00
Blue Bikes[2,4] (Boston, MA)	$2.50	$10.00	NA	$20.00	$99.00
Nice Ride[4] (Minneapolis, MN)	$2.00	$6.00	NA	NA	$75.00
Bay Wheels[3,4] (San Francisco, CA)	$2.00	NA	NA	$15.00	$149.00

[1] Operates in New York City and Jersey City, NJ. The largest Bikeshare provider in the USA
[2] Previously known as Hubway. Rebranded as Blue Bikes in March 2018
[3] Previously known as Go Bike (by Ford) and rebranded in June 2019 as Bay Wheels
[4] Operated by Lyft or its subsidiary company Motivate for the bikeshare provider in the city/region

(Source: Bay Wheels [*15*]; Blue Bikes [*16*]; Capital Bikeshare [*17*]; Citi Bike [*18*]; Divvy [*19*]; Metro Bikeshare [*20*]; and Nice Ride [*21*])

Table 2 Ariely's experiment on decoy pricing for subscription to The Economist magazine

Subscription Option	Description	Annual Price	Survey responses	
			ACS1	ACS2
1	Web only subscription to economist.com	$59	68	16
2*	Print only subscription to "The Economist"	$125	Not given	0
3	Subscription to printed copy of "The Economist" + web subscription to economist.com	$125	32	84
*Decoy option	Total revenue Σ (Price x Responses)		$8,012	$11,444

(Source: Ariely [28])

The list of fare options presented in the choice sets is described in the survey forms as follows:

1. *A bunch of single-trips, each 30-min trip costing $2*
2. *24-hour pass for an unlimited number of 30-min trips, costing $8*
3. *3-day pass for an unlimited number of 30-min trips, costing $17*

Version 1 (CS-1) displays only fare options 1 and 2. In the real world, it replicates the choice sets (not the prices) that are currently available for riders at Divvy, Metro, Blue Bikes, and Nice Ride systems (Table 1). Version 2 (CS-2), which displays all three options, replicates the choice sets available at CaBi and Citi Bike.

Both versions of the survey were randomly and evenly distributed to the attendees during a lecture series on bikeshare pricing at six different universities across USA (Table 3). Even distribution of respondents between the two choice sets would ensure equal allocation of total ridership between the sets. The respondents (n=157) included full- and part-time students, faculty, staff, and other seminar attendees at the six universities. When taking the survey, the respondents did not know that two versions of the questionnaire exist. Analysis (χ^2 test results) presented in Table 3 show that the respondent choices are independent of their gender, income, prior bikeshare experience, or location.

For estimating revenues using the survey data, we assumed that a typical casual user makes an upfront decision to accommodate his/her travel needs with only one of the fare options available, with a limit of three or fewer trips. The reasoning for limiting the number of single trips (ST) purchases to three is that purchasing a 24-hour pass for $8 for unlimited rides per day would make more economic sense than purchasing four or more single trips at $2 each. Whichever may be the fare option purchased, some casual users might only use bikeshare once (one single trip, or OST), while others may take dual single trips (DST) or triple single trip rides (TST). Table 4 illustrates normalized revenues for the DST scenario.

Revenue estimates were made for three extreme cases in which every STF buyer would make either only one, two or three trips (i.e., 100% of STF trips would be either OST, DST or TST). Figure **1** illustrates revenue estimates for OST and TST (calculations are shown in Table 4 for DST). For each of the three scenarios and at each location, revenue estimates for the choice set with the 3-day pass option (CS-2) are significantly higher (α = 5%) than estimates for CS-1. The observed increase in estimated revenues ranges from 25% to 84%. The 95% confidence interval band for normalized revenues narrows and converges towards the estimated mean as the number of single trips increases from one to three (Figure 2).

These observations imply that the value-based pricing strategy for bikeshare pricing has a consistently positive impact on revenues, an impact that is comparable to that of decoy pricing. While decoy pricing may be perceived as 'deceptive marketing' or even 'profiteering,' value-based pricing has the potential to be well received by bikeshare users.

4 PRESENTATION OF FARE OPTIONS

The casual bikeshare users in the United States typically receive information on the product mix through websites, kiosks, and mobile apps. For the presentation aspect (the fourth P in 4Ps of the marketing mix), we hypothesized that, just as the composition of the product mix affects revenues, presentation of fare options at various points of sale may also impact revenues.

To test this hypothesis, two fictional web designs were developed. The first design mimics actual fare-selection screens on websites of Capital Bikeshare and Citibike. This design requires the user to navigate to a second page to discover a value-priced option. The second design displays all three casual fare options on the same page (Figure 3).

In a short 30-second survey, potential casual users at four different venues were asked to choose a fare from the two designs. To avoid any sample size bias, an attempt was made to distribute the sample evenly between two designs. As the results in Table 5 indicate, user selection of fare option between the two designs is independent of the location of the survey but dependent on the presentation of choices.

The results of this experiment (n = 73) at four different locations indicated estimated revenues with Design 2 are 13% to 149% higher than the estimates for Design 2. On an aggregate basis, the revenue increase with Design 2 is 43% over Design 1. Chi-Square test indicates that the consistent increase in revenues for Design 2 are independent of location of the survey.

Table 3 Summary statistics of user choices of bikeshare fare options

Class Variables		Fare Choice Set 1					Fare Choice Set 2					
		DP[7]	STF[8]	Total	χ^2	p	DP[7]	3DP[9]	STF[8]	Total	χ^2	p
Gender	Female	12	6	18	1.721	0.181	8	2	7	17	1.348	0.509
	Male	28	34	62			24	15	21	60		
Income	<$35k	34	35	69	0	1	21	11	23	55	2.479	0.289
	>$35k	6	5	11			11	6	5	22		
Prior user?	No	30	29	59	0	1	26	12	23	61	0.995	0.608
	Yes	10	11	21			6	5	5	16		
Sample Location	UA[1]	13	11	24	4.224	0.518	10	5	10	25	6.103	0.806
	UNCC[2]	11	10	21			6	5	7	18		
	Clemson[3]	2	3	5			2	1	4	7		
	GMU[4]	5	10	15			6	4	3	13		
	Memphis[5]	6	2	8			5	1	1	7		
	TCNJ[6]	3	4	7			3	1	3	7		
Overall Sample Total		40	40	80			32	17	28	77		

[1]University of Alabama at Tuscaloosa; [2]University of North Carolina at Charlotte; [3]Clemson University; [4]George Mason University; [5]University of Memphis; [6]The College of New Jersey

[7] DP: 24-Hour or Day Pass ($8); [8] STF: Single Trip ($2); [9] 3DP: 3-Day Pass ($17)

Interpretation example: At a significance level (α) = 0.05, the p-value of 0.181 indicates that the user choice of fare option is independent of the gender of the respondent

Conclusion: Gender, income, prior usage, and location of the sample have no significant influence on the respondent choices

Table 4 Estimated revenues from 100 casual users with each single-trip buyer making dual single trips (DST)

Location and Aggregate Statistics	Fare Choice Set 1			Fare Choice Set 2				Percent Increase with CS-2
	24-hour pass ($8)	Single trips ($2/trip)	Total	24-hour pass ($8)	3-day pass ($17)	Single trips ($2/trip)	Total	
University of Alabama	$433	$183	$616[a]	$320	$340	$160	$820	33.0%
UNC Charlotte	$419	$191	$610	$267	$472	$156	$894	46.7%
Clemson University	$320	$240	$560	$229	$243	$229	$700	25.0%
George Mason University	$267	$267	$533	$369	$523	$92	$985	84.6%
University of Memphis	$600	$100	$700	$571	$243	$57	$871	24.5%
The College of New Jersey	$343	$229	$571	$343	$243	$171	$757	32.5%
Totals at all locations	***$400***	***$200***	***$600***	***$332***	***$ 375***	***$145***	***$853***	***42.2%***
Mean	$397	$202	$599	$350	$344	$144	$838	40.0%
SE	$117	$59	$59	$120	$126	$61	$102	
Lower 95% CI	$274	$140	$537	$224	$212	$80	$731	36.2%
Upper 95% CI	$520	$263	$660	$476	$476	$208	$945	43.1%

This illustration assumes all users opting for single trips would purchase two single trips at $2 each. Confidence intervals as based on the t-distribution assumption for the sample.

[a] Example calculation: Of the 24 respondents of choice set 1 at University of Alabama, 13 and 11 opted for 24-hr pass and single-trip, respectively. Thus, for this case revenue for 100 users when choosing from CS-1 would be: 100 × [(13 ÷ 24) × $8 per pass + (11÷24) × $2 per trip × 2 trips by each user] = $433 + $183 = $616

5 CONCLUSIONS AND DISCUSSION

The research shows that, for a given ridership level, changes to fare options could result in significant variations in revenues. The changes are attributable to bikeshare users' perception of value among available fare options. Statistically significant revenue increases are feasible with a fare-choice set containing an additional value-priced option when compared to a binary choice set.

Though the range of projected revenue increases attributable to value-pricing is rather wide (25%-84%), the experiment underscores the point that the mere introduction of a value-based pricing option may have a consistently positive and statistically significant impact on revenue. Similarly, the second experiment on presentation of fare options to users demonstrates that the user's choice is influenced by the presentation of product menu at points of sale, such as websites, kiosks, and mobile apps.

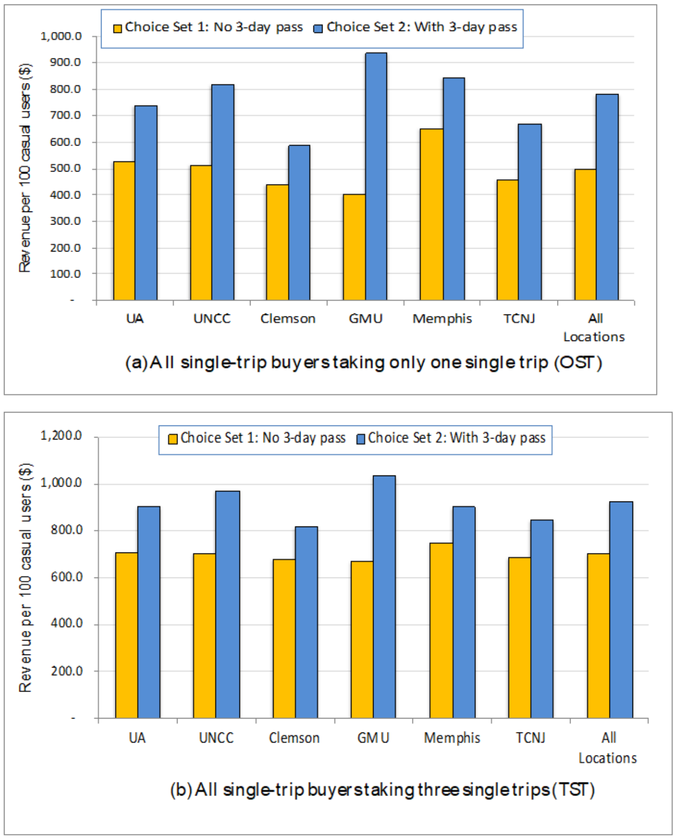


Figure 1 Revenues per 100 casual users with and without value-based price option

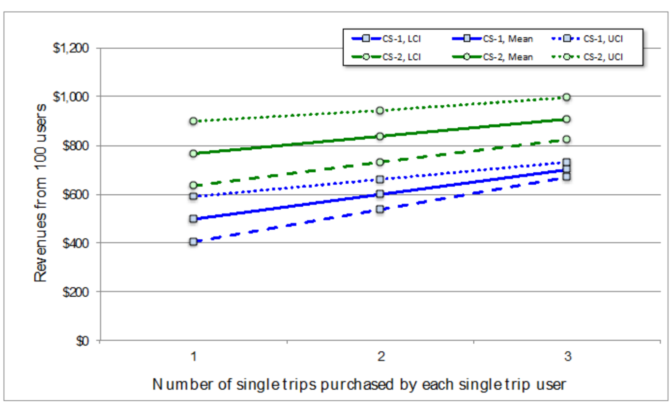


Figure 2 Revenues per 100 users based on the number of single trip fares purchased by users

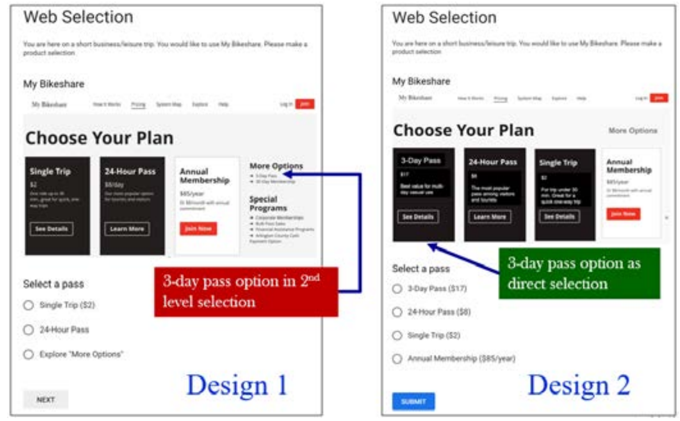


Figure 3 Alternative designs for the web-based presentation of value-based options to users

Table 5 User selection of fare options for two alternate designs

Location	Fare	Design 1 *n*	Design 1 %	Design 2 *n*	Design 2 %	Change in Rev. with Design 2
1.	Total	7	100%	6	100%	
	NS	0	0%	0	0%	
	STF	3	43%	2	33%	
	DP	1	14%	1	17%	
	3DP	3	43%	3	50%	
	Rev.	*$928*		*$1,050*		*13%*
2.	Total	14	%	11	100%	
	NS	1	7%	4	36%	
	STF	8	57%	3	27%	
	DP	4	29%	1	9%	
	3DP	1	7%	3	27%	
	Rev.	*$464*		*$591*		*27%*
3.	Total	9	100%	11	100%	
	NS	0	0%	0	0%	
	STF	5	71%	4	67%	
	DP	3	43%	2	33%	
	3DP	1	14%	5	83%	
	Rev.	*$729*		*$1,817*		*149%*
4.	n	6		9		
	NS	0	0%	2	22%	
	STF	3	50%	3	33%	
	DP	3	50%	0	0%	
	3DP	0	0%	4	44%	
	Rev.	*$500*		*$822*		*64%*
All	n	36		37		
	NS	1	3%	6	16%	
	STF	19	53%	12	32%	
	DP	11	31%	4	11%	
	3DP	5	14%	15	41%	
	Rev.	*$586*		*$841*		*43%*

<u>Survey Locations</u>
1. TRB Annual Meeting, Washington DC
2. University of California, Irvine, CA
3. George Mason University (GMU-1)
4. George Mason University (GMU-2)

n: Sample size
NS: No Selection
STF: Single Trip ($2)
DP: 24-Hour Pass ($8)
MDP: 3-Day Pass ($17)
Rev.: Revenue/100 Riders

Mason Publishing

5.1 Key Takeaway

The key takeaway from this research is that a scientific method that leverages the concepts of consumer pricing research and behavioural economics to set bikeshare pricing could significantly increase revenues from casual users of bikeshare. Changes to fare products can be tested with a simple survey of carefully sampled potential users using the methods described in this paper. That is, bikeshare systems across the world could use the methodology and/or results of this study in strategizing and redesigning product-mix; product-testing and then presenting various price options for bikeshare users. For example, five of the seven largest bikeshare systems in the USA (Table 1) could potentially increase revenues from their casual users by simply introducing a multi-day pass.

It should be noted that the conclusions of this research are subject to a few limitations. The respondents in the sample are not from diverse population groups. The research found that income has no impact on users' choice, which is akin to stated preference. However, income influence may have been absent in the responses as they were based on a hypothetical situation which does not actually involve spending money (i.e., a stated preference instead of a revealed choice). Though the research only shows promise of improving economic sustainability through increased revenues, more work is needed in this regard for establishing the suitable set of pricing options for a given bikesharing service.

REFERENCES

[1] Venigalla, M.M., T. Brennan, S. Rayaprolu, & S. Kaviti. (2020a). Increasing Bikeshare Revenue through Value-Based Pricing: Lessons from Behavioral Economics. 99th Annual Meeting of the Transportation Research Board. National Research Council. Washington DC. (Jan 12-15, 2020).

[2] Venigalla, M.M., Kaviti, S., & Brennan, T. (2020b). Impact of bikesharing pricing policies on usage and revenue: An evaluation through curation of large datasets from revenue transactions and trips. Journal of Big Data Analytics in Transportation (in print).

[3] Pucher, J., Dill, J., & Handy, S. (2010). Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine*, *50*, S106–S125. https://doi.org/10.1016/J.YPMED.2009.07.028

[4] de Nazelle, A., Nieuwenhuijsen, M. J., Antó, J. M., Brauer, M., Briggs, D., Braun-Fahrlander, C., … Lebret, E. (2011). Improving health through policies that promote active travel: A review of evidence to support integrated health impact assessment. *Environment International*, *37*(4), 766–777. https://doi.org/10.1016/J.ENVINT.2011.02.003

[5] Kaviti, S., Venigalla, M. M., Zhu, S., Lucas, K., & Brodie, S. (2018). Impact of pricing and transit disruptions on bikeshare ridership and revenue. *Transportation*, 1–22. https://doi.org/10.1007/s11116-018-9904-5

[6] Kaviti, S., & Venigalla, M. M. (2019). Assessing service and price sensitivities, and pivot elasticities of public bikeshare system users through monadic design and ordered logit regression. *Transportation Research Interdisciplinary Perspectives, 1(1)*, 100015. https://doi.org/10.1016/J.TRIP.2019.100015

[7] Kaviti, S., Venigalla, M. M., & Lucas, K. (2019). Travel behavior and price preferences of bikesharing members and casual users: A Capital Bikeshare perspective. *Travel Behaviour and Society*, *15*, 133–145. https://doi.org/10.1016/J.TBS.2019.02.004

[8] Venigalla, M.M., Kaviti, S., Pierce, W. and Zhu, S. (2018). Analysis of Single-trip fare data for Capital Bikeshare. *District Department of Transportation (DDOT), Final Report*.

[9] McCarthy, E.J., Shapiro, S.J., & Perrealt, W.D. (1979). Basic marketing. (pp. 29-33). Irwin-Dorsey.

[10] Citi Bike (2019b). December 2018 Monthly Report. Retrieved August 8, 2019 from http://citibikenyc.com/system-data/operating-reports

[11] McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior. In P. Zarembka, ed., Frontiers in Econometrics, Academic Press, New York, pp. 105– 142.

[12] Lovelock, C.H. (1975). "Researching and Modeling Consumer Choice Behavior in Urban Transportation", in NA - Advances in Consumer Research Volume 02, eds. M.J Schlinger and A. Abor, MI: *Association for Consumer Research, Pages: 851-862.*

[13] Hinterhuber, A. (2008). Customer value-based pricing strategies: why companies resist. *Journal of business strategy, 29(4), 41-50.*

[14] Venigalla, M., Kaviti, S., & Brennan, T. (2020). Impact of Bikesharing Pricing Policies on Usage and Revenue: An Evaluation Through Curation of Large Datasets from Revenue Transactions and Trips. *Journal of Big Data Analytics in Transportation*, 1-16.

[15] Bay Wheels (2020). Bay Wheels Pricing. Retrieved January 30, 2020 from http://lyft.com/bikes/bay-wheels/pricing

[16] Blue Bikes (2020). Choose Your Plan. Retrieved January 30, 2020 from http://bluebikes.com/pricing

[17] Capital Bikeshare (2020). Choose Your Plan. January 30, 2020 from http://capitalbikeshare.com/pricing

[18] Citi Bike (2020). Choose Your Plan. Retrieved January 30, 2020 from http://citibikenyc.com/pricing

[19] Divvy (2019). Choose Your Plan. Retrieved August 8, 2019 from http://divvybikes.com/pricing

[20] Metro Bikeshare (2020). Pricing. Retrieved January 30, 2020 from http://bikeshare.metro.net/pricing/

[21] Nice Ride (2020) Choose Your Plan. Retrieved January 30, 2020 from http://niceridemn.com/pricing

[22] Boz, H., Arslan, A., & Koc, E. (2017). Neuromarketing aspect of tourısm pricing psychology. *Tourism Management Perspectives*, *23*, 119–128. https://doi.org/10.1016/J.TMP.2017.06.002

[23] Gonzalez-Prieto, D., Sallan, J. M., Simo, P., & Carrion, R. (2013). Effects of the addition of simple and double decoys on the purchasing process of airline tickets. *Journal of Air Transport Management*, *29*, 39–45. https://doi.org/10.1016/J.JAIRTRAMAN.2013.02.002

[24] Huber, J., Payne, J. W., & Puto, C. (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research*, *9*(1), 90. https://doi.org/10.1086/208899

[25] Heide, M., White, C., Gr∅nhaug, K., & Østrem, T. M. (2008). Pricing Strategies in the Restaurant Industry. *Scandinavian Journal of Hospitality and Tourism*, *8*(3), 251–269. https://doi.org/10.1080/15022250802451065

[26] Kasavana, M. L., Smith, D. I., & Schmidgall, R. S. (1990). *Menu engineering : a practical guide to menu analysis*. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US19920039139

[27] McFadden, D., Machina, M. J., & Baron, J. (1999). Rationality for Economists? In *Elicitation of Preferences* (pp. 73–110). https://doi.org/10.1007/978-94-017-1406-8_4

[28] Ariely, D. (2008). *Predictably irrational: the hidden forces that shape our decisions*. Retrieved from https://search.proquest.com/docview/235821459?pq-origsite=gscholar

[29] Rayaprolu, R., & Venigalla, M.M. (2020). Motivations and Mode-choice Behavior of Micromobility Users in Washington, DC. *Journal of Modern Mobility Systems, 1,* (pp. 110-118).