Using Mobile Probes to Inform and Measure the Effectiveness of Traffic Control Strategies on Urban Networks



Morgan State University The Pennsylvania State University University of Maryland University of Virginia Virginia Polytechnic Institute & State University West Virginia University

The Pennsylvania State University The Thomas D. Larson Pennsylvania Transportation Institute Transportation Research Building & University Park, PA 16802-4710 Phone: 814-865-1891 & Fax: 814-863-3707 www.mautc.psu.edu

Using mobile probes to inform and measure the effectiveness of traffic control strategies on urban networks

Prepared for:

Mid-Atlantic Universities Transportation Center

Prepared by:

Vikash V. Gayah Assistant Professor

Andrew S. Nagle Graduate Research Assistant

Department of Civil and Environmental Engineering Pennsylvania State University 217 Sackett Building University Park, PA 16802 <u>gayah@engr.psu.edu</u>

Hesham Rakha Samuel Reynolds Pritchard Professor of Engineering

> **Jianhe Du** Senior Research Associate

Center for Sustainable Mobility, Virginia Tech Transportation Institute 3500 Transportation Research Plaza Blacksburg, VA 24061 <u>hrakha@vt.edu</u>

June 2015

Technical	Report	Documentation	Page
reennour	roport	Dooumontation	. ugu

rechnical Report Documentation rage						
1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.				
MAUTC 2012-02						
4. Title and Subtitle Using mobile probes to inform and measur strategies on urban networks	 5. Report Date July 1, 2015 6. Performing Organization Code 					
7. Author(s)	8. Performing Organization Report No.					
vikash v. Gayan, Hesham Rakha, Jianne	Du, Andrew S. Nagle					
9. Performing Organization Name and A	ddress	10. Work Unit No. (TRAIS)				
The Pennsylvania State University University Park, PA 16802		11. Contract or Grant No.				
Virginia Polytechnic Institute & State Unive Blacksburg, VA 24061	DRTR12-G-UTC03					
12. Sponsoring Agency Name and Addr	ess	13. Type of Report and Period Covered				
US Department of Transportation Research & Innovative Technology Admin	Final Report: 8/1/2012 – 6/30/2015					
UTC Program, RDT-30 1200 New Jersey Ave., SE Washington, DC 20590		14. Sponsoring Agency Code				
15. Supplementary Notes						
16. Abstract						
Urban traffic congestion is a problem that plagues many cities in the United States. Testing strategies to alleviate this congestion is especially challenging due to the difficulty of modeling complex urban traffic networks. However, recent work has shown that these complicated systems can be modeled in relatively simple ways by leveraging consistent relationships that exist between network-wide averages of pertinent traffic properties, such as average network flow, network density and the rate at which trips are completed. Using these "macroscopic" traffic models, various control strategies can be developed to mitigate congestion and improve network performance. However, the effectiveness of many of these strategies depends on the ability to estimate traffic conditions on the network in real-time. This jointly proposed research between Penn State and Virgin Tech investigated how real-time mobile vehicle probes can be combined with macroscopic urban traffic models to inform more efficient network-wide traffic control strategies. Additionally, this work will examine how the effectiveness of these strategies can be directly measured in the field using only mobile vehicle probe data. These two efforts can lead to more efficient control of downtown traffic networks and a reduction in vehicular delay during rush hour periods.						
17. Key Words Macroscopic fundamental diagram; urban to vehicles; traffic state estimation	18. Distribution Statement No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161					

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	122	N/A

Form DOT F 1700.7

Disclaimer

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

TABLE OF CONTENTS

Table of Contentsiii
Study background1
Paper 1: Using mobile probe data and the Macroscopic Fundamental Diagram to estimate network densities: Tests using microsimulation (published in <i>Transportation Research Record</i>)
Paper 2: The accuracy of network-wide traffic state estimations using mobile probe data (published in <i>Transportation Research Record</i>)
Paper 3: Deriving Macroscopic Fundamental Diagrams from probe data: Issues and proposed solutions (under consideration for publication)
Paper 4: Comparing the use of link and probe data to inform perimeter metering control (in the proceedings of the 94 th Annual Meeting of the Transportation Research Board)
Paper 5: Design and evaluation of network control strategies using the Macroscopic Fundamental Diagram (in the proceedings of the <i>2015 IEEE Intelligent Transportation</i> <i>Systems Conference</i>)
REFERENCES
APPENDIX A1
APPENDIX B22
APPENDIX C47
APPENDIX D75
APPENDIX E97

STUDY BACKGROUND

Urban traffic congestion in the United States is a significant drain on productivity and the environment. One study estimates that urban drivers spend about 36 hours annually stuck in congestion and that results in waste of about 24 gallons of fuel [1]. While in the past this congestion has been mitigated by expanding the roadway network, roadway infrastructure investments are incredibly expensive and have been shown to actually exacerbate congestion [2, 3] Instead, network-wide control strategies that seek to manage traffic on the network are slowly gaining traction. For example, several cities have implemented congestion pricing strategies that charger users to enter the downtown area [4, 5, 6]. Adaptive signal control strategies are also used in Zurich to restrict flow into congested portions of the network and instead direct this flow to underutilized areas [7]. However, the effects of these strategies have been difficult to predict due to the complexity of urban street networks.

For this reason, there has been a recent movement towards modeling urban traffic networks macroscopically by examining the relationships that arise between aggregate traffic relationships. It has been shown that robust and reproducible relationships exists between the average vehicle flow, average vehicle density and average rate at which trips are completed and leave the network [8, 9, 10]. The relationship between flow and density is known as the Macroscopic Fundamental Diagram (MFD) and the relationship between trip completion rate and density (or network accumulation) is known as the Network Exit Function (NEF). These macroscopic relationships can then be used to create various network-wide traffic control strategies and analyze their effects in simple ways. For example, vehicle entry into the network can be restricted by timing traffic signals on the periphery of the network to restrict flow into the network, or through pricing [8, 11, 12, 13].

In order for these strategies to be implemented efficiently, accurate estimation of the traffic state is required. The use of mobile probe vehicles to estimate the traffic states has been suggested in the literature [14]. Vehicle probes are preferred to fixed detectors (e.g., loop detectors) in urban networks because queues forming at signalized intersections can spill over onto the detector and lead to erroneous estimates. But while the use of mobile vehicle probes to estimate localized traffic conditions has been studied, very little research has been done to determine how effective this type of data can be to inform network-wide control strategies. In light of this, the purpose of this study was to develop methods to use mobile probe data to describe network-wide traffic conditions and examine ways these probe-based estimates can be integrated into network-wide traffic control strategies.

This report presents five research efforts that have either been included in a set of conference proceedings, published in an academic journal or is in consideration for publication. The remainder of this report provides a short description of each of these

studies. The latest draft of each of these studies are then included in the Appendix of this report.

Paper 1: Using mobile probe data and the Macroscopic Fundamental Diagram to estimate network densities: Tests using microsimulation (published in *Transportation Research Record*)

This paper presents a method of indirectly estimating average vehicle densities across a network in real time by combining travel speed information from few circulating probe vehicles with the Macroscopic Fundamental Diagram of urban traffic. The proposed method is advantageous because it requires relatively little data and involves few calculations. Tests of this methodology on a simulated network show that while the results are not accurate when the network is uncongested, reliable density estimates can be obtained when the network is congested or approaching congestion, even if only a small fraction of vehicles serve as probes. This is promising since congested states are the most critical. Therefore, this methodology seems useful as a traffic monitoring scheme to complement network-wide control strategies, provided that the network exhibits a well-defined and reproducible Macroscopic Fundamental Diagram.

This paper is presented in Appendix A.

Paper 2: The accuracy of network-wide traffic state estimations using mobile probe data (published in *Transportation Research Record*)

This work proposes how data from mobile probe vehicles can be used to estimate relevant network-wide traffic metrics, including average vehicle flow, density, speed, vehicle accumulation and exit flow. The method requires very little data—just the distances traveled by probes at various times and both probe and non-probe vehicle counts at fixed locations. The former piece of information is becoming increasingly available through advances in Intelligent Transportation Systems, GPS and mobile computing. The latter can be estimated by combining probe data with fixed detector sources. In addition, the uncertainty of these measurements can be estimated using data from the probe vehicles themselves. This information can be used to directly estimate the MFD and other network-wide relationships or monitor traffic in real-time. This methodology is tested on a micro-simulated network and has been shown to be very accurate when mobile probe penetration rates reach about 20%.

This paper is presented in Appendix B.

Paper 3: Deriving Macroscopic Fundamental Diagrams from probe data: Issues and proposed solutions (under consideration for publication)

We propose here a method to estimate a network's MFD using mobile probe data when the market penetration level of such technologies is not the same across an entire network. This method relies on the determination of an appropriate average probe penetration rate that can be used to represent conditions across the network. These average probe penetration levels are weighted harmonic means using the individual probe vehicle travel times and distances as weights. The results of this method are then tested in the INTEGRATION micro-simulation environment. The estimated MFDs are compared to the ground truth MFD obtained using a 100% market penetration of probe vehicles to evaluate the proposed methods. The results show that in general, the weighted harmonic mean probe penetration levels outperform a simple (arithmetic) average probe penetration level. This conclusion especially holds true when the imbalance of demand and penetration level increases. In addition, an algorithm to estimate the average probe penetration level is proposed, as this data might not generally be known. This algorithm links count data from sporadic fixed detectors in the network to probe vehicle information that pass the detectors. The simulation results indicate that this algorithm is very effective. Since the data needed for this algorithm are readily available and easy to collect (specifically the detected probe penetration level at randomly selected links and probe vehicle travel information), the proposed algorithm is practically feasible and offers a better approach for the estimation of the MFD using mobile probe data that is becoming increasingly available in urban environments.

This paper is presented in Appendix C.

Paper 4: Comparing the use of link and probe data to inform perimeter metering control (in the proceedings of the 94th Annual Meeting of the Transportation Research Board)

This work compares the use of traffic state and MFD estimations from two methods to inform a simple perimeter boundary-flow control scheme using micro-simulation. The first uses point estimates on links traditionally available from fixed detectors. The second uses trajectory information from GPS-enabled mobile probe vehicles that can be more spatially distributed. We find here that both methods can be adequately used to inform the perimeter flow control schemes. Furthermore, the performance of the network is remarkably consistent when reduced information—from a subset of detectors or probe vehicles—is used to inform the control scheme. Additionally, our results suggest that accounting for uncertainty in the state estimates can improve network performance when very few probe vehicles or detectors are used to inform the control scheme. These results are very promising for the implementation of MFD-based network-wide control in practice.

This paper is presented in Appendix D.

Paper 5: Design and evaluation of network control strategies using the Macroscopic Fundamental Diagram (in the proceedings of the 2015 IEEE Intelligent Transportation Systems Conference)

This paper discusses how to estimate the MFD using probe data with varied penetration rates across a network. Subsequently, congestion control strategies are applied to the network and the MFDs are plotted for each control strategy. The results demonstrate that it is feasible to use the MFD estimated using limited probe data as an effective tool

to monitor and control a network. The most effective strategy is a network-wide adaptive traffic signal control system, which decreases delays by up to 40%. Average fuel consumption levels decrease by up to 10%. Furthermore, combining control strategies without fully integrating them produces system-wide dis-benefits relative to running each system independently.

This paper is presented in Appendix E.

REFERENCES

- 1. Schrank, D. and Lomax, T. (2009) *2009 Urban Mobility Report*, Techinical report for the Texas Transportation Institute.
- 2. Lee, D., Klein, L., and Camus, G. (1999) Induced traffic and induced demand. *Transportation Research Record*, No. 1659, pp. 68-75.
- 3. Cervero, R. (2002) Induced travel demand: Research design, empirical evidence, and normative policies. *Journal of Planning Literature*, Vol. 17, No. 1, pp. 3-20.
- 4. Phang, S.Y. and Toh, R. (2004) Road congestion pricing in Singapore: 1975 to 2003. *Transportation Journal*, Vol. 43, No. 2, pp. 16-25.
- 5. Eliasson, J. and Mattsson, L.-G. (2006) Equity effects of congestion pricing and a case study for Stockholm. *Transportation Research Part A*, Vol. 40, No. 7, pp. 602-620.
- 6. Litman, T. (2011) *London congestion pricing: Implications for other cities.* Technical report for the Victoria Transport Policy Institute.
- 7. Nash, A.B. and Silvia, R. (2001) *Implementation of Zurich's transit priority program*. Technical report for the Mineta Transprotation Institute.
- 8. Daganzo, C.F. (2007) Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B*, Vol. 41, No. 1, pp. 49-62.
- 9. Daganzo, C.F. and Geroliminis, N. (2008) An analytical approximation for the macroscopic fundmental diagram of urban traffic. *Transportation Research Part B*, Vol. 42, No. 9, pp. 771-781.
- 10. Geroliminis, N. and Daganzo, C.F. (2008) Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B*, Vol. 42, No. 9, pp. 759-770.
- 11. Geroliminis, N. and Daganzo, C.F. (2007) Macroscopic modeling of traffic in cities. In *Proceedings of the 86th Annual Meeting of the Transportation Research Board*.
- 12. Geroliminis, N. and Levinson, D. (2009) Cordon pricing consistent with physics of overcrowding. In *Proceedings of the 18th International Symposium of Transportation and Traffic Theory*.
- 13. Daganzo, C.F., Gonzales, E.J., and Gayah, V.V. (2011) Traffic congestion in networks, and alleviating it with public transportation and pricing. *UC Berkeley Volvo Working Paper UCB-ITS-VWP-2011-7*.
- 14. Daganzo, C.F., Gayah, V.V., and Gonzales, E.J. (2012) The potential of parisomious models for understanding large scale transprotation systems and answering big picture questions. *EURO Journal of Transportation and Logistics*, doi 10.1007/s13676-012-0003-z.

APPENDIX A

Using mobile probe data and the Macroscopic Fundamental Diagram to estimate network densities: Tests using microsimulation

This article may be cited as: Gayah, V.V. and Dixit, V.V. (2013) Using mobile probe data and the macroscopic fundamental diagram to estimate network densities: Tests using microsimulation. *Transportation Research Record*, 2390:76--86.

INTRODUCTION

Macroscopic models of urban traffic have existed for nearly five decades. However, many earlier models were at best not comprehensive and at worst physically inaccurate since they were unable to describe congested network conditions (1-5). Those models that were able to accurately describe both uncongested and congested conditions were unable to describe network dynamics (6-11). Moreover, the lack of available and reliable traffic data at the time meant that these early models (and their underlying assumptions) could not be verified across multiple datasets.

Recently, a macroscopic model of urban traffic was proposed that was both physically realistic and was able to describe network dynamics. This model conjectured that a unique, reproducible relationship exists between the average flow of vehicles across a network, q, and the average density of vehicles within the network, k (12-13). This relationship has come to be known as the Macroscopic Fundamental Diagram (or MFD). Furthermore, if the average trip length within the network is constant over time, this model suggests that a similar relationship should exist between the rate at which trips are completed in the network and the number of vehicles currently traveling within it (network accumulation). This relationship has come to be known as the Network Exit Function (or NEF). Both of these relationships have since been shown to exist theoretically (13), with simulation (14) and, due to recent advances in ITS technologies, empirically (15-16) for networks in which drivers distribute themselves relatively evenly across the links in the network. Such even distributions should be especially likely on networks of well-connected streets if drivers change routes to avoid locally congested areas (17-19).

The use of these macroscopic models has shown significant potential in the study and control of urban traffic networks. For example, Daganzo (12) showed: that networks have an innate tendency towards gridlock when they become congested; and, how they can be optimally managed during a rush hour period to avoid gridlock states and minimize the total amount of vehicular delay by metering vehicle entry into the network. The optimal strategy is quite simple: 1) allow vehicles to enter the network until the density of vehicles within the network reaches a critical value associated with the maximum trip completion rate; then, 2) limit the rate at which vehicles are allowed to enter the network so that the density never exceeds this critical value. If at any point the density somehow exceeds this critical value (e.g., due to stochastic fluctuations in internal trip demand), vehicle entry into the network can be further reduced to avoid gridlock. The potential of this strategy has also been confirmed using micro-simulation (14).

Other control strategies have also been introduced that make use of the MFD or NEF. These include: optimal pricing strategies for cars entering an urban network (20); optimal operation and pricing of cars and transit in urban networks (21); and, optimal allocation of space to transit in congested networks (22). While these strategies tend to be elegant and simple to apply in theory, they all require knowledge of the current traffic state (e.g., current average network density) in real time. Even if this information is not available, at a minimum these strategies require knowledge of whether or not the network is currently congested.

However, direct estimation of average network densities in real time is problematic given current detection technologies. Fixed loop detectors, the most common method used to collect traffic data, tend to be inaccurate in urban areas. The reason is that these detectors tend to be placed near intersections and the presence of queues at signals can spill back onto the detector causing incorrect density estimates. Because of this, loop detector location has even been found to significantly affect macroscopic relationships (*16, 23, 24*). Other fixed detectors, such as cameras, can mitigate this, but they are unable to cover the entire network without significant expense. In addition, the time required to process this type of data makes real-time estimation nearly impossible.

Probe vehicles traveling within the traffic stream can also be a useful way to collect traffic information. As early as twenty years ago, several demonstration projects were implemented to determine travel time information from probe vehicle data (25-28). However, heuristic algorithms (29) and simulation studies (30) suggest that a large number of probe vehicles are needed to accurately determine network-wide traffic conditions. For example, Srinivasan and Jovanis (29) found that 3,500 probe vehicles would be needed to estimate travel times on the major arterials and freeways in Sacramento. Such a large number might be infeasible if vehicles have to be specially outfitted to provide real-time probe data.

The recent proliferation of GPS enabled handheld devices has made it possible for vehicles already traveling in the network to serve as mobile probes. Since these GPS devices are placed within vehicles driven by regular drivers, equipped vehicles travel through the network with the same driver behavior and origin-destination patterns as the general traffic stream. These devices are able to simultaneously collect and communicate a wealth of information about the trip being made. This includes current location, speed, acceleration as well as the history of the current trip. This data can then be aggregated and analyzed in real time to estimate traffic conditions with little additional infrastructure. Herrera et al *(31)* provided an example of such an approach for a single freeway. This method can be extended to provide traffic information across the individual links in a network. However, applying this method to all links in a network would require extensive computational effort. Instead, it might be possible to make use of macroscopic network properties to indirectly estimate aggregate network conditions.

In light of this, this paper proposes a methodology that combines data obtained from mobile probes with the macroscopic fundamental diagram of urban traffic to indirectly estimate network densities. The accuracy of this methodology is then tested using microsimulation. In particular, this study aims to determine what types of network states can be accurately estimated, what type of data need to be collected from the probe vehicles, how many probe vehicles are required and how often does this data need to be collected. These results can be used to inform the necessary traffic monitoring required to implement networkwide control policies in the field.

The rest of this paper is organized as follows. We first describe the simulated traffic network and mobile probe data that is used as a part of this study. We then describe the

methodology that is used to estimate network densities using the mobile probe data and macroscopic fundamental diagram. This is followed by a discussion of the results, and, finally, some concluding remarks.

STUDY DATA

This section describes the micro-simulation network used as a part of this study and simulated probe vehicle data that is made available to estimate network densities.

Orlando micro-simulation network

A simulated traffic network was used to examine the combination of probe vehicles and the MFD to estimate network densities. The network was created in the VISSIM microsimulation software for use as a part of a study for the Florida Department of Transportation (*32*). The simulated network contains a portion of the downtown Orlando street network, shown in Figure 1; it covers a roughly 1.7 mile x 1.7 mile area and contains about 120 signalized intersections. Note that only the surface streets (and not the two freeways shown in Figure 1) were included in this simulation, and that speed limits are generally low (30-35 mi/hr) on these urban streets. The simulation represents the AM peak period and covers a 3-hour rush period.

The network was created using traffic data obtained from the City of Orlando. In addition, the network was calibrated to match empirical data obtained by mobile "chase cars". These chase cars followed individual vehicles randomly as they traveled within the network boundaries and recorded data on the time spent moving, time spent stopped and total travel distance. Parameters of the macroscopic two-fluid model (7-8) were calculated for the chase cars and then these parameters were replicated within the simulation. In this way, the macroscopic traffic properties of the downtown Orlando network were reflected within the simulation. Full details on the calibration and validation procedure can be found in Dixit et al (32-33).



FIGURE 1 Orlando downtown network

Mobile probe data

Vehicles in the simulated network were randomly assigned one or two roles as they were created: unequipped (non-probe) or equipped (probe). Unequipped vehicles did not report any information. Probe vehicles were assumed to be able to report back their exact location, instantaneous speed and acceleration, average speed through its trip, and the total time and total distance traveled within the network at short, discrete intervals. Ideally, these discrete intervals should be on the order of a fraction of a second; however, extremely short intervals produced unmanageable file sizes and decreased computational speed of the micro-simulation significantly. Several tests were performed to determine a reporting interval that provided accurate results within the data constraints of the software. A 3-second interval was selected as it was able to produce data at high enough frequency to be accurate, while still remaining manageable.

The fraction of vehicles equipped as probes is an important determinant of the accuracy of traffic state estimations. Several different proportions (i.e., penetration rates) were tested; see Table 1 for a list of mobile probe penetration rates considered in this study.





5
7.5
10
15
20
25
30
40
50

METHODOLOGY

The proposed estimation methodology combines data obtained from the probe vehicles with the MFD of the urban traffic network to estimate the average network density, instead of trying to measure it directly. This procedure makes use of the fact that for well-defined MFDs, each state predicted by the curve is associated with a unique value of average vehicle

speed, \boldsymbol{v} ; see Figure 2. Therefore, we propose to estimate traffic states by first estimating average travel speeds of vehicles in the network using the probes, and then identifying the corresponding average network density on the MFD associated with that speed. Even if the exact traffic state cannot be determined with a sufficient degree of accuracy, perhaps because the mobile probes are not able to accurately estimate average travel speeds of all vehicles, it might be possible to at least determine when the network is congested or approaching congestion.



FIGURE 2 Typical MFD showing unique average vehicle speed associated with a given density

The remainder of this section describes how average vehicle speeds can be calculated using very little mobile probe data, the macroscopic fundamental diagram of the Orlando Network and how the probe data and Orlando MFD were combined to predict traffic states within the network.

Average speed calculation

As per Edie's generalized definitions of traffic stream measurements (34), the average speed of the vehicles in some measurement period is simply the ratio of the total distance traveled within that period and the total time spent within the network. Therefore, the average speed of the probe vehicles can be expressed as:

$$v_{p} = \frac{\sum_{i=1}^{I} d_{i}}{\sum_{i=1}^{I} t_{i}}$$
(1)

where d_i and t_i are the total distance traveled and total time spent, respectively, in the

network by probe vehicle *i* during the measurement period, and *I* is the total number of

vehicles that were present in the network during this period. Equation (1) shows that only two pieces of information are required from each probe vehicle to measure the average speed. Note that this can be easily determined if vehicles report their odometer readings at the beginning and end of any measurement period, and odometer readings and time when the vehicle enters or exits the network. The average speed of probe vehicles, $w_{\mu\nu}$ can then be used as an estimate for the average speed of all vehicles

traveling within the network, \$\$, during the measurement period.

We define each of the periods over which average vehicle speeds are calculated as a sampling interval. The length of a sampling interval can affect the accuracy of average speed and network state estimations. Table 2 presents the different sampling interval lengths considered in this study. Note that each of the many 3-hour simulation periods was broken up into multiple unique sampling intervals.

Sampling Interval length (seconds)
15
30
45
60
90
120
150
300

TABLE 2 Sampling interval lengths considered

Macroscopic flow-density relationship

Edie's generalized definitions (34) can also be used to determine the relationship between average network flow and average network density. Figure 3a shows this relationship calculated at discrete 5-minute intervals for the Orlando network. Note that this figure was created using data from many unique simulations of the 3-hour rush period.

As shown in Figure 3a, the data are highly chaotic and a unique, reproducible curve that describes average flow as a function of average density does not appear to exist. Recent work has shown that network behavior can be significantly different during the beginning of a rush than at the end of a rush (19). During the beginning of the rush, traffic tends to be more uniformly distributed across the network and it should be more

likely for a reproducible MFD curve to arise. However, during the end of the rush vehicles tend to be less evenly distributed across the network and these uneven distributions cause inefficient and chaotic network states. For this reason, in this study we focused only on traffic states during the beginning of the rush (defined as the first two hours of the rush period before average densities start to decrease). The flowdensity relationship for the beginning of the rush is shown in Figure 3b. Notice that this relationship is much less chaotic and a unique and reproducible relationship exists between average vehicle flow and average vehicle density. The solid line in Figure 3b denotes this reproducible curve.





FIGURE 3 Average flow-density relationships for the: (a) entire 3-hour rush period; and, (b) first two hours of the rush period. Note that (b) represents the MFD.

Focusing only on traffic states at the beginning of the rush hinders the applicability of this work somewhat, since state estimations cannot be made at the end of the rush. However, as shown in Daganzo (12), it is critical to control networks as they transition from uncongested to congested states in order to avoid the tendency towards gridlock and to minimize vehicular delay. Since these transitions primarily occur at the beginning of the rush, our study should shed light on how accurately this critical transition can be identified.

Estimation of network state

Once the average vehicle speed, \hat{v} , is estimated, the average network density can be

estimated using information from the MFD. The MFD shown in Figure 3b can be manipulated to yield a relationship between average network density and average vehicle speed. This relationship is shown in Figure 4. Over the range of observed traffic states, the data in Figure 4 can be approximated by the following curve:

$$k = -0.4549 v^2 + 7.8753 v + 3.1525 v \in [10.1, 17.8]$$
 (2)

Note that this curve fits the data very well with an \mathbb{R}^2 value of 0.992.



FIGURE 4 Average speed-density relationship for the beginning of the rush

Values of **v** inside this observed range are inserted for **v** in Equation (2) to estimate the

average vehicle density within the network, \hat{k} . Values of \hat{v} outside these bounds cannot

be used and we keep track of the number of times this occurs for a particular combination of sampling interval length and mobile vehicle penetration rate. We will later use the proportion of out-of-bounds estimates as an indicator of the accuracy of the methodology.

RESULTS

This section examines how accurately network densities can be predicted using this methodology. The first subsection discusses the ability of the methodology to accurately predict all states during the beginning of the rush period. The second subsection discusses the ability of the methodology to identify only congested or near-congested states.

Overall state estimation

The proposed methodology was applied to determine the estimated network density (\hat{k}) for all sampling intervals over many unique simulated rush periods. To determine the accuracy of this estimation, the true network density (k) was also concurrently calculated for each sampling using Edie's generalized definitions (34). The ratio \hat{k}/k was then calculated as a measure of the estimation accuracy. Table 3 presents the mean and variance of all \hat{k}/k values for a given combination of mobile probe penetration rate and sampling interval length, as well as the percentage of speed estimates that were out of the range predicted by the MFD (out-of-bounds or OOB). Accurate and reliable vehicle density estimations should have a mean value of \hat{k}/k near 1, variance of \hat{k}/k near 0 and percentage of out-of-bounds estimates near 0.

		Mobile Probe Penetation Rate (%)										
			2.5	5	7.5	10	15	20	25	30	40	50
		mean(kî/k)	1.029	1.001	1.000	1.004	1.012	1.004	1.005	1.011	1.013	1.009
	300	variance(kî/k)	0.1798	0.0232	0.0172	0.0136	0.0096	0.0068	0.0054	0.0044	0.0039	0.0030
		frac OOB	15.79%	0.00%	0.00%	0.13%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
		mean(kî/k)	1.079	0.997	0.996	0.998	1.008	0.999	1.001	1.007	1.008	1.004
	150	variance(kî/k)	0.2721	0.0396	0.0307	0.0252	0.0180	0.0126	0.0115	0.0095	0.0078	0.0067
		frac OOB	23.04%	0.54%	0.20%	0.20%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
		mean(kî/k)	1.091	0.996	0.996	0.997	1.007	0.998	1.000	1.006	1.007	1.003
ec)	120	variance(kî/k)	0.2887	0.0444	0.0345	0.0283	0.0207	0.0148	0.0125	0.0113	0.0088	0.0072
s) (frac OOB	25.05%	0.87%	0.43%	0.11%	0.00%	0.05%	0.00%	0.00%	0.00%	0.00%
둾		mean(kî/k)	1.097	1.000	0.996	0.995	1.004	0.995	0.997	1.003	1.004	1.000
Ler	90	variance(kî/k)	0.3261	0.0590	0.0457	0.0388	0.0298	0.0242	0.0203	0.0171	0.0153	0.0130
val		frac OOB	28.11%	1.48%	0.94%	0.45%	0.04%	0.04%	0.00%	0.00%	0.00%	0.00%
ter		mean(kî/k)	1.123	1.009	1.002	1.000	1.005	0.995	0.998	1.004	1.005	1.001
Ē	60	variance(kî/k)	0.3613	0.0759	0.0599	0.0514	0.0408	0.0327	0.0269	0.0235	0.0200	0.0171
ling		frac OOB	35.44%	2.94%	1.76%	1.07%	0.38%	0.16%	0.19%	0.08%	0.03%	0.05%
du		mean(kî/k)	1.102	1.021	1.010	1.006	1.006	0.998	0.998	1.003	1.003	0.999
Sa	45	variance(kî/k)	0.3705	0.0899	0.0735	0.0623	0.0558	0.0457	0.0408	0.0375	0.0338	0.0311
		frac OOB	39.73%	4.61%	3.10%	2.15%	0.85%	0.87%	0.43%	0.41%	0.19%	0.14%
		mean(kî/k)	1.097	1.033	1.019	1.015	1.010	1.001	0.999	1.004	1.003	0.999
	30	variance(kî/k)	0.3967	0.1101	0.0910	0.0782	0.0717	0.0606	0.0561	0.0516	0.0478	0.0451
		frac OOB	44.52%	6.67%	4.53%	3.59%	1.86%	1.55%	1.15%	0.86%	0.68%	0.58%
		mean(k ^ˆ /k)	1.089	1.049	1.038	1.027	1.023	1.014	1.010	1.012	1.011	1.006
	15	variance(k [^] /k)	0.4073	0.1352	0.1107	0.0985	0.0889	0.0750	0.0707	0.0676	0.0627	0.0598
		frac OOB	50.82%	8.89%	6.88%	5.49%	3.75%	3.30%	2.91%	2.31%	2.08%	2.02%

TABLE 3 Summary statistics for estimation of all average vehicle densities

Higher penetration rates mean that more vehicles supply information to estimate average travel speed, and this would help reduce the impact of a few slower or faster vehicles biasing the estimate. Similarly, longer sampling intervals mean that

fluctuations in the speed of individual vehicles over time are averaged out. This is particularly important in urban networks because vehicles spend significant time stopped at traffic signals, and the sampling interval needs to be long enough to incorporate time spent both moving and stopped. Thus, we would expect state estimations to become increasingly more accurate as the mobile probe penetration rate and sampling interval length increases.

The results presented in Table 3 verify this expectation. In general, the accuracy of the estimation increases with the mobile probe penetration rate and length of the sampling

interval. Also note that the means of \hat{k}/k are very near 1 for most combinations of

mobile probe penetration rates greater than 2.5% and sampling intervals greater than

15 seconds. However, the values of the variance of \hat{k}/k show that the individual \hat{k}/k

values are often very different from 1. This implies that individual estimates are not very accurate.

In order to examine this more closely, box plots of the distributions of \hat{k}/k are

presented in Figure 5. Figure 5a shows plots for various sampling interval lengths for a penetration rate of 50% while Figure 5b shows plots for various penetration rates for a sampling interval of 300 seconds. The plots are not very promising. Even for the highest

penetration rates and longest sampling intervals, the scatter in individual values of \hat{k}/k

is incredibly high. For example, in the case expected to be the most accurate, individual estimates of density can be as little as 80% or as much as 130% of the actual value. This level of accuracy might not be enough if the control mechanism being implemented is sensitive like a pricing scheme. In the more likely case that penetration rates are low (say 10-20%), the plots show that individual estimates are highly inaccurate even at the largest sampling interval.



a)



FIGURE 5 Box-plots showing accuracy of estimations of all average vehicle densities for: a) mobile vehicle probe penetration rate of 50%; and, b) sampling interval length of 300 seconds.

Identification of congested states

The previous section shows that individual density estimates during the first two hours of the rush period can be very inaccurate. One of the reasons for this is revealed by considering the relationship between density and average speed in Figure 4. Note that the slope of the fitted curve becomes steeper as the average speed increases. This means that network density is more sensitive to speed estimates for higher average travel speeds than for lower average travel speeds. Thus, differences between the actual and estimated average vehicle speeds will lead to much larger differences between actual and estimated network densities when speeds are high than when speeds are low. For this reason, the methodology will be less accurate when the network is operating in free flow conditions (i.e., higher average speeds) than when the network is operating at or near congestion (i.e., lower average speeds). Note that this is not unique to this network; this property should be expected for any network with a typical unimodal, concave shaped MFD.

Fortunately, most control strategies that use the MFD or NEF do not require exact density estimates when the network is operating in free flow. In fact, state estimations are typically only needed when the network gets congested (12, 14, 20-22). Since average travel speeds at (or near) congested states are low, we would expect the density estimates to be more accurate during these more critical times. Moreover, for most control strategies, such as the optimal metering strategy proposed by Daganzo (12), we do not need to know the exact density with certainty, but instead we only need to identify when the network enters the congested state.

We now examine the accuracy of this methodology at predicting congested states. To do this, we first select a critical density that defines the boundary between free flow and congestion. For the particular MFD shown in Figure 4, this critical density is selected as 31 vehicles per mile since this represents the density for which the MFD curve starts to flatten out and scatter begins to appear in the observed flow-density data *(18)*.

We now compare the estimated and actual vehicle densities with the critical density of 31 vehicles per mile. For each combination of mobile probe vehicle penetration rate and sampling interval length, we calculate two metrics over many unique rush periods: 1) the percentage of estimated densities correctly predicted as being greater than the critical density (labeled "correct"); and, 2) the percentage of actual densities greater than the critical density that were predicted as being lower than the critical density (labeled "missed"). These two values are presented for each combination of penetration rate and sampling interval length in Table 4. The "correct" value gives an indication of the reliability of states identified as being congested, and one minus this value shows what fraction of states are incorrectly identified (false positives). The "missed" value shows the fraction of congested states that are not identified by this methodology (false negatives). Higher "correct" percentages and lower "missed" percentages are both indicative of more accurate estimations of congested network conditions.

TABLE 4 Summary statistics for identification of traffic states with densities greater than a critical

		Mobile probe penetration rate										
			2.5	5	7.5	10	15	20	25	30	40	50
	300	correct	64.1%	93.8%	96.2%	96.0%	96.5%	98.0%	97.3%	98.0%	98.0%	98.0%
		missed	20.7%	3.5%	4.8%	4.2%	3.0%	3.2%	3.3%	3.2%	3.3%	3.2%
	150	correct	57.0%	93.0%	95.2%	93.7%	96.0%	96.5%	96.2%	96.6%	97.1%	96.6%
្ល	1.50	missed	36.0%	5.7%	5.6%	4.8%	4.0%	4.3%	4.7%	4.3%	4.3%	4.5%
ŝ	120	correct	55.5%	91.2%	94.1%	93.6%	94.8%	96.7%	96.4%	96.8%	97.5%	97.2%
τω	120	missed	39.4%	5.5%	5.8%	5.2%	3.5%	4.0%	3.8%	3.3%	3.8%	3.5%
Fe	90	correct	53.8%	91.2%	93.5%	93.2%	95.3%	96.8%	96.3%	97.2%	97.4%	97.4%
g	50	missed	46.4%	6.4%	6.6%	5.7%	4.4%	4.3%	4.1%	3.9%	4.0%	4.0%
Ę	60	correct	50.3%	88.0%	90.7%	91.5%	93.3%	94.7%	95.2%	95.6%	96.8%	96.9%
Ē	8	missed	47.8%	8.6%	8.8%	7.2%	6.7%	5.5%	5.1%	5.1%	5.2%	4.8%
l ling	45	correct	48.0%	86.0%	89.0%	89.7%	91.8%	93.1%	93.4%	94.0%	95.2%	95.2%
Ê		missed	47.1%	9.5%	9.4%	7.6%	6.6%	5.8%	5.6%	4.9%	4.8%	4.4%
Sa	30	correct	43.6%	82.2%	86.6%	86.9%	88.8%	90.3%	90.7%	91.4%	92.1%	92.4%
	5	missed	46.2%	12.8%	11.2%	9.4%	8.4%	7.8%	7.4%	6.8%	6.8%	6.4%
	15	correct	39.6%	79.3%	82.9%	84.3%	86.1%	87.9%	88.3%	88.9%	90.0%	90.3%
	CT	missed	43.3%	16.6%	13.9%	11.7%	10.3%	9.4%	9.0%	8.3%	8.1%	7.5%

density (*k* = 31 veh/mi)

As expected, the accuracy of the identification of congested states increases with both the mobile probe penetration rate and the sampling interval length. These results also show that the methodology is able to accurately predict congested network states, and does not fail to identify many congested states, for a wide range of sampling intervals and penetration rates. For example, the highlighted cells show the combination of penetration rates and sampling intervals that have a prediction accuracy greater than 95% and a false negative rate of less than about 5%. Note that very accurate predictions of congested states can be achieved with as few of 7.5% of vehicles being equipped as probe vehicles as long as the sampling interval is large (300 seconds). If state estimations are needed at a finer resolution, say 90 seconds, then only 15% of vehicles need to be equipped before accurate identification of congested states can be achieved.

To examine the accuracy of exact state estimations near congestion, Figure 6 presents box plots of the values of \hat{k}/k for the states near the critical density (states whose actual or estimated density exceeds the critical density). Note that the distributions of these \hat{k}/k values are much smaller than those shown in Figure 5, confirming that predictions

are much more accurate when the density is at or near the critical density. These results are very promising. It appears that highly accurate state estimations can be obtained using this methodology near critical states. For example, for the highest sampling interval (300 seconds) most density estimates are within 10% of the actual value even for penetration rates as low as 5%.



FIGURE 6 Box-plots showing accuracy of estimations of average vehicle densities near the critical density (k = 31 veh/mi) for: a) mobile vehicle probe penetration rate of 50%; and, b) sampling interval length of 300 seconds.

CONCLUDING REMARKS

This study describes a novel methodology that combines mobile vehicle probe data with the macroscopic fundamental diagram of urban traffic to estimate the average vehicle density within an urban network in real time. By using a macroscopic model, the average density can be estimated without determining the density on individual links of the network, requiring little information from probe vehicles and few calculations. However, it does rely on the existence and knowledge of a well-defined and reproducible MFD. In general, determination of an MFD requires an extensive data collection effort so this scheme is well suited for cases where the MFD is already known (e.g., traffic monitoring for a control scheme based on a known MFD relationship).

This methodology also demonstrates that traffic estimations based on average travel speeds are fairly inaccurate in free flow states. This reason for this is that network densities are highly sensitive to average travel speeds when the network is in free flow. This behavior should be expected on all networks with a well-defined MFD due to its unimodal, concave shape. However, once average travel speeds within the network start to decrease (i.e., as the network become congested), densities become much more predictable. This is fortunate because congested states are the most critical to determine. Therefore, the accuracy of this methodology is highest in cases where high accuracy is essential.

Based on the results of this simulation study, it appears that accurate density estimates near congestion can be obtained with a mobile vehicle probe penetration rate as low as 7.5%, as long as estimates are needed only once every 5 minutes. If this sampling resolution becomes smaller, the penetration rate required to achieve accurate estimates will increase significantly. However, since most current network-wide control strategies are designed for network conditions that are changing slowly with time, these longer sampling intervals should be enough to accurately implement these strategies.

While this methodology does appear to be very promising and yields encouraging results in the micro-simulation environment, it will become important to verify these results using empirical field data. Field conditions include some complications that may yield slightly less accurate results. For one, simulated networks are very clearly defined and vehicles disappear once they leave the street network. However, in reality, it might not be easy to discern a vehicle stopped at an intersection from another vehicle parked along a street or in an adjacent garage. Even still, field conditions may also have some benefits. For example, real drivers tend to route themselves more adaptively than simulated drivers and the resulting MFDs may be more reproducible and less chaotic during the end of the rush (18, 19). Future work should also aim to determine the minimum estimation accuracy needed to effectively implement some of these network-wide control measures, and how often the control scheme needs to be updated. Once this is determined, the appropriate probe penetration rate needed can be identified using field or simulation efforts.

REFERENCES

- 1. Smeed, R. and Wardrop, J. An exploratory comparison of the advantages of cars and buses for travel in urban areas. *Journal of the Institute of Transport*, Vol. 30, No. 9, pp. 301-315, 1964.
- 2. Smeed, R. Road capacity of city centers. *Traffic Engineering and Conrol*, Vol. 9, pp. 455-458, 1966.
- 3. Thomson, J. Speeds and flows of traffic in Central London: Part 2 Speed-flow relations. *Traffic Engineering and Control*, Vol. 8, No. 12, pp. 721-725, 1967.
- 4. Wardrop, J. Journey speed and flow in central urban areas. *Traffic Engineering and Control*, Vol. 9, pp. 528-532, 539, 1968.
- 5. Zahavi, Y. Traffic performance evaluation of road networks by he alpharelationship: Part 2. *Traffic Engineering and Control*, Vol. 14, pp. 292-293, 1972.
- 6. Godfrey, J. The mechanism of a road network. *Traffic Engineering and Control*, Vol. 11, pp. 323-327, 1969.
- 7. Herman, R. and Prigogine, I. A two-fluid approach to town traffic. *Science*, Vol. 204, No. 4389, pp. 148-151, 1979.
- 8. Ardekani, S. and Herman, R. Urban network-wide traffic variables and their relations. *Transportation Science*, Vol. 21, No. 1, pp. 1-16, 1987.
- 9. Mahmassani, H., Williams, J. and Herman, R. Performance of urban traffic networks. In: *Proceedings of the 10th International Symposium on Transportation and Traffic Theory*, 1987.
- 10. Mahmasani, H. and Peeta, S. Netwrok performance under system optimal and user equilibrium dynamic assignments: Implications for advanced traveler information systems. *Transportation Research Record*, No. 1408, pp. 83-93, 1993.
- 11. Olszewski, P., Fan, H. and Tan, Y. Area-wide traffic speed-flow model for Singapore CBD. *Transportation Research Part A*, Vol. 29, No. 4, pp. 273-281, 1995.
- 12. Daganzo, C.F. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B*, Vol. 41, No. 1, pp. 49-62, 2007.
- 13. Daganzo, C.F. and Geroliminis, N. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B*, Vol. 42, No. 9, pp. 771-781, 2008.
- 14. Geroliminis, N. and Daganzo, C.F. Macroscopic modeling of traffic in cities. In: *Proceedings of the 86th Annual Meeting of the Transportation Research Board*, 2007.
- 15. Geroliminis, N. and Daganzo, C.F. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B*, Vol. 42, No. 9, pp. 759-770, 2008.

- 16. Buisson, C. and Ladier, C. Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. *Transportation Research Record*, No. 2124, pp. 127-136, 2009.
- 17. Gayah, V.V. and Daganzo, C.F. Exploring the effect of turning maneuvers and route choice on a simple network. *Transportation Research Record*, No. 2249, pp. 15-19, 2011.
- Daganzo, C.F., Gayah, V.V. and Gonzales, E.J. Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. *Transportation Research Part B*, Vol. 45, No. 1, pp. 278-288, 2011.
- 19. Gayah, V.V. and Daganzo, C.F. Clockwise hysteresis loops in the macroscopic fundamental diagram: An effect of network instability. *Transportation Research Part B*, Vol. 45, No. 4, pp. 643-655, 2011.
- 20. Geroliminis, N. and Levinson, D. Cordon pricing consistent with the physics of overcrowding. In: *Proceedings of the 18th International Symposium on Transportation and Traffic Theory*, 2009.
- 21. Gonzales, E.J. and Daganzo, C.F. Morning commute with competing modes and distributed demand: User equilibrium, system optimum, and pricing. *UC Berkeley Volvo Working Paper UCB-ITS-VWP-2011-2*, 2011.
- 22. Daganzo, C.F., Gayah, V.V. and Gonzales, E.J. The potential of parsimonious models for understanding large scale transportation systems and answering big picture questions. *EURO Journal of Transportation and Logistics*, Vol. 1, No. 1-2, pp. 47-65, 2012.
- 23. Wu, X., Liu, H. and Geroliminis, N. An empirical analysis of the arterial fundamental diagram. *Transportation Research Part B*, Vol. 45, No. 1, pp. 255-266, 2011.
- 24. Courbon, T. and Leclercq, L. Cross-comparison of macroscopic fundamental diagram estimation methods. *Procedia Social and Behavioral Sciences*, Vo. 20, pp. 417-426, 2011.
- 25. Aplin, N. Overview of IVHS activities in Australia. *IVHS America Annual Proceedings*, pp. 286-291, 1993.
- Ligas, J., Bowcott, S.F., Boyce, D., Nelson, P., Schofer, J., Koppleman, F. and Kirson, A. An update on ADVANCE. *IVHS American Annual Proceedings*, pp. 621-626, 1993.
- 27. Clelland, A., Wendtland, M., Tedesco, S., Mammano, F. and Endo, G. Evaluations results and lessons learns from the Pathfinder operational test. *IVHS American Annual Proceedings*, pp. 448-457, 1993.
- 28. Barbaresso, J. and Grubba, J.L. Overview of the FAST-TRAC IVHS program. *IVHS American Annual Proceedings*, pp. 630-638, 1993.

- 29. Srinivasan, K.K. and Jovanis, P.P. Determination of number of probe vehicles required for reliable travel time measurement in urban network. *Transportation Research Record*, No. 1537, pp. 15-22, 1996.
- 30. Chen, M. and Chien, S.I.J. Determining the number of probe vehicles for freeway travel time estimation by microscopic simulation. *Transportation Research Record*, No. 1719, pp. 61-68, 2000.
- 31. Herrera, J.C., Work, D.B., Herring, R., Ban, X., Jacobson, Q. and Bayen, A.M. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C*, Vol. 18, No. 4, pp. 568-583, 2010.
- 32. Dixit, V.V., Crowe, J. and Radwan, E. Evalution and microscopic simulation of the Orlando downtown network. Final report submitted to the City of Orlando, Department of Transportation, 2009.
- 33. Dixit, V.V., Crowe, J. and Radwan, E. The calibration and validation of microsimulation using the two-fluid model. In: *Proceedings of the 90th Annual Meeting of the Transportation Research Board*, 2011.
- 34. Edie, L.C. Discussion of traffic stream measurements and definitions. In: Proceedings of the 2nd International Symposium on the Theory of Traffic Flow, 1965.

APPENDIX B

The accuracy of network-wide traffic state estimations using mobile probe data

This article may be cited as: Nagle, A.S. and Gayah, V.V. (2014) Accuracy of networkwide traffic states estimated from mobile probe data. *Transportation Research Record*, 2421:1--11.

INTRODUCTION

Aggregate models of urban traffic have long existed (1-4), although early attempts were limited by lack of theoretical justification, physical realism, verification of underlying assumptions and reliable network data. Recently, Daganzo (5) proposed conditions under which well-defined relationships between urban traffic variables should arise: uniform congestion distributions and average trip lengths that were invariant with time. Under these conditions, average flow and density measured across a network should follow a reproducible curve, known commonly as the Network or Macroscopic Fundamental Diagram (MFD). This relationship can also be scaled to describe a relationship between the rate trips at which are completed and the number of vehicles on a network, known commonly as the Network Exit Function (NEF). Evidence suggesting the existence of these relationships was first presented using traffic data from the city of Yokohama, Japan in Geroliminis and Daganzo (6). The MFD and NEF models are very useful tools in the management, design and control of urban traffic networks. For example, Daganzo (5) shows how carefully metering the vehicle entry rate into a network can avoid gridlock and minimize the total vehicle travel time. Other strategies use these models to determine more realistic metering schemes (7-9), pricing strategies (10-11), space allocation (12), street network design (13) and vehicle routing (14). However, many of these strategies require that the functional form of the MFD is known a priori and that network-wide traffic conditions can be estimated in real-time. Fortunately, the entire MFD or NEF might not be needed; instead, a reduced version using a subset of data might be sufficient (15).

Unfortunately, the MFD and NEF are difficult to estimate in practice. Very few studies have been able to obtain the requisite data to estimate MFDs of urban networks (16-17). Estimates from fixed measurement sources, like inductive loop detectors and cameras, require expensive infrastructure. Furthermore, estimates from these fixed sources are also impacted by their placement due to queues at intersections (17-19). While studies (15) have shown that limited detector information is sufficient to implement gating control in a specific network, the impact of detector location might result in MFDs that are inaccurate or not comparable across networks. If the control implemented is very sensitive, like pricing, more accurate MFDs may be required. Daganzo and Geroliminis (20) provide a methodology to estimate the functional form of the MFD for networks consisting of a single route and proposed that these relationships would hold on more complex and realistic multi-route networks. However, simulations suggest that these formulae overestimate the MFD (21-22). Furthermore, theory (22-23), simulations (24-25) and empirical data (17) suggest that MFDs might be more complex, containing phenomena like hysteresis. Even if the MFD is known, monitoring network-wide traffic conditions is still an issue: real-time network data is rarely available, and using fixed sources to cover a large spatial region can be difficult or expensive to achieve.

Recent advances in GPS-enabled devices and mobile computing may help overcome some of these obstacles. These "probe" vehicles can provide information on their location, speed, and distance traveled at regular intervals, which serves as a rich source of traffic data in real-time. Probe vehicles are not location constrained, providing much better spatial coverage than fixed sensors. They also require little additional infrastructure and are becoming increasingly common in fleet vehicles and private automobiles. In fact, the Federal Highway Administration (FHWA) Connected Vehicle (CV) initiative is providing a usable framework that facilitates collecting and processing probe vehicle data (*26*). Previous studies have used probe data to estimate traffic conditions, but these focus primarily on travel times on individual arterial roadways (*27*) or speeds on freeways (*28*) and do not consider network-wide conditions. A recent study proposed combining probe data and a known MFD to monitor network densities in real-time (*29*), but this methodology is not accurate in free-flow travel conditions and cannot be used if the MFD exhibits hysteresis patterns.

This paper proposes a way to combine data from probe vehicles with limited data from fixed detectors to estimate network-wide traffic characteristics like average flow, density, speed and exit rate in real-time. The methodology does not rely on the existence of a well-defined MFD and provides accurate estimates even if hysteresis phenomena are present. This method can also estimate a network's MFD in a way that is directly comparable across different networks. In addition, analytical formulae can be developed to quantify the uncertainty of the measurements. These rely only on measurable pieces of probe data and can yield insights as to the penetration rate needed for a given accuracy. Simulations are used to verify all formulae, and as expected, estimates become more accurate as the fraction of vehicles that serve as mobile probes increases.

The remainder of this paper is organized as follows. First, we describe the variables of interest and the estimation procedure. Then, we develop formulae to determine the uncertainty of the estimates. Next, we examine the accuracy of these estimations using data generated from a micro-simulation of an idealized network. Then, we discuss how this methodology can be used to estimate the MFD and quantify its accuracy. And finally, we provide some concluding remarks.

estimation of network-wide traffic variables

This section discusses the network-wide traffic variables of interest, and the estimation procedure for these variables. We first define and discuss how to measure each variable. Then, we explore how to estimate each using data from probe vehicles, assuming the probe penetration rate is known a priori. Next, we discuss how to estimate this penetration rate in real-time by combining data from fixed and mobile sources.

Variables of interest

To estimate a network's MFD, the average flow, q [veh/hr], and the average density, k [veh/mi], of vehicles in the network must be determined. For a given analysis period, these properties are defined using the generalized definitions of Edie (*30*):

$$q = d_T / LT = N d / LT, \text{ and}$$
(1)

$$k = t_T / LT = N\bar{t} / LT, \tag{2}$$

where d_T [veh-mi] is the total distance traveled by all vehicles on the network during the analysis period, \overline{d} [mi] is the average distance traveled, t_T [veh-hr] is the total time vehicles spend in the network, \overline{t} [hr] is the average time spent, N [veh] is the number of vehicles to use the network during this period, L [mi] is the total length of streets and T[hr] is the length of the analysis period.

Another metric to describe network-wide traffic conditions is average vehicle speed, **v** [mi/hr]. For networks with well-defined MFDs, the average speed can serve as a proxy for the level of congestion or density (*12*), since each state on the MFD is associated with a unique value of average vehicle speed. As per Edie:

$$v = d_T / t_T = \bar{d} / \bar{t}. \tag{3}$$

For modeling purposes, it is often more convenient to use the NEF, which relates the average rate that vehicles are able to exit the network, f [veh/hr], with the average number of vehicles traveling inside the network, n [veh]. These two metrics are defined as follows:

$$f = qL/l = N\bar{d}/Tl = N^{e}/T \text{ and}$$
(4)

$$n = kL = N\bar{t}/T, \tag{5}$$

where *N*^{*e*} [veh] is the total number of vehicles that exit the network during the analysis period and *l* [mi] is the average length of a trip. The variables *q*, *k*, *v*, *f* and *n* can fully describe average traffic conditions within a network.

Estimating variables using probe data

If the trajectories of all vehicles traveling within the network are available, the values of

 d_T , t_T , N^e and N can be calculated for any analysis period to directly determine the

metrics of interest. Unfortunately, if only a subset of trajectory information is available (which occurs if only a fraction of the vehicles serve as mobile probes), then these values cannot be calculated. However, it may be possible to estimate these quantities

using only probe data if the fraction of vehicles that serve as probes, ρ , is known a

priori. In this case, the number of probe vehicles within the network during any analysis period should be equal to a fixed proportion of the total number of vehicles during that time: $N_p = N\rho$. Furthermore, if the fraction of probe vehicles is the same for all O-D pairs (i.e., if probes are uniformly distributed across space), the average distance and

time traveled in the network by the probe vehicles, ${ar d}_p$ and ${ar t}_p$, respectively, should

accurately represent the same quantities measured over all vehicles; i.e., $\vec{d} = \vec{d}_{p}$ and

$\bar{t} = \bar{t}_{p}$.

By substituting the estimates from probe vehicles into Equations 1-3 and 5, we obtain estimates of the average flow, density, speed and number of vehicles in the network:

$\hat{q} = N_p \bar{d}_p / \rho LT$	(6)

$\hat{k} = N_n \bar{t}_n / \rho LT$	(7)
p-p-r	(•)

$$\hat{v} = \bar{d}_p / \bar{t}_p \tag{8}$$
Similarly, the number of exiting vehicles during an analysis period should be proportional to the number of probe vehicles that exit the network, $N_p^e = N^e \rho$. Substituting this into Equation 4, we find:

$$\hat{f} = N_p^e / \rho T \tag{10}$$

The quantities \bar{d}_p , \bar{t}_p , N_p and N_p^s are directly measureable from the trajectories of

probe vehicles, which can be accurately obtained in real-time. In fact, entire trajectories are not needed: these quantities can also be obtained if probe vehicles simply report their odometer reading at discrete time periods that represent the start and end of all analysis intervals, and both the odometer reading and time when the vehicle enters and exits the network as in (29). Therefore, we can estimate the network metrics of interest, as well as the MFD and NEF, using data from probe vehicles and Equations 6-10 if the

probe penetration rate, *p*, is known.

Estimation of mobile probe penetration rate

In general, the probe penetration rate may not be known a priori and will most likely vary with time. For example, the fraction of Connected Vehicles may increase at different rates across the country, and these rates may not generally be known. If fleet vehicles like taxis are used as probes, the proportion of traffic that these vehicles represent should be expected to change significantly over time.

Fortunately, analysts can combine data from probes with traditional fixed sensors to estimate the fraction of probe vehicles during any particular analysis period, as in

Herrera et al. (31). To estimate p in this way, detectors can be placed at various

locations throughout the network to count the total number of vehicles that cross all

detectors during an analysis period, N^{d} . Simultaneously, the GPS devices can be used to

track the number of probe vehicles crossing detector locations during the same period,

 N_{p}^{d} . An estimate of the mobile probe penetration rate is simply:

This provides an easy way to calculate ρ using only the data from the probe vehicles and existing loop detector data. The vehicle counts from detectors, especially those near signals, are far more reliable than occupancy or density estimates, as long as the analysis period is larger than a cycle length.

ACCURACY OF ESTIMATIONS

We now examine the uncertainty of traffic variables estimated in this way. As will be shown, the uncertainties can often be estimated well using just probe data.

Accuracy of probe penetration rate

Let us first consider the accuracy of the estimate of p. If N^d vehicles travel over

detectors during some time period, we would expect $N_p^d = \rho N^d$ probe vehicles to do so

on average. However, N_p^d would have some variation due to randomness. Treating N_p^d as a binomial random variable, where each vehicle crossing a detector has the same probability, ρ , of being a probe, we find $var(N_p^d) = \rho(1-\rho)N^d$. Thus, the variance of

the estimator $\hat{\rho}$ is:

$$v\hat{a}r(\hat{\rho}) = \rho(1-\rho)/N^d \tag{12}$$

Notice that Equation 12 tends toward 0 as N^d increases. Therefore, the accuracy of the estimation should increase as more vehicles are counted by detectors during an analysis period. This can be achieved by increasing either the number of detectors in the network or, more realistically, the length of the analysis interval. Equation 12 can be used to determine the number of vehicles that need to be sampled by detectors to achieve a given level of accuracy by assuming the worst case $\rho = 0.5$ that yields the largest $v \hat{\alpha} r(\hat{\rho})$. For example, to determine $\hat{\rho}$ with a maximum standard deviation of

0.005, the detectors must sample 10,000 vehicles during an analysis period. Given the number of detectors on the network, and using a minimum expected flow at each, the length of the analysis period needed for this level of accuracy can be determined.

Accuracy of traffic variables

We now consider the accuracy of the average flow, density, speed, exit rate and number of vehicles on the network. During any analysis period, the number of probe vehicles, average distance and time traveled by each and the fraction of probe vehicles are all random variables. As shown in the previous section, if we select a large enough analysis period and use enough

fixed detectors, ρ can be estimated quite accurately; therefore, we treat ρ as a constant and

express uncertainties as a function of p. The uncertainty of each estimate is found by taking the variance of Equations 6-10:

$$v\hat{a}r(\hat{q}) = q^2 var(d_i)(1-\rho)^2 / N_p^2 \bar{d}_p^2 + qvar(d_i)(1-\rho) / \bar{d}_p \rho LT + q^2(1-\rho) / N_p$$
(13)

$$v\hat{a}r(\hat{k}) = k^2 var(t_i)(1-\rho)^2 / N_p^2 \bar{t}_p^2 + kvar(t_i)(1-\rho) / \bar{t}_p \rho LT + k^2 (1-\rho) / N_p$$
(14)

$$v\hat{a}r(\hat{v}) = v^{2}var(d_{i})(1-\rho)/N_{p}\bar{d}_{p}^{2} - 2v^{2}cov(\bar{d}_{p},\bar{t}_{p})(1-\rho)/\bar{d}_{p}\bar{t}_{p}$$

$$+v^{2}(1-\rho)var(t_{i})/N_{p}\bar{t}_{p}^{2}$$
(15)

$$v\hat{a}r(\hat{n}) = n^2 var(t_i)(1-\rho)^2 / N_p^2 \bar{t}_p^2 + nvar(t_i)(1-\rho) / \bar{t}_p \rho T + n^2 (1-\rho) / N_p$$
(16)

$$v\hat{a}r(\hat{f}) = f(1-\rho)/T\rho \tag{17}$$

These equations account for the fact that the distance and time traveled by each vehicle, d_i and t_i , are not independent and identically distributed since the individual probe vehicles are selected from the set of all vehicles without replacement. Note that Equation 15 holds by applying a first-order Taylor Series approximation to Equation 8

before taking the variance. Since, in general, the quantities q, k, v and f are unknown, we can substitute the estimates provided by Equations 6-10 into Equations 13-17. Equations 13-14 and 16-17 can be calculated in real-time using measurable data from the individual probe vehicles. Unfortunately, Equation 15 contains the term $cov(\bar{d}_p, \bar{t}_p)$ that cannot be calculated with the probe data in real-time. However,

 $cov(\bar{d}_{p}, \bar{t}_{p}) > 0$ since \bar{d}_{p} and \bar{t}_{p} should have a strong positive correlation (i.e., since the

average distance traveled generally increases with the average time spent in the system). Removing this term provides an upper bound for the variance of the average speed estimator:

$$v\hat{a}r(\hat{v}) \ge v^2 var(d_i)(1-\rho)/N_p \bar{d_p}^2 + v^2(1-\rho)var(t_i)/N_p \bar{t_p}^2,$$
 (18)

which can now be calculated using probe data in real-time.

TESTS USING MICRO-SIMULATION

Data from a micro-simulation model are now used to test the accuracy of the estimation method. We first describe the simulation network that is used. Then, we compare the simulation results with the analytical equations previously developed.

Simulated network

To test these methodologies, an idealized 16x16 square grid network of alternating one-way streets was developed using AIMSUN micro-simulation software. The network consisted of 544 links each with length 400 ft. A total of 512 detectors were placed near the intersection of each non-exiting link; such setups could be expected in networks with actuated traffic signal control. Origins and destinations were assumed to exist at all entry and exit links and at all internal intersections. O-D patterns were assumed to be uniform for simplicity, but the magnitude of traffic demands were adjusted to simulate a typical morning rush with a clearly defined one-hour peak period.

Vehicles entering the network were randomly assigned as a probe vehicle with some

fixed probability, p. The value of p was held constant during each simulation run, but

due to randomness the actual probe penetration rate varied across analysis periods. Probe vehicles produced sets of vehicle trajectory data for many hours over several days. The data sets consisted of information that is easily attainable using GPS technologies, such as vehicle ID, time, position and speed at discrete 0.75-second intervals. For the tests and plots presented in this section, a single 5-minute analysis period was randomly selected to simplify the presentation of results. During this

selected period, *q* = 462.4 veh/hr, *k* = 30 veh/mi, *v* = 15.4 mi/hr, *n* = 1,235.8 veh, and *f*

= 19,920 veh/hr. However, the general trends and results presented here do not hinge on this particular period and extend to the entire range of conditions considered.

Simulation Results

Estimation of probe penetration rate

To verify the accuracy and assumptions of Equations 11 and 12, estimates of the probe penetration rate and the variance of the estimates were computed using the proposed methodology. For each run, the aggregate count of all vehicles, N^{d} , and probe vehicles, N_{p}^{d} , that crossed a detector was determined and used to calculate $\hat{\rho}$. This process was repeated several times to compute the *var(\hat{\rho}*). The actual variances follow the same trends that were expected by examining Equation 12. Namely, as the number of vehicles counted, N^{d} , increased, the variance of the estimated probe penetration rate decreased. Similarly, as the sampling interval, T, increased, the variance of the estimate also decreased.



FIGURE 1 Theoretical and actual variance of the estimated mobile probe penetration rate

The estimated values predict the actual values quite well. Furthermore, the estimates were quite consistent and contained very little variation. The actual and predicted variances are presented in Figure 1 for a given time period. Notice that Equation 12 accurately estimates the observed variance, and the variances are generally very low for

the entire range of probe penetration rates. In fact, the highest variance for $\rho = 0.5$ was

about 0.0001. These results verify the accuracy of predicting *p* in this way and justify

the treatment of *p* as a constant.

Comparing probe data with all vehicle data

average time spent were found to be statistically equivalent.

Another critical assumption in our methodology to estimate the remaining variables of interest is that the probe vehicles accurately represent the average distance and time traveled

by all vehicles; i.e., $\bar{d} = \bar{d}_p$ and $\bar{t} = \bar{t}_p$. The simulation was used to verify that these assumptions hold when all O-D pairs have the same fraction of probe vehicles. Statistical t-tests were performed for the range of $\rho = \{0.05, 0.10, 0.15, \dots, 0.75\}$ to see if the two sets of values were statistically different. For all values of ρ , the average distance traveled and

Traffic variable estimations

We now examine the accuracy of using Equations 6-10 to estimate traffic conditions and verify how well Equations 13-18 predict the uncertainty of the estimates. Many simulation instances were performed for the range of $\rho = \{0.05, 0.10, 0.15, \dots, 0.75\}$, and the estimates

of \hat{q} , \hat{k} , \hat{v} , \hat{n} , and \hat{f} were calculated for each instance. The actual values were simultaneously calculated using Equations 1-5 and the trajectory data produced by all vehicles. For each simulation instance, the ratios \hat{q}/q , \hat{k}/k , \hat{v}/v , \hat{n}/n and \hat{f}/f were calculated as a measure of estimation accuracy. Values near 1.0 indicate that the estimated value is very near the actual value. Box plots of these results are provided in Figure 2. The bottom and top of the boxes represent the lower and upper quartiles, respectively, and the band in the middle of the boxes is the median. Extreme outliers are illustrated with a plus sign. A tolerance of $\pm 10\%$ of the actual value was selected and drawn as a horizontal line on the box plot to further illustrate the accuracy (with the exception of average speed, where a tolerance of $\pm 3\%$ was used).

The estimations for average flow and density produced in the idealized grid network are within 10% of the actual flow and density values with 95% confidence for probe penetration rates of about 15%. Very similar levels of accuracy were achieved for the average exit flow estimations. The estimations for average exit flow are within 11% of the actual exit flow with 95% confidence for probe penetration rates of at least 15%. The results produced by the average number of vehicles in the system were identical to the average density (Figure 2b and 2d) because the former is proportional to the latter

by a factor of the network length, *L*.

The individual estimates of speed were shown to be much more accurate, with estimates within 3% of the actual average speed with 95% confidence for probe penetration rates of at least 10%. This increased accuracy is explained by examining Equations 6-8. The estimated speed relies only on the average distance and time spent in the network by the probe vehicles, which were shown to be statistically equivalent to the average distance and time spent in the network by all vehicles. The average flow and density estimates also rely on these data as well as the number of probes vehicles in the network. The inclusion of this last variable introduces additional sources of error to the flow and density estimates. Thus, these estimates should be less accurate than the estimates of average speed, as we have found.

While the results shown in Figure 2 are for an analysis period near capacity, this methodology was reproduced for a variety of states, including those at the beginning and end of the simulation period. Similar trends and results are obtained during the entire range of network states; thus, the method is accurate during both free flow and congested conditions, even though a wide range of conditions were considered. This is contrary to the methodology proposed by Gayah and Dixit (29), which was limited to only near-capacity states during the middle of a rush period. This illustrates that the presented methodology is promising for estimating network states that might lie on both the uncongested and congested branches of the MFD. Tests on a more realistic micro-simulation of downtown Orlando, Florida are consistent with these idealized simulations (32).

The variances of the estimates across the many simulation instances are also calculated and plotted against the estimated variances from Equations 13-18; see Figure 3. As

expected, the estimates become more accurate as *p* increases. Promisingly, the

variances are quite low: estimating all traffic variables from probe data in this way is very accurate when probe fractions are as little as 15%. For example, the period chosen for presentation here had the highest flow and density observed, and Equations 13-14 suggest that this period should have the most uncertainty in these estimates. However, flow estimates are within ± 48.5 veh/hr, and density estimates are within ± 3.2 veh/mi,

with 96% confidence when ρ = 0.15. Speed estimations are even more accurate. If ρ

increases to 0.30, the error decreases to about \pm 30.7 veh/hr and \pm 2.0 veh/mi, respectively. These results are consistent with the results shown in Figure 2.

Confidence intervals for these variances can be computed easily. If **x** represents one of the metrics considered here, the confidence interval for the true variance of **x**, **var(x)**, is:

$$(n-1)v\hat{a}r(\hat{x})/\chi^2_{1-(\alpha/2)} < var(x) < (n-1)v\hat{a}r(\hat{x})/\chi^2_{\alpha/2},$$

19)

where $\chi^2_{\alpha/2}$ is the Chi-squared value associated with n-1 degrees of freedom and a

(

significance value of *a*. The 95% confidence interval for all metrics calculated in this

way is illustrated in Figure 3 as the shaded region between the dashed lines. It is clearly shown that the estimated variances using Equations 13-17 fall within the 95% confidence interval for all probe percentages, which confirms the accuracy of the variance formulae. Thus, the theoretically derived Equations 13-17 provide a very accurate indication of the accuracy of the traffic state estimations. Unfortunately, the upper bound for the accuracy of speed estimations is not tight. Still, the magnitude of the upper bound is fairly low, so it can give a useful indication of the accuracy of speed measurements.

Applications using fractional error

Since the variance of these metrics can be accurately estimated using Equations 13-17, statistical methods can be used to determine the fraction of probes necessary to accurately estimate the average network flow, density, speed, exit flow and number of vehicles in real-time within a desired fractional error. First, we know that the confidence interval surrounding

the mean of variance x is $\mu_x \pm z\sigma_x$ and the fractional error is bounded by $\varepsilon = z\sigma_x/\mu_x$,

where \boldsymbol{z} represents the critical value of a normal distribution associated with the desired level of significance. If an agency wishes to estimate how much probe information is needed to estimate \boldsymbol{x} within some desired fractional error, \boldsymbol{z} , these expressions and Equations 13-17 can

be used to find the minimum value of p required. For example, the fraction of probes

necessary to ensure that the estimate of average exit rate is within ε of the true value with 95% confidence is:

$$\rho \ge \frac{z^2}{z^2 f T + z^2}.$$

20)

Note that Equation 20 confirms expectations: the minimum penetration rate increases with the level of significance and decreases with the fractional error allowed. Additionally, the minimum penetration rate decreases as the time period increases, reflecting that fewer probes are needed if data is collected over a longer period of time. This method may be repeated for each of the metrics considered in this study.



FIGURE 2 Boxplots showing accuracy of estimations for average: a) flow; b) density; c) speed; d) number of vehicles; and e) exit flow



FIGURE 3 Comparison of theoretical and actual variances for: a) average flow; b) average density; c) average speed; d) average number of vehicles; and e) average exit flow

ESTIMATING THE MACROSCOPIC FUNDAMENTAL DIAGRAM

The methodology presented thus far is able to estimate average network flows and densities in real-time using just limited trajectory data from probe vehicles and detectors. Since these estimations are very accurate, this methodology could be used to directly estimate a network's MFD. We now examine the potential of estimating the MFD in this way using simulated data.

The average flows and densities were calculated at discrete 5-minute intervals for the entire simulation period using all vehicle data to obtain the actual MFD; see the dark lines in Figure 4. The clockwise hysteresis shape in the MFD is similar to the shapes previously observed in many existing studies that used both empirical and simulated network data (*17-18, 29*). The primary reason for this clockwise hysteresis loop is that networks are inherently more unstable during the end of the rush hour (when densities are decreasing) than during the beginning of the rush hour (when densities are increasing), and higher levels of instability are associated with lower average flows in the network. A detailed explanation of the existence of hysteresis loops can be found in Gayah and Daganzo (*22*). In addition, the estimation methodology was performed for

many instances over the entire simulation period for $p = \{0.05, 0.10, 0.15, ..., 0.75\}$.

Samples of these MFD estimates are shown as light gray lines in Figure 4. As expected,

Figure 4 shows that the MFD estimates converge upon the actual MFD as *p* increases.

Notice that estimates are quite accurate for the range of observed states.

Three metrics are used to quantify the fit of the MFD estimates to the true MFD. The first two metrics are the root mean square error (RMSE) of the average flow and density estimates for every 5-minute interval. This provides an indication of how well the estimated MFD fits the flow and density, respectively. The RMSE of average flow and

density were simply calculated by $RMSE(q) = \sqrt{\Sigma(\hat{q} - q)^2/N}$ and

 $RMSE(k) = \sqrt{\sum (\hat{k} - k)^2 / N}$, respectively, where N is the number of time intervals

during the simulation. The third metric quantifies how well the average flow and density are estimated simultaneously. This was done by converting all values of flow and density to dimensionless values by dividing by the network capacity and jam density, respectively. These two parameters were calibrated using simulation and found

to be $q_{o} = 900$ veh/hr and $k_{j} = 320$ veh/mi. The combined error was then calculated as

the square of the distance between the actual and estimated scaled flow and density point on the MFD. The root mean square error was then calculated to determine the error associated with the entire estimate for each probe penetration rate, as such:

$$RMSE(q,k) = \sqrt{\sum \left[(\hat{q}/q_{o} - q/q_{o})^{2} + (\hat{k}/k_{j} - k/k_{j})^{2} \right] / N}.$$
(21)

The error associated with the MFD estimations are illustrated in Table 1.

TABLE 1 Root mean square error of flow, density, and dimensionless simultaneous error fo
variable probe percentages

Probe %	RMSE(q) (veh/hr)	RMSE(k) (veh/mi)	RMSE(q,k)
5	33.66	2.02	3.79E-02
10	23.09	1.39	2.60E-02
15	18.58	1.12	2.09E-02
20	15.34	0.94	1.73E-02
25	13.41	0.82	1.51E-02
30	11.74	0.71	1.32E-02
35	10.24	0.62	1.15E-02
40	9.46	0.57	1.07E-02
45	8.59	0.52	9.68E-03
50	7.58	0.46	8.55E-03
55	7.02	0.42	7.91E-03
60	6.22	0.38	7.01E-03
65	5.67	0.35	6.39E-03
70	5.04	0.30	5.68E-03
75	4.49	0.27	5.06E-03

The analytical comparison of the results shows the increase in accuracy as the probe percentage increases. The results show that the average flow and density estimate is off by 33.7 veh/hr and 2.0 veh/mi on average, respectively, with a probe penetration rate of 5%. The error is greatly reduced with just an increase to 15% probe vehicles, where the average flow and density estimates are off by 18.6 veh/hr and 1.1 veh/mi, on

average. The RMSE(q, k) value is not meaningful on its own but can be used to compare

the accuracy of MFD estimations for various values of *p*. For example, Table 1 shows

that the accuracy increases rapidly until $\rho = 0.2$, at which point the gains start to

decrease. Overall, these results show that the methodology presented in this paper is fairly accurate at estimating the MFD using limited mobile probe data.



FIGURE 4 Actual and estimated MFD for probe penetration rates of: a) 5%; b) 15%; c) 30%; d) 40%; and e) 50%

CONCLUDING REMARKS

This study provides a general methodology to estimate network-wide traffic conditions using data from mobile probe vehicles. The amount and type of data required is fairly unobtrusive and consists of each vehicle's odometer reading at discrete points in time that define periods of interest and when the vehicle enters and exits the network. From this information, realtime traffic metrics, such as average vehicle speed, flow, density and exit rate can be determined. The method requires that the fraction of circulating vehicles that serve as probes during the analysis period is known. However, we show that this fraction can be estimated quite accurately by simply combining data from probe vehicles and fixed detectors. Analytical formulae are also developed that can estimate the uncertainty of these measurements using only data provided by the probe vehicles. In this way, the accuracy of the measurements can be determined simultaneously with the measurements themselves. Simulated probe vehicle trajectories from an idealized micro-simulation network are used to examine the accuracy of this estimation methodology. The simulation tests confirm that the methodology increases in accuracy with the fraction of mobile probe vehicles on the network. In general, estimates of flow, density, exit rate and number of vehicles in the network can be obtained within 10% of the true value when as little as 15% of the circulating vehicles serve as probes. Estimates of average speed are even more accurate: when only 15% of vehicles serve as probes, average speeds are estimated within 3% of the true value. The probe penetration rates necessary for accurate speed estimates in this study are comparable to previous studies that focus on estimating travel times or speeds at specific roadway segments. However, the other metrics have more uncertainty and thus require more probe data. The simulation tests also verify that the analytical formulae of estimation accuracy provide a true indication of the uncertainty that should be expected. These results are consistent with tests done using a more realistic micro-simulation network of Orlando, Florida (32). To our knowledge, network-wide estimations of average density have only been conducted recently in Gayah and Dixit (29), which showed that probe penetration rates of 7.5% or lower produce accurate estimations; however, the method relies on the knowledge of the MFD a priori, which might not be realistic in many cities.

In addition, the method used in this study only provides accurate estimates when the network is near the critical density associated with the formation of congestion. The estimation methodology proposed here, however, appears to be accurate during the entire range of traffic states, including free-flow, congestion and even during unstable periods characterized by hysteresis phenomena. Therefore, the estimation procedure should be very robust and can even be used to directly estimate a network's Macroscopic Fundamental Diagram (MFD) and Network Exit Function (NEF). Estimations of the MFD and NEF in this way should also be more consistent and comparable across various networks as opposed to estimates that rely primarily on

loop detectors, since the latter method has shown to be highly sensitive to detector location.

Several limitations exist in this methodology. For one, it was assumed that probe vehicles maintain uniform spatial distributions across the network. In reality, some O-D pairs may have higher probe penetration rates, and this would result in heterogeneous distributions. Further work is being performed to characterize these heterogeneous spatial distributions and quantify how they might influence estimation uncertainty. For example, dynamic partitioning algorithms could be used to divide the network into subregions with uniform probe distributions to better estimate traffic conditions in the larger region. This work also assumed that enough detectors were available to accurately estimate the probe penetration rate. Future work should examine how uncertainty in this quantity can impact network traffic estimates. Furthermore, detector data and probe data can be combined in other ways than described here-for example, detectors can measure flows well in urban areas, and probes can be used solely for average speed measurements. However, this could create consistency problems since flows would only be measured for a subset of the network (i.e., only on certain links), while speeds would be measured across the entire network. This would provide MFDs or NEFs for a portion of the network but may not reflect conditions across the entire network. The method proposed in this paper uses probe data as much as possible and limits the use of detector data to avoid these potential inconsistencies while it measures conditions over the entire network over which vehicles are traveling. Still, these other potential combinations of detector and probe data are worthy of exploration.

ACKNOWLEDGMENTS

This research was supported by the Mid-Atlantic Universities Transportation Center. The authors would also like to thank Vinayak V. Dixit for his thoughtful discussions.

REFERENCES

- 1. Smeed, R. and Wardrop, J. An exploratory comparison of the advantages of cars and buses for travel in urban areas. *Journal of the Institute of Transport*, Vol. 30, No. 9, pp. 301-315, 1964.
- 2. Wardrop, J. Journey speed and flow in central urban areas. *Traffic Engineering and Control*, Vol. 9, pp. 528-532, 539, 1968.
- 3. Godfrey, J. The mechanism of a road network. *Traffic Engineering and Control*, Vol. 11, pp. 323-327, 1969.
- 4. Herman, R. and Prigogine, I. A two-fluid approach to town traffic. *Science*, Vol. 204, No. 4389, pp. 148-151, 1979.
- 5. Daganzo, C.F. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B*, Vol. 41, No. 1, pp. 49-62, 2007.
- 6. Geroliminis, N. and Daganzo, C.F. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B*, Vol. 42, No. 9, pp. 759-770, 2008.
- 7. Haddad, J. and Geroliminis, N. On the stability of traffic perimeter control in tworegion urban cities. *Transportation Research Part B*, Vol. 46, No. 9, pp. 1159-1176, 2012.
- 8. Zheng, N., Waraich, R.A., Axhausen, K.W., and Geroliminis, N. A dynamic cordon pricing scheme combining the Macroscopic Fundamental Diagram and an agent-based traffic model. *Transportation Research Part A*, Vol. 46, No. 8, pp. 1291-1303, 2012.
- 9. Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., and Papageorgiou, M. Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B*, Vol. 46, No. 10, pp. 1393-1403, 2012.
- 10. Geroliminis, N. and Levinson, D. Cordon pricing consistent with the physics of overcrowding. In: *Proceedings of the 18th International Symposium on Transportation and Traffic Theory*, 2009.
- 11. Gonzales, E.J. and Daganzo, C.F. Morning commute with competing modes and distributed demand: User equilibrium, system optimum, and pricing. *Transportation Research Part B*, Vol. 46, No. 10, pp. 1519-1534, 2012.
- 12. Daganzo, C.F., Gayah, V.V. and Gonzales, E.J. The potential of parsimonious models for understanding large scale transportation systems and answering big picture questions. *EURO Journal of Transportation and Logistics*, Vol. 1, No. 1-2, pp. 47-65, 2012.

- 13. Gayah, V.V. and Daganzo, C.F. Analytical capacity comparison of one-way and two-way signalized street networks. *Transportation Research Record*, No. 2301, pp. 76-85, 2012.
- 14. Knoop, V.L., Hoogendoorn, S.P., and Van Lint, J.W.C. Routing based strategies based on Macroscopic Fundamental Diagram, *Transportation Research Record*, No. 2315, pp. 1-10, 2012.
- 15. Keyvan-Ekbatani, M., Papageorgiou, M., and Papamichail, I. Urban congestion gating control based on reduced operational network fundamental diagrams. *Transportation Research Part C*, Vol. 33, pp. 74-87, 2013.
- 16. Geroliminis, N. and Daganzo, C.F. Macroscopic modeling of traffic in cities. In: *Proceedings of the 86th Annual Meeting of the Transportation Research Board*, 2007.
- 17. Buisson, C. and Ladier, C. Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. *Transportation Research Record*, No. 2124, pp. 127-136, 2009.
- 18. Geroliminis, N. and Sun, J. Hysteresis phenomena of a Macroscopic Fundamental Diagram in freeway networks. *Transportation Research Part A*, Vol. 45, pp. 966-979, 2011.
- Courbon, T. and Leclercq, L. Cross-comparison of macroscopic fundamental diagram estimation methods. *Procedia Social and Behavioral Sciences*, Vo. 20, No. 9, pp. 417-426, 2011.
- 20. Daganzo, C.F. and Geroliminis, N. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B*, Vol. 42, No. 9, pp. 771-781, 2008.
- 21. Gonzales, E.J., Chavis, C., Li, Y., and Daganzo, C.F. Multimodal transport modeling for Nairobi, Kenya: Insights and recommendations with an evidence based model. UC Berkeley, Volvo Working Paper UCB-ITS-VWP-2009-5, 2009.
- 22. Gayah, V.V. and Daganzo, C.F. Clockwise hysteresis loops in the macroscopic fundamental diagram: An effect of network instability. *Transportation Research Part B*, Vol. 45, No. 4, pp. 643-655, 2011.
- 23. Daganzo, C.F., Gayah, V.V. and Gonzales, E.J. Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. *Transportation Research Part B*, Vol. 45, No. 1, pp. 278-288, 2011.
- 24. Gayah, V.V. and Daganzo, C.F. Effects of turning maneuvers and route choice on a simple network. *Transportation Research Record*, No. 2249, pp. 15-19, 2011.
- 25. Mazloumian, A., Geroliminis, N., and Helbing, D. The spatial variability of vehicle densities as determinant of urban network capacity. Philosophical Transactions of the Royal Society, 368(1928), 4627-4647, 2010.

- 26. U.S. Department of Transportation, Federal Highway Administration. Connected Vehicle Research. Last Accessed June 25, 2013, http://www.its.dot.gov/connected_vehicle/connected_vehicle.htm.
- 27. Srinivasan, K.K. and Jovanis, P.P. Determination of number of probe vehicles required for reliable travel time measurement in urban network. *Transportation Research Record*, No. 1537, pp. 15-22, 1996.
- 28. Hofleitner, A., Herring R., Abbeel P., and Bayen A. Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network. IEEE Transactions on Intelligent Transportation Systems, Vol. 13. Issue 4, pp. 1679-1693, 2012.
- 29. Gayah, V.V. and Dixit, V.V. Using mobile probe data and the macroscopic fundamental diagram to estimate network densities: tests using microsimulation. Paper presented at the Transportation Research Board 92nd Annual Meeting, Washington, D.C., January 13-17, 2013.
- 30. Edie, L.C. Discussion of traffic stream measurements and definitions. In: *Proceedings of the 2nd International Symposium on the Theory of Traffic Flow*, 1965.
- 31. Herrera, J.C., Work, D.B., Herring, R., Ban, X., Jacobson, Q. and Bayen, A.M. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C*, Vol. 18, No. 4, pp. 568-583, 2010.
- 32. Nagle, A.S. and Gayah, V.V. A method to estimate the macroscopic fundamental diagram using limited mobile probe data. In: *Proceedings of the 16th International IEEE Conference on Intelligent Transport Systems*, 2013.

APPENDIX C

Deriving Macroscopic Fundamental Diagrams from probe data: Issues and proposed solutions

This article is currently under review and may be cited as a working version: Du, J., Rakha, H. and Gayah, V.V. Deriving Macroscopic Fundamental Diagrams from probe data: Issues and solutions. Working paper.

INTRODUCTION AND BACKGROUND

Network-wide traffic relationships have been the subject of study for at least several decades (Smeed, 1966, Godfrey, 1969, Zahavi, 1972, Herman and Prigogine, 1979, Ardekani and Herman, 1987, Mahmassani et al., 1987, Mahmassani et al., 1984, Olszewski et al., 1995). However, earlier efforts failed to provide a comprehensive model that was physically realistic, dynamic and verified by empirical data. Only recently did researchers finally verify with simulation and empirical data that a well-defined relationship exists between the average flow and density measured across an urban network (Geroliminis and Daganzo, 2008, Geroliminis and Sun, 2011a, Geroliminis and Daganzo, 2007, Daganzo, 2007, Daganzo and Geroliminis, 2008). This relationship, known now as the Network or Macroscopic Fundamental Diagram (NFD or MFD), is very helpful for researchers and traffic management agencies to monitor the status of a traffic network, design efficient traffic control strategies, and measure the effectiveness of network efficiency improvement strategies.

Recent research has examined several different applications of the MFD for improved traffic control. One set of examples include gating strategies that carefully limit vehicle inflow into a network to avoid congested states and maximize overall efficiency (Keyvan-Ekbatani et al., 2012, Keyvan-Ekbatani et al., 2013, Keyvan-Ekbatani et al., 2014, Haddad et al., 2013, Geroliminis et al., 2013). Other strategies include congestion pricing schemes that make use of the MFD to determine optimal pricing (Geroliminis and Levinson, 2009, Zheng et al., 2012, Simoni et al., 2015). In general, these control strategies are sensitive to the functional form of the MFD, and several studies have examined the factors that influence the attributes of an MFD. Geroliminis and Sun studied the spatial variability of vehicle density and found that it affects the shape, the scatter, and the existence of an MFD (Geroliminis and Sun, 2011b). Various researchers have also explored how driver routing influences the shape and reliability of MFDs (Saberi et al., 2014, Daganzo et al., 2011, Gayah and Daganzo, 2011, Mahmassani et al., 2013). Studies of arterial road networks controlled by different types of adaptive traffic signal systems reveal that the shape of the MFD depends on the particular signal system used and the level of heterogeneity in the system (Gayah et al., 2014, Zhang et al., 2013). Ji and Geroliminis (Ji and Geroliminis, 2012) studied how networks could be partitioned to obtain more reliable MFDs based on the variance of link densities and spatial compactness of the partitioned regions. Buisson and Ladier (Buisson and Ladier, 2009) found that measurement heterogeneity, such as differences between the surface and highway network and distances between the loop detector and traffic signal, has a strong impact on the shape of the MFD.

While it is convenient to use an MFD to describe the traffic status across a network and design traffic control strategies, the data needed to plot the MFD are not always readily available. In theory, the MFD of a network can be easily calculated if the trajectories of all vehicles traveling in the network are known. However, trajectories for all vehicle are generally impossible to acquire. Detector data may be used to estimate MFDs, but detector-based MFDs can have significant errors over the generalized

definitions and are dependent on where the detector is placed on each link (Courbon and Leclercq, 2011, Leclercq et al., 2014). Fortunately, the availability of GPS technology has made it relatively simple to record the kinematic data from at least a portion of the vehicles traveling in the network. Recent research has shown that the trajectories of a subset of vehicles traveling in the network can be used to satisfactorily estimate current traffic states (Gayah and Dixit, 2013) or even the overall MFD (Nagle and Gayah, 2014) in a way that is both reliable and accurate when probe vehicles are more or less uniformly distributed across a network. Leclercq et al. also explored the combination of detector data for flow calculations and detector data for average speed calculations (Leclercq et al., 2014), which will also work well if probe vehicles are uniformly distributed across a network. Unfortunately, the assumption of uniform probe vehicle distributions might be too restrictive in practice. GPS technologies are likely to be included in newer vehicles and certain origin and/or destinations are thus more likely to have higher penetration rates of these types of vehicles than others. Methods to estimate the MFD are needed that account for this heterogeneous distribution of probe vehicles across a network to estimate more reliable MFDs.

In this study, we test the feasibility of using probe information to estimate MFDs when the probe penetration rate is not uniform across the network and when the demands in the network are unbalanced. The generalized definitions of traffic flow parameters are used to develop proper weighted averages of probe penetration rate that should be used to estimate average flow and density in a network assuming the probe penetration rates are known for specific OD pairs or regions. Simulation results using INTEGRATION (Van Aerde and Rakha, 2013a, Van Aerde and Rakha, 2013b, Chamberlayne et al., 2012, Rakha et al., 2012, Rakha et al., 2004, Rakha and Zhang, 2004b, Van Aerde et al., 1996, Rakha and Zhang, 2004a) verify that these weighted averages of probe penetration rate provide an obvious advantage over simple arithmetic average probe penetration rates in estimating MFDs from a subset of probe vehicle information. Furthermore, an algorithm that combines fixed detector volume count data and probe vehicle travel times and travel distances is proposed to estimate the probe penetration rate of different regional OD pairs since this information is not likely to be known a priori. The algorithm relies on data from just a limited number of links in the network, in this paper tested for just 10% of the links, in addition to the data from the few probe vehicles themselves. It works well when the probe penetration rate varies from area to area, which is promising for heterogeneous distributions of probe vehicles in a network. Given the availability of fixed detector data and probe vehicle data in most networks, this algorithm is significantly practical for advancing the algorithm of estimating MFD using probe data from a simulation environment to the real world.

The rest of this paper is organized as follows. First, the weighted probe penetration rates needed are derived from the generalized definitions of traffic flow parameters. Then, the idealized micro-simulation test network used to test our methods and data that are used in the methods are described. The results of the proposed algorithm using trajectory data in the simulation results are then described and compared for different scenarios. Following this, a practical method for estimating probe penetration rates for different regional OD pairs is proposed and tested. Finally, the conclusions of the paper are summarized.

MFD ESTIMATION USING PROBE VEHICLES

The MFD provides a relationship between network-wide averages of flow and density, both of which can be used to describe operating conditions within a network. The remainder of this section describes how the MFD can be estimated using probe data when the probe penetration rate is consistent across the network and then extends this methodology to varied probe penetration rates across individual regions or OD pairs.

Estimation of MFD Assuming Uniform Probe Penetration Rates

The generalized definitions of Edie (Edie, 1965) can be used to calculate the average

density (k) and flow (q) in a network, respectively, as follows:

$$k = \frac{\sum_{i=1}^{I} t_{i}}{L_{n} * T} \tag{1}$$

$$q = \frac{\sum_{i}^{I} d_{i}}{L_{n} * T} , \tag{2}$$

where *I* is the total number of trips recorded in that analysis period (e.g. 15 minutes);

 t_i and d_i are the travel time (seconds) and distance (miles), respectively, for trip i; and,

 L_n and T are the network length (miles) and analysis period length (seconds),

respectively. Equations (1) and (2) only differ in the numerator: the definition of density uses the total time vehicles spend traveling within the network during the analysis period as the numerator, while the definition of flow uses the total distance vehicles travel during the analysis period as the numerator.

In most real-world applications, the total time and distance traveled will only be known if detailed trajectories of all vehicles are provided. However, if these data are only provided for a subset of vehicles in the network (i.e., those serving as mobile probes) then Equations (1) and (2) cannot be applied directly. To overcome this limitation, Nagle and Gayah (2014) proposed the following equations to estimate density and flow in a network if the fraction of vehicles serving as mobile probes is

known (*p*) and the same across individual regions or OD pairs:

$$\widehat{k} = \frac{\sum_{1}^{l'} t_{il'}}{\rho L_n * T} \tag{3}$$

$$\widehat{q} = \frac{\sum_{i}^{l'} d_{i'}}{\rho L_n \ast T} \tag{4}$$

where t_{ij} and d_{ij} are the travel time (seconds) and distance (miles), respectively, for

probe vehicle *i*'; and, *I*' is the total number of probe vehicles recorded within the

analysis period. This method inherently assumes that the average travel time and travel distance of probe and non-probe vehicles are the same, which would be true if probe vehicles are uniformly distributed across the network. The total travel time (distance) is then approximated by scaling the travel time (distance) of probe vehicles by the

penetration rate (p).¹

Estimation of MFD from Varied Probe Penetration Rates

Although convenient for modeling purposes, the assumption of a single constant probe penetration rate that describes probe distribution across the entire network is not realistic in the real world. For one, vehicles with GPS devices that can be used as probe vehicles are distributed non-uniformly in real traffic networks. Generally, more affluent areas usually have a higher rate of well-equipped vehicles that can collect travel time and speed information that can be used as probe data to inform network estimations. Additionally, the running frequencies of fleet vehicles that are used as probe vehicles (such as taxis, freight vehicles, and buses), vary by time over the course of the day and, in many cases, the routes used by such vehicles are imbalanced in the network. Consequently, the assumption of a single uniform penetration rate may not accurately represent the distribution of probe vehicles throughout the network.

We now propose a method to overcome this limitation that specifically addresses the issue of non-uniform probe penetration rates across regions or OD pairs in a network. We start by assuming that the probe penetration rate for vehicles

¹ Note that the probe penetration rate may change with time and thus should have a subscript *t* to reflect this

dependence. Instead, all equations are provided for a single analysis period and thus this subscript is removed to simplify the presentation of equations provided.

traveling between a specific origin o and destination d, ρ_{od} , is fixed for some analysis period and is known a priori².

Let us first focus on the estimate of density, which is proportional to the total travel time of all vehicles in the network (Equation 1). Define the total travel time of all vehicles traveling between OD pair **od** within a particular analysis interval as T_{od} and the total travel time of probe vehicles traveling between this OD pair within the same time period as T'_{od} . We assume that the average travel time for probe and non-probe vehicles between any OD pair is equal or very similar, which is realistic if probe and non-probe drivers generally act in the same way. In this case, $T_{od} = T'_{od}/\rho_{od}$.

Furthermore, the total travel time for all vehicles is $T = \sum_{o,d}^{O,D} (T'_{od}/\rho_{od})$ where O and D represent the total number of origins and destinations, respectively.

Equation (3) suggests that $T = \sum_{1}^{l'} t_{i'} / \rho_{eq}^{k}$, where we use ρ_{eq}^{k} to refer to an equivalent average penetration rate of probe vehicles in the network used to estimate density. These two definitions of T can be combined to estimate this equivalent average penetration rate:

$$\rho_{eq}^{k} = \frac{\sum_{i=1}^{l'} t_{i'}}{\sum_{o,d}^{0,DT'od} / \rho_{od}},$$
(5)

which turns out to be the harmonic mean of penetration rates of individual OD pairs using the distance traveled by probe vehicles between each OD pair as the weight. This average penetration can then be substituted into Equation (3) for *p* to provide a more accurate estimate of density when probe penetration rates vary within the network.

Similar logic can be used to provide an equivalent average probe penetration rate to estimate flow:

² This latter assumption is especially restrictive as this quantity will not be known for all OD pairs in a network; therefore, a method to estimate these values is proposed in one of the subsequent sections of this paper. However, for now, such an assumption is used to simplify the description of the estimation procedure.

$$\rho_{eq}^{q} = \frac{\sum_{i}^{l'} d_{i'}}{\sum_{o,d}^{0,DD'} od/\rho_{od}},$$

where D'_{od} represents the total distance traveled by probe vehicles between any OD

pair od. This average penetration rate can then be substituted into Equation (4) for p to

then provide a more accurate estimate of flow when probe penetration rates vary within the network.

RESULTS FROM COMBINATIONS OF VARIED PROBE PENETRATION

RATES

In this section, a set of simulation tests are used to verify that that weighted average penetration rates provided by Equations (5) and (6) that take into account heterogeneous probe distributions yield more accurate estimates of flow and density in a network than the results generated under the assumption of a uniform penetration rate across the network. We first describe the simulation framework and accuracy measures. Then, we examine the results for various heterogeneous probe penetration rates when traffic demands are perfect uniform. Lastly, we examine the results for various heterogeneous probe penetration rates when traffic demands are also heterogeneous.

Network Description and Measures of MFD Accuracy

An idealized 16 × 16 square grid network with alternating one-way streets is used for this study. The network consists of a 544 links, each 400 feet long. For simplicity, origins and destinations are placed at all entry and exit links, respectively; therefore, a total of 64 origin/destination zones are used. A uniform demand is initially generated between each origin-destination (OD) pair. A 3-hour simulation is used in which the first hour represents a warm-up period, the second hour is the peak demand rate, and the third hour is a recovery period. The demand rate for the first and third hours is 6 vehicles/hour per OD pair, while the peak demand is 17 trips/hour for each OD pair.

To evaluate the accuracy of the estimated MFD, three measures of effectiveness (MOEs) that incorporate the root mean square error (RMSE) of the average flow, density, and flow plus density are proposed (Nagle and Gayah, 2104).

$$RMSE(q) = \sqrt{\sum(q - \hat{q})^2/N}$$
(7)

$$RMSE(k) = \sqrt{\sum (k - \hat{k})^2 / N}$$
(8)

$$RMSE(q,k) = \sqrt{\sum \left[\left(\frac{q-\hat{q}}{q_o} \right)^2 + \left(\frac{k-\hat{k}}{k_j} \right)^2 \right] / N}$$
(9)

Here q and k are the real flow rate and density calculated using all the trajectory data as

per the generalized definitions of Edie, \hat{q} and \hat{k} are the estimated flow rate and density calculated using Equations (3) and (4), where the market penetration rate is estimated using one of the weighted average probe penetration rates from Equations (5) or (6).

The terms q_{o} and k_{i} are the maximum flow and jam density observed in the network,

respectively. *N* is the number of time intervals within the simulation.

Varied Penetration Rate by OD

Six scenarios were considered in which the penetration rates of mobile probe vehicles differed based on the OD pairs to examine the effects of various overall penetration rates and levels of heterogeneity between penetration rates for OD pairs in different regions. To simplify the analysis, the network was broken down into the four quadrants shown in FIGURE 1. Probe penetration rates for specific OD pairs were defined based on the regional origins and destinations. The specific scenarios considered were:

- A. ODs from region **I** to region **II** have a penetration rate of 0.8. All other ODs have a penetration rate of 0.1.
- B. ODs from region I to region II have a penetration rate of 0.5. All the other ODs have a penetration rate of 0.1.
- C. ODs within region I have a penetration rate of 0.8. All the other ODs have a penetration rate of 0.1.
- D. ODs within region I have a penetration rate of 0.5. All the other ODs have a penetration rate of 0.1.
- E. ODs that have origins in region I have a penetration rate of 0.8. All the other ODs have a penetration rate of 0.1.
- F. ODs that have origins in region **I** have a penetration rate of 0.5. All the other ODs have a penetration rate of 0.1.

The scenarios are designed to represent idealized probe vehicle distribution patterns in typical networks. Scenarios A and B are designed to represent the situation where certain OD pairs have higher penetration rates than the rest of the areas. Scenarios C and D are representatives of situations where a condensed area has a higher probe penetration rate compared to other areas. Scenarios E and F are used to simulate the situation where affluent areas have a higher market penetration rate of probe vehicles. The higher probe penetration rates are set to 0.8 and 0.5, individually, versus the lower penetration rate of 0.1 to examine the impacts of a difference in probe penetration rates by regions on the results.



FIGURE 1: Idealized grid network with varied penetration rate.

The simulation was run once and the trajectory of all vehicles were extracted to determine the true MFD. Subsets of the vehicle trajectories were then extracted based on the scenarios above to provide the data used to estimate the MFD based on Equations (3) through (6). To eliminate the fluctuations generated by the randomness of sampling, the process of sampling was repeated 50 times and the average RMSE is calculated over the 50 samples. FIGURE 2 presents the estimated MFD using the proposed methodology and the true MFD using one of the 50 samples. Table 1 presents the MOEs of estimation accuracy for all MFDs obtained for Scenarios A through F using the average results of the 50 samples. Note that for all MOEs, smaller values indicate better performance.



FIGURE 2: MFD for Scenario A to F

	RMSE	Scenario A	Scenario B	Scenario C	Scenario D	Scenario E	Scenario F
Arithmetic Average Rate	Flow Rate (Veh/h)	41.90	28.58	66.83	50.73	14.36	14.66
	Density (veh/mi)	2.02	1.66	4.84	3.73	2.07	2.13
	RMSE (q,k)	0.05	0.03	0.08	0.06	0.02	0.02
Waightad	Flow Rate (veh/h)	22.47	23.72	25.22	25.20	15.59	15.83
Average	Density (veh/mi)	2.12	2.02	1.82	1.87	1.66	1.72
	RMSE (q,k)	0.03	0.03	0.03	0.03	0.02	0.02

TABLE 1: RMSE for Scenarios A-F

As can be seen, the weighted probe penetration rate outperforms the arithmetic average penetration rate for almost all cases. Table 2 lists the Two-Sample T-test results for the six scenarios. Except for scenarios E and F, where the weighted average method is similar to the arithmetic average method, in all the other four scenarios, the RMSE generated by the arithmetic average method is significantly larger than the weighted average. The results indicate that:

TABLE 2: Two Sample T-Test for the Difference of RMSEs
--

	Scenario	Scenario	Scenario	Scenario	Scenario	Scenario
	А	В	С	D	Е	F
Mean						
	0.0211	0.00477	0.0493	0.0305	-2E-05	0.00009
(Difference of RMSE)						
t Value	27.25	(7)	124.20	75 70	0.10	0.(7
t value	27.25	0.72	124.39	/5./8	-0.19	0.67
Pr > t	<.0001	<.0001	<.0001	<.0001	0.8532	0.5038

- 1) When the higher probe penetration rate is accompanied with OD pairs that are not distributed evenly within the network, such as scenarios A through D, the advantages of using a weighted probe penetration rate increases with the imbalance of probe penetration rates across individual OD pairs.
- 2) When a higher probe penetration rate exists in ODs that are distributed evenly across the entire network, the imbalance between higher and lower probe penetration rates will not affect the accuracy of the estimations as significantly. This can be observed in scenarios E and F, as higher penetration rates exist for trips starting from region I are distributed to all other locations in the grid network. Under these circumstances, the weighted average probe rate does not have any significant advantage over the results estimated using arithmetic average probe rate.
- 3) When a higher probe penetration rate exists in ODs that are limited in a concentrated subnetwork, as in the case of C and D, the advantage of the results estimated using weighted average probe penetration rates is significant.

For scenarios A and C, the difference in penetration rate for different ODs is bigger (0.8 versus 0.1) compared to scenarios B and D. The calculation with the weighted probe penetration rates provides a significantly higher estimation accuracy. We further explore this phenomenon by fixing the total number of probe vehicles and just varying the distribution of penetration rate in different regions. We start by assuming that the average probe penetration rate is 20%. Keeping this average penetration rate fixed, we then increase the percentage of the probes for the ODs in region I from 0.2 to 0.9. While the percentage of probes in region I increases, the probes from other regions decreases accordingly to maintain the same overall average penetration rate at 20%. The RMSEs are then calculated for each combination of probe penetration rates. TABLE 3 illustrates the results using the arithmetic average and the proposed weighted harmonic average penetration rate in each case. FIGURE 3 shows the resulted RMSE of using the arithmetic average and the proposed weighted average penetration rate. The second vertical axis with the orange triangles represents the percentage improvement associated with the use of the weighted average rate. As can be seen, more significant advantages are achieved using the weighted average penetration rate the larger the difference exists between the higher penetration rates versus the lower penetration rates. Additionally, the larger the difference in penetration rates is by OD pairs, the poorer the results generated using the arithmetic average penetration rate. These results confirm the significant negative impacts generated by the heterogeneous distribution of probe penetration rates in the network, especially if the arithmetic mean is used in estimating the MFD assuming a homogenous penetration rate.

	High Rate in								
	Region I	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
	Total Prohes	596	594	593	594	592	589	593	594
	Total Trobes	1	3	7	7	5	6	4	9
	Flow Rate	14.2	17.0	19.4	23.0	27.4	33.3	37.4	44.0
Arithmetic	(veh/h)	4	3	7	7	8	0	9	3
Average	Density								
Rate	(veh/mi)	1.06	1.24	1.40	1.71	1.91	2.34	2.57	3.08
	RMSE (q,k)	0.02	0.02	0.02	0.03	0.03	0.04	0.04	0.05
	Flow Rate	14.1	15.8	15.7	16.0	15.0	16.5	15.7	18.0
Maightad	(veh/h)	6	5	7	0	7	6	3	1
Average	Density								
Average	(veh/mi)	1.04	1.17	1.18	1.17	1.06	1.18	1.13	1.30
	RMSE (q,k)	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

TABLE 3: Impacts of Imbalanced Penetration Rates





Varied Penetration Rate by OD with Imbalanced Congestion Levels

The tests performed in Section 3.2 are all based on uniform demand conditions, where each OD pair generates the same number of trips. However, in reality network demands are unbalanced (and often highly unbalanced), which results in some areas in a network being more congested than others. This congestion can be contained within a specific area in the network, such as the downtown central business district (CBD) or along a certain corridor. How this imbalanced demand will jointly affect the estimated MFD using varied probe penetration rates will be explored in this section. For these tests, the following scenarios are tested:

I. The original balanced ODs are changed in such a way that if the ODs are within region I, the demand increases by 50% while all the other ODs decrease to 90% of the original demand. The penetration rate for ODs within region I is 0.8 while for all the others it is 0.1.

II. The demand stays the same as scenario I, but the penetration rates are changed to 0.5 versus 0.1.

III. The demand stays the same as scenario I, but the penetration rates are changed to 0.3 versus 0.1.

The results, shown in TABLE 4 below, indicate that when the congestion in the network is contained within a specific area, using the weighted penetration rate does have an advantage over the arithmetic average. However, this advantage diminishes as the difference in penetration rates decreases between the congested and other areas. This observation is meaningful to the practice of traffic congestion monitoring and control in that usually reproducible congestion occurs within a specific area (such as the downtown area of urban networks) while at the same time the mobile probe penetration rate in such areas might be much higher than the rest of the network. Consequently, the advantage of using a weighted penetration rate will be significant in improving network monitoring in such realistic situations.

	RMSE	0.8/0.1 Upper Right OD/Others (Scenario I)	0.5/0.1 Upper Right OD/Others (Scenario II)	0.3/0.1 Upper Right OD/Others (Scenario III)
Arithmetic	Flow Rate (Veh/h)	42.63	37.02	32.30
Average Rate	Density (Veh/Mile)	1.94	1.91	1.96
	RMSE (q,k)	0.05	0.04	0.04
Weighted	Flow Rate (Veh/h)	24.90	26.32	27.11
Average	Density (Veh/Mile)	1.91	1.88	1.96
	RMSE (q,k)	0.03	0.03	0.03

TABLE 4: RMSE for Imbalanced Demands and Varied Penetration Rate in Different Areas(Areas)

Another set of scenarios are tested where the congestion within the network is limited to a specific corridor as described below:

IV. Trips originating from region I and traveling to region II are increased to 1.5 times the original numbers while all the others stay the same. The penetration rate for ODs from region I to region II is 0.8, while for all the others it is 0.1.

V. The demands stays the same as scenario IV, but the penetration rates are changed to 0.5 versus 0.1.

VI. The demand stays the same as scenario IV, but the penetration rates are changed to 0.3 versus 0.1.

TABLE 5 illustrates the results for the corridor congestion case. The weighted average probe penetration rate yields significantly better results.

	RMSE	0.8/0.1 Diagonal OD/Others	0.5/0.1 Diagonal OD /Others	0.3/0.1 Diagonal OD/Others
		(Scenario IV)	(Scenario V)	(Scenario VI)
Arithmotic	Flow Rate (Veh/h)	104.99	66.31	36.21
Average Rate	Density (veh/mi)	6.30	3.96	2.24
	RMSE (q,k)	0.12	0.08	0.04
	Flow Rate (veh/h)	22.79	23.79	24.92
Weighted Average	Density (veh/mi)	1.92	1.92	1.91
	RMSE (q,k)	0.03	0.03	0.03

TABLE 5: RMSE for Imbalanced Demands and Varied Penetration Rate in Different Areas(Corridors)

In reality, the congested area does not necessarily cover the area where the probe penetration rate is higher. For example, the congested area may be along the corridor, but the ODs within an affluent residential area may have a higher penetration rate. Consequently, one more set of scenarios is tested where the OD demands are identical with the case of IV to VI. The combinations of penetration rates are varied as shown in TABLE 6. Again, the weighted average probe penetration rate performs better than the arithmetic average rate across the board.
	RMSE	0.8/0.1 Upper Right OD/Others	0.5/0.1 Upper Right OD/Others	0.3/0.1 Upper Right OD/Others
Anithmatia	Flow Rate (veh/h)	71.85	54.15	40.12
Average Rate	Density (veh/mi)	4.74	3.65	2.76
	RMSE (q,k)	0.08	0.06	0.05
	Flow Rate (veh/h)	26.29	26.34	26.00
Weighted Average	Density (veh/mi)	1.91	1.87	1.86
	RMSE (q,k)	0.03	0.03	0.03

TABLE 6: RMSE for Congested Area Occurred Outside of Higher Penetration Rate Area

PROBE PENETRATION RATE ESTIMATION

The discussions above show that using a weighted average penetration rate has an advantage over the arithmetic average penetration rate in estimating the MFD. This advantage is especially obvious when the probe penetration rate is not fixed across the network and/or the demand is very imbalanced. The analysis above, however, is based on the assumption that the probe market penetration rates for each OD pair are known a priori. The calculation of the MFD used data extracted at a known percentage as the probe market penetration rate. Therefore, it is vital to accurately estimate the penetration rate in a more practical way when such parameters are unknown. Previous research showed that, assuming a uniform probe penetration rate, it is reliable to estimate the penetration rate by sampling probe vehicles at certain fixed traffic detectors in the network (Nagle and Gayah, 2014). The aggregate count of all vehicles,

 N^{d} , and probe vehicles, N_{p}^{d} , that crossed a detector was determined and used to

calculate the penetration rate $\hat{\rho}$. The method was tested to be effective and accurate

when the probe penetration rate is uniform across the network. However, since the penetration rate is not uniform in this study, as would be the case in a real world application, a method is needed to generate an aggregated number to be used in Equations (3) through (6) as the aggregated pseudo penetration rate.

A close study of the detected probe rates at some randomly selected links and the number of probe trips passing those links and aggregated by OD pair shows that there is a similar pattern in the two groups. FIGURE 4 shows the histograms of the link

detected probe penetration rate (*p*) at 10% of the randomly picked links in the network

(a) and the aggregated count of probe vehicles passing those links by OD pairs (b). Both plots show a "High-Medium-Low" pattern. When the penetration rate is not a uniform number, it is reasonable to foresee that links that are closer to the origins where the higher probe penetration rates are located should be more likely to have a higher

detected probe penetration rate $\hat{\rho}$. At the same time, when the trips made by probe

vehicles are aggregated by OD pairs, higher penetration rates for certain OD pairs are more likely to have larger observations of trips in the network. Consequently, we propose an algorithm to match the two sets (observed penetration rates and observed counts of probe vehicles by OD pairs) using k-means clustering analysis, which

partitions observations into n clusters in which the within-cluster sum of squares is

minimized. These k-mean results are then used to calculate the pseudo weighted average probe penetration rate. The following steps were adopted:

1) Use k-means clustering analysis to group the detected penetration rate on randomly picked links in the network into **n** clusters and calculate the **n**-cluster means of the detected penetration rate. The result will be vector 1;

2) Aggregate the number of probe vehicles passing those links by OD pairs;

3) Use k-means clustering analysis to group the trip number counts (by OD

pairs) into **n** clusters and calculate the **n**-cluster trip count mean. The result will be vector 2;

4) Sort vector 1 and vector 2 generated in step 1 and step 3, individually. Pair the two sorted vectors such that each element in vector 1 will have a corresponding counterpart at the same rank position in vector 2. Assign each OD pair in vector 2 the corresponding mean penetration rate value at the same ranking position in vector 1; and,

5) Calculate the weighted average penetration rate by summing over the clusters using Equations (5) and (6).

To illustrate how this algorithm works, we now set the number of clusters to three as an example (as illustrated in FIGURE 4). Three unique groups are identified in FIGURE 4(a) (ρ_{max} , ρ_{median} , ρ_{min}) and three groups are identified in FIGURE 4(b)

(*Count_{max}*, *Count_{median}*, *Count_{min}*). They are paired together such that the highest

values in each vector are matched together, median values are matched together, and minimum values are paired as well. FIGURE 5 is the flowchart when the cluster number is set at 3. The results shown in TABLE 7 illustrate a robust output using weighed average probe rates. The magnitude of the errors is comparable with the calculations above; however, the advantage here is that these results do not require knowledge of the probe penetration rate by OD pair a priori.



FIGURE 4: Histograms of (a) Detected penetration rate and (b) Observed aggregated probe trips.



FIGURE 5: K-mean clustering analysis for estimating weighted average probe penetration rate.

TABLE 7: RMSE of MFD from Estimated Weighted Average Probe Penetration Rate

Arithmetic	Flow Rate (veh/h)	66.83							
Average Rate	Density (veh/mi)	4.84							
	RMSE (q,k)	0.08							
		2- Clusters	3- Clusters	4- Clusters	5- Clusters	6- Clusters			
Weighted by Travel Time	Flow Rate (veh/h)	199.74	28.61	33.12	111.67	276.08			
	Density (veh/mi)	16.03	2.10	2.43	2.71	3.48			
	RMSE (q,k)	0.24	0.03	0.04	0.13	0.31			

(0.8/0.1 Upper Right OD/Others)

To examine the impacts that the number of clusters selected may have on the

accuracy level of this proposed method, four additional values of *n* are further tested:

two, four, five, and six. The RMSE of each case is plotted in FIGURE 6 and also listed in TABLE 7. These results all verify that the optimal number of clusters that should be selected is 3. FIGURE 7 shows the MFD resulting from this method (which assumes the demand pattern from Scenario C described in Section 3.2). Notice that the matching k-means clustering analysis proves to be very effective in estimating the MFD. Comparing to the results generated when the penetration rate is a known priori (Scenario C from Table 1), the RMSE only increases slightly.



FIGURE 6: RMSE using varied number of clusters



FIGURE 7: MFD Estimated Using Cluster Method

CONCLUSIONS AND DISCUSSION

In this study, a methodology is proposed to estimate a network's MFD using data from a limited number of probe vehicle trajectories when probe penetration rates are not uniformly distributed across a network. This methodology incorporates the distance and time probe vehicles traverse a network into the calculation of the equivalent probe penetration rates observed in the network using harmonic means. These equivalent penetration rates are then used to estimate network-wide traffic quantities such as average flow and average density using aggregated data obtained from all mobile probe vehicles. When compared with using the simple arithmetic mean of probe penetration rates previously proposed in the literature, the proposed weighted average method is generally more accurate. This is especially true in situations when the demands in the network are unbalanced and when the penetration rates in the network vary significantly from area to another. Scenarios tested in this study are designed to simulate realistic typical congestion phenomena such as a congested CBD area or a congested corridor. The varied penetration rates across the network are designed to represent situations that include higher penetration rates in affluent areas where travelers may travel within the area (Region I), travel to certain destinations in the network (Region I to region II), or the destinations of the probe vehicles are distributed evenly over the entire network (origins from Region I). Since the imbalanced demands and varied penetration rates described in the scenarios in this study are very common in the real world, the results from this study are of practical importance. The results also show that when the imbalance of penetration rate in the network increases, the errors associated with the estimation of the MFD also increases. The same conclusion holds when the demand is imbalanced in the network. The advantage of using distance- and time-weighted harmonic average penetration rates is more significant under these circumstances. The benefit of the proposed methodology is the most significant when the penetration rate imbalance and the variation of probe travel time and distance is the largest as can be seen from Scenarios A and C in TABLE 1. As would be expected, the advantage of the proposed method is not as significant when the probe travel time and distance is more balanced comparing to others, as in the case of Scenarios E and F,.

Another observation is that errors in the estimates of the MFD vary significantly using the arithmetic average penetration rate especially when the spatial heterogeneity of traffic demand and penetrate rates in the network is large. On the contrary, the methodology proposed in this study generates a more stable errors regardless of these imbalances in either demand or penetration rate. This stability of the proposed methodology makes it a more confident estimation tool that is robust to be used in different networks with different congestion patterns.

Because it is not realistic to know the exact probe penetration rate for each OD pair across the network in practical applications, this study proposes an algorithm to estimate the weighted probe penetration rate using fixed detector data and sample probe data. This methodology involves linking two k-means clustering analyses: one for the detected percentage of probe vehicles passing randomly selected links in the network and the other for the probe trip counts by OD pairs. The significance of this proposed algorithm is that the two data sources are likely to be easily collected or readily available in practice. Estimating the probe penetration rate by combining loop detector counts with the probe vehicle traveling data, such as origins, destinations, travel times, and distances, is feasible and realistic. The resulting MFD using the

estimated weighted average probe penetration rate generates a much smaller RMSE compared to using an average link detected probe penetration rate.

As with any research effort, further research is required as summarized below:

- 1. Identifying the impacts of locations of the links where the detectors should be placed to estimate probe penetration rates. In this study, the selection of such links was completely random. In the next step of research, identifying locations where the detectors should be placed to better capture the probe vehicle information needs to be studied. Preliminary suggestions for selecting links are to find locations that are geographically dispersed across the network, experiencing varied congestion levels that represent the traffic conditions in the adjacent areas, or highly likely to be used by re-routing drivers when the congestion occurs. Furthermore, how varied number of links with detectors will impact the results and how the number links is related to the optimum number of clusters will also be examined.
- 2. Improving the k-mean clustering model by identifying the number of clusters based on features of the network and/or demands in the network. In the work of Leclercq et al. (Leclercq et al., 2014), they concluded that due to the spillbacks, using loop dector data to estimate the MFD is not a optimum solution comparing to probe data. In the next step of the research, we will vary the penetration rates in the network such that there are more varied penetration rates existing in different areas of the network to identify and quantify the impacts this might have on the proposed methodology
- 3. Studying the changes and shifting on MFDs generated by network-wide traffic control strategies and better understanding how the MFD can be used as a monitoring and control tool to alleviate traffic congestion.

ACKNOWLEDGEMENTS

This research was supported by the Mid-Atlantic University Transportation Center (MAUTC). The authors would like to thank the anonymous reviewers for their constructive feedback and input they provided us. The paper has benefited significantly from their input.

REFERENCES

- ARDEKANI, S. & HERMAN, R. 1987. Urban network-wide traffic variables and their relations. *Transportation Science*, 21, 1-16.
- BUISSON, C. & LADIER, C. 2009. Exploring the Impact of Homogeneity of Traffic Measurements on the Existence of Macroscopic Fundamental Diagrams. *Transportation Research Record*, 2124, 127-136.
- CHAMBERLAYNE, E., RAKHA, H. & BISH, D. 2012. Modeling the Capacity drop Phenomenon at Freeway Bottlenecks using the INTEGRATION Software. *Transportation Letters: The International Journal of Transportation Research*, 4, 227-242.
- COURBON, T. & LECLERCQ, L. 2011. Cross-comparison of macroscopic fundamental diagram estimation methods. *Procedia Social and Behavioral Sciences*, Vol. 20, pp. 417 426.
- DAGANZO, C. F. 2007. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, 41, 49-62.
- DAGANZO, C. F., GAYAH, V. V. & GONZALES, E. J. 2011. Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. *Transportation Research Part B: Methodological*, 45, 278-288.
- DAGANZO, C. F. & GEROLIMINIS, N. 2008. An Analytical Approximation for the Macroscopic Fundamental Diagram of Urban Traffic. *Transportation Research Part B: Methodological*, 42, 771-781.
- EDIE, L. C. Discussion of traffic stream measurements and definitions. the 2nd International Symposium on the Theory of Traffic Flow, 1965.
- GAYAH, V. & DIXIT, V. 2013. Using Mobile Probe Data and the Macroscopic Fundamental Dagram to Esimate Network Densities. *Transportation Research Record*, 2390, 76-86.
- GAYAH, V. V. & DAGANZO, C. F. 2011. Clockwise hysteresis loops in the Macroscopic Fundamental Diagram: An effect of network instability. *Transportation Research Part B: Methodological*, 45, 643-655.
- GAYAH, V. V., GAO, X. & NAGLE, A. S. 2014. On the impacts of locally adaptive signal control on urban network stability and the Macroscopic Fundamental Diagram. *Transportation Research Part B: Methodological*, 70, 255-268.
- GEROLIMINIS, N. & DAGANZO, C. F. 2007. Macroscopic modeling of traffic in cities. 86th Annual Meeting of the Transportation Research Board. Washington, D.C.
- GEROLIMINIS, N. & DAGANZO, C. F. 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42, 759-770.
- GEROLIMINIS, N., HADDAD, J. & RAMEZANI, M. 2013. Optimal Perimeter Control for Two Urban Regions With Macroscopic Fundamental Diagrams: A Model Predictive Approach. *IEEE Transactions on Intelligent Transportation Systems*, 14, 348-360.
- GEROLIMINIS, N. & LEVINSON, D. 2009. Cordon Pricing Consistent with the Physics of Overcrowding. *Transportation and Traffic Theory 2009.* Springer Science.
- GEROLIMINIS, N. & SUN, J. 2011a. Hysteresis phenomena of a Macroscopic Fundamental Diagram in freeway networks. *Transportation Research Part A: Policy and Practice*, 45, 966-979.
- GEROLIMINIS, N. & SUN, J. 2011b. Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Research Part B: Methodological*, 45, 605-617.

- GODFREY, J. 1969. The mechanism of a road network. *Traffic Engineering and Control 11, 323–327.*, 11, 323–327.
- HADDAD, J., RAMEZANI, M. & GEROLIMINIS, N. 2013. Cooperative traffic control of a mixed network with two urban regions and a freeway. *Transportation Research B*, 54, 17-36.
- HERMAN, R. & PRIGOGINE, I. 1979. A two-fluid approach to town traffic. Science, 204, 148-181.
- JI, Y. & GEROLIMINIS, N. 2012. On the spatial partitioning of urban transportation networks. *Transportation Research Part B: Methodological*, 46, 1639-1656.
- KEYVAN-EKBATANI, M., KOUVELAS, A., PAPAMICHAIL, I. & PAPAGEORGIOU, M. 2012. Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B: Methodological*, 46, 1393-1403.
- KEYVAN-EKBATANI, M., PAPAGEORGIOU, M. & PAPAMICHAIL, I. 2013. Urban congestion gating control based on reduced operational network fundamental diagrams. *Transportation Research Part C*, 33, 74-87.
- KEYVAN-EKBATANI, M., PAPAGEORGIOU, M. & PAPAMICHAIL, I. 2014. Perimeter traffic control via remote feedback gating. *Procedia – Social and Behavioral Sciences*, Vol. 111, pp. 645-653.
- LECLERCQ, L., CHIABAUT, N. & TRINQUIER, B. 2014. Macroscopic fundamental diagrams: A cross-comparison of estimation methods. *Transportation Research Part B*, Vol. 62, pp. 1-12.
- MAHMASSANI, H., WILLIAMS, J. & HERMAN, R. Performance of urban traffic networks The 10th International Symposium on Transportation and Traffic Theory, 1987. 1-20.
- MAHMASSANI, H., WILLIAMS, J. C. & HERMAN, R. 1984. Investigation of network-level traffic flow relationships: some simulation results. *Transportation Research Record* 971, 121-130.
- MAHMASSANI, H. S., SABERI, M. & ZOCKAIE, A. 2013. Urban network gridlock: Theory, characteristics, and dynamics. *Transportation Research Part C: Emerging Technologies*, 36, 480-497.
- NAGLE, A. & GAYAH, V. 2014. The accuracy of network-wide traffic state esimations using mobile probe data. *forthcoming, Transportation Research Record.*
- OLSZEWSKI, P., FAN, H. & TAN, Y. 1995. Area-wide traffic speed-flow model for the Singapore CBD. *Transportation Research Part A*, 29, 273-281.
- RAKHA, H., PASUMARTHY, P. & ADJERID, S. The INTEGRATION framework for modeling longitudinal vehicle motion. TRANSTEC, Sept. 1-3 2004 Athens, Greece.
- RAKHA, H. & ZHANG, Y. 2004a. INTEGRATION 2.30 framework for modeling lane-changing behavior in weaving sections. *Transportation Research Record: Journal of the Transportation Research Board*, 1883, 140-149.
- RAKHA, H. & ZHANG, Y. H. 2004b. INTEGRATION 2.30 framework for modeling lane-changing behavior in weaving sections. *Traffic Flow Theory and Highway Capacity and Quality of Services 2004*, 140-149.
- RAKHA, H. A., AHN, K. & MORAN, K. 2012. INTEGRATION Framework for Modeling Ecorouting Strategies: Logic and Preliminary Results. *International Journal of Transportation Science and Technology*, 1, 259-274.
- SABERI, M., MAHMASSANI, H. S. & ZOCKAIE, A. 2014. Network capacity, traffic instability, and adaptive driving: findings from simulated urban network experiments. *EURO Journal of Transportation and Logistics*, 3, 289-308.

SIMONI, M. D., PEL, A. J., WARAICH, R. A. & HOOGENDOORN, S. P. 2015. Marginal cost congestion pricing based on the network fundamental diagram. *Transportation Research Part C: Emerging Technologies*, 56, 221-238.

SMEED, R. 1966. Road capacity of city centers. Traffic Engineering and Control, 9, 455-458.

- VAN AERDE, M., HELLINGA, B., BAKER, M. & RAKHA, H. 1996. INTEGRATION: An overview of traffic simulation features. *Transportation Research Records*.
- VAN AERDE, M. & RAKHA, H. 2013a. INTEGRATION © Release 2.40 for Windows: User's Guide Volume I: Fundamental Model Features. Blacksburg: M. Van Aerde & Assoc., Ltd.
- VAN AERDE, M. & RAKHA, H. 2013b. INTEGRATION © Release 2.40 for Windows: User's Guide Volume II: Advanced Model Features. Blacksburg: M. Van Aerde & Assoc., Ltd.
- ZAHAVI, Y. 1972. Traffic performance evaluation of road networks by the a-relationships, Parts 1 and 2. *Traffic Engineering and Control*, 14, 228–231, 292–293.
- ZHANG, L., GARONI, T. M. & DE GIER, J. 2013. A comparative study of Macroscopic Fundamental Diagrams of arterial road networks governed by adaptive traffic signal systems. *Transportation Research Part B: Methodological*, 49, 1-23.
- ZHENG, N., WARAICH, R. A., AXHAUSEN, K. W. & GEROLIMINIS, N. 2012. A dynamic cordon pricing scheme combining the Macroscopic Fundamental Diagram and an agent-based traffic model. *Transportation Research Part A: Policy and Practice*, 46, 1291-1303.

APPENDIX D

Comparing the use of link and probe data to inform perimeter metering control

This article may be cited as: Nagle, A.S. and Gayah, V.V. (2015) Comparing the use of link and probe data to inform perimeter metering control. *94th Annual Meeting of the Transportation Research Board*, 11-15 January, Washington DC.

INTRODUCTION

Daganzo (1) recently proposed a model of aggregate network behavior that was used to unveil the gridlock tendencies of congested traffic networks. Using this model, a control strategy was proposed to mitigate this gridlock behavior and maximize overall network efficiency. This strategy limits vehicle entries into a network when the density exceeds a critical value associated with congestion. This work also suggested the types of networks for which MFDs should arise: homogeneous networks with spatially uniform demand and congestion patterns. Under these conditions, the average flow and density in a network should be related by a well-defined curve known as the Macroscopic Fundamental Diagram (or MFD).

This relationship was then verified with empirical data in Geroliminis and Daganzo (2). Using a combination of data from loop detectors and GPS-enabled local taxis in Yokohama, Japan, an MFD linking the space-mean flow, density, and speed was obtained. Empirically derived MFDs were later obtained by others (*3; 4*). Thus, well-defined and reproducible MFDs may arise even when demand and congestion patterns are not perfectly uniform across a region of space (*5*). This is especially true if networks are redundantly designed and drivers are willing and able to alter routes in real-time to avoid locally congested areas (*6-10*). When MFDs exist, they can be used to inform network-wide control strategies to improve efficiency or as a means to examine the large-scale impacts of local improvement strategies. Recent examples include control strategies such as pricing (*11-13*), dynamic routing (*14*), adaptive signal control (*15; 16*), and perimeter metering (*1; 17-22*), also known as gating or perimeter boundary flow control.

The aforementioned control strategies all require the ability to estimate both the network's MFD and traffic conditions on the network in real-time. Unfortunately, this is not a trivial task. Typically, aggregate network properties are measured using fixed inductive loop detectors. However, fixed detectors tend to overestimate traffic densities in urban areas due to queue spillbacks near signalized intersections (*3*; *23*). Furthermore, not all links have embedded detectors so that traffic conditions are only measured at a subset of locations within the network. Thus, MFDs derived from reduced sets of detectors may be very inaccurate, especially if the links with detectors are selected randomly (*24*).

Keyvan-Ekbatani et al. (20) tested a perimeter flow control scheme informed solely by an MFD estimated from a reduced set of links that were equipped with detectors and traffic state estimations using these reduced detectors. These tests were indicative of how perimeter control would actually be implemented in a real city, since using a subset of the network's links greatly reduces data requirements and makes real-world implementation more feasible. Overall, the authors found that metering based on reduced MFDs provided substantial benefits in overall network efficiency. These reduced MFDs were created by carefully selecting detector locations based on prevailing traffic flows. However, if detector placement is random (e.g., the detectors are placed independently of traffic patterns or traffic patterns changes after the detectors are placed), the benefits when using reduced MFDs may not be as high. Thus, the use of randomly placed detectors as a method to inform MFD-based control deserves additional study.

More recent research has also shown that traffic state estimations and the MFD may be obtained accurately using data from mobile probe vehicles (*25-27*). If a known proportion of vehicles within an urban network report their odometer readings and travel times at specific time intervals, accurate estimates of average network flow and density can be calculated, along with estimates of their uncertainty. The fraction of probes can be reliably determined by combining probe and limited detector data. Accurate measurements were obtained when relatively few vehicles circulating in the network served as probes. Thus, the use of probe-based MFDs and state estimates based on reduced probe data also deserves further study as a method to inform perimeter control strategies.

In this paper, we compare the use of link-based and probe-based traffic state estimations to inform a simple perimeter control strategy using micro-simulations of an idealized network. The primary focus is to compare how the different types of data used to inform this perimeter control strategy impacts its efficiency. The detector method provides full information at a small number of locations distributed spatially across the network, while the probe method provides full trajectory information for a little number of vehicles that can cover the entire spatial extents of the network. The efficiency of each is measured by the delay savings obtained over the base-case when no metering is applied. Additionally, these results are compared to another scheme that accounts for the uncertainties that arise in traffic state estimations. The results suggest that both methods perform remarkably well even when limited data is available. Comparison of the methods also suggests that either method should provide adequate information to inform perimeter control strategies in practice.

The remainder of the paper is organized as follows. First, we describe the simulated network and the perimeter control strategy. Next we discuss the methods to estimate the network's traffic state. Then, we discuss the efficiency of the perimeter control strategy on an idealized simulated network. Finally, we provide some concluding remarks.

SCENARIO

This section describes the simulated network used in our tests and the perimeter control scheme to which probe-based and link-based traffic state estimations are applied.

Simulation network

In our simulations, we considered an idealized grid network simulated in the AIMSUN micro-simulation software. The network consisted of alternating one-way streets

arranged to form a 16x16 square grid. A 14x14 square subnetwork in the center of the larger 16x16 network was used to represent a typical city center within an urban area; see Figure 1a. This center was selected as the protected region around which the perimeter metering scheme was applied. The entire network consisted of a total of 544 links, each 400 feet long. Traffic detectors were placed in the middle of all links to represent fixed detectors: a total of 512 detectors were included and 420 of these were located within the subnetwork. Demands were calibrated to represent a typical morning rush period; see Figure 1b. Origins and destinations were placed at all intersections and at the upstream ends of all entry links into the network. All trips were equally likely to end at any of the destinations within the subnetwork. 20% of the trips had origins outside the subnetwork and the remaining 80% originated inside the subnetwork. Note that these latter trips were not subject to any metering control. Routing in the network was performed using the stochastic logit model implemented in AIMSUN in which drivers select the route that provides them with the smallest travel time. Furthermore, 25% of the drivers were assumed to be able to adaptively alter their routes during their trip in response to prevailing traffic conditions.



FIGURE 1 a) Network representation; b) Demand profile.

All intersections were signalized and operated with a 60-second cycle length, no offsets between adjacent signals and two-phase operation. Since one-way streets were used, no conflicting turning movements existed. Signal timings at intersections with entry links into the protected subnetwork were determined using the control logic detailed in the next section. Signal timings at the remaining intersections were fixed and operated with 26 seconds of green, 3 seconds yellow, and 1 second all red per phase.

Perimeter control strategy

The simple perimeter (gating/metering) control strategy employed in this work is illustrated in Figure 1. This strategy mimics the bang-bang strategy original proposed in Daganzo (1), which shows that the efficiency of a network is maximized when the

network is never allowed to enter the congested (decreasing) portion of the MFD. The strategy adopted here relies only on knowledge of the MFD and real-time estimates of

the average network density in the subnetwork. At the beginning of each time period, t,

the average density from the previous time period, t-1, was calculated. If the average

density exceeded a critical density threshold, $k_{
m c}$, the green time at every peripheral

intersection (i.e., the metered locations highlighted in Figure 1a) was reduced to 0 seconds to halt peripheral vehicle entries into the subnetwork. However, if average

density during the previous time period did not exceed k_{c} , the signal timings were reset

to the original values shared by all intersections. Note that all non-peripheral intersections always operated under fixed timings, so vehicles within the subnetwork were uninhibited by the perimeter control. More sophisticated strategies exist in the literature to select an appropriate metered entry rate; however, we adopt this simply strategy since our primary focus here is the data used to inform the strategy as opposed to the strategy itself.



FIGURE 2 Perimeter control strategy logic

TRAFFIC STATE ESTIMATIONS

This section describes the two methods used to calculate the network-wide metrics that describe the MFD and inform the perimeter control. We first define these metrics, then

discuss the estimation procedure used when only a subset of the requisite data is available, and finally discuss the uncertainty of the measurements.

Estimating the MFD using the generalized definitions

Here, we use the average flow and density to represent the MFD instead of other metrics, such as the total time spent (TTS) in the system and the total travel distance (TTD) allowed. The reason for this choice is that TTS and TTD could vary based on how many links or probe vehicles are sampled in the estimation process, whereas flow and density are normalized versions that are consistent across different probe and link percentages.

The average flow, q [veh/hr], and average density, k [veh/mi], within the network were defined using the generalized definitions of Edie. For a given analysis period, the metrics are calculated using the following formulas: $q = d_T/LT$ and $k = t_T/LT$, where

 d_T [veh-mi] is the total distance traveled by all vehicles in the network, t_T [veh-hr] is

the total time spent by all vehicles in the network, *L* [mi] is the total length of streets,

and T [hr] is the length of the analysis interval. However, the true values of d_T and t_T

can only be obtained if the trajectories of all vehicles traveling within the network are known. Since these data are generally unavailable, alternate methods to estimate the average network flow and density must be used to gain an understanding of the network's traffic state.

Two traffic state estimation methods are considered here that have illustrated their usefulness to estimate a network's traffic state in real-time using limited or reduced data. The first method is the current state of practice, which uses loop detector measurements placed on links throughout the network to provide an estimate of the network's average flow and density. The second strategy uses data from mobile probe vehicles.

Estimating traffic states using link data

The MFD of a network can be derived using spot measurements from loop detectors that are placed on links throughout the network. In this case, the average flow and density are calculated from the loop detector measurements using the following equations:

$$\hat{q} = \sum_{i} q_{i} T L_{i} / LT, \text{ and}$$

$$\hat{k} = \sum_{i} [o_{i} N_{i} L_{i} T / 100 \lambda] / LT,$$
(2)

where q_i [veh/hr] is the measured loop detector flow on link *i*, L_i [mi] is the length of

link i, o_i [%] is the measured time-occupancy on link i, N_i [lanes] is the number of lanes

on link *i*, and *l* [mi] is the average vehicle length. This method assumes that traffic

states measured at each detector accurately reflects conditions across that entire link. Previous work has shown that detectors should be placed near the middle of the link to achieve this result (*19*; *28*).

For the MFD or real-time traffic state estimations to accurately represent traffic conditions of the entire network, a loop detector is required on every link. However, Equations 1 and 2 can also be used to estimate the MFD if detectors are only available for a fraction of the links within the network. If both detectors and vehicles are uniformly distributed across the network, the reduced MFDs should describe aggregate traffic conditions well. In this work, we will explore informing the perimeter control strategy using full link estimations, where 100% of the links in the subnetwork are equipped with detectors, and reduced link estimations, where only a subset of the links in the subnetwork are equipped with detectors.

Estimating traffic states using probe data

The generalized definitions can be used to estimate the MFD and aggregate traffic measures if all vehicles serve as mobile probes and provide detailed trajectory data. However, if just a fraction of vehicles serve as probes, only a subset of trajectory information would be available for traffic state estimations. Recent work has found that trajectory data from a subset of vehicles serving as probes can be combined with counts at fixed detector locations to estimate average flows and densities in a network quite accurately (*25*). In this case, estimates of average network flow and density in the network are calculated as follows:

$$\hat{q} = d_p / \rho LT = N_p \bar{d}_p / \rho LT$$
, and (3)

$$\hat{k} = t_p / \rho LT = N_p \bar{t}_p / \rho LT, \tag{4}$$

where d_p [mi] is the total distance traveled by the probe vehicles in the network, t_p [hr]

is the total time spent by the probe vehicles in the network, \vec{d}_{p} [mi] is the average

distance traveled by each probe vehicle in the network, \bar{t}_{p} [hr] is the average time spent

by each probe vehicle in the network, N_p [veh] is the number of probe vehicles that traveled in the network, and p is the mobile probe penetration rate during a specified time interval. The quantities \bar{d}_p , \bar{t}_p , and N_p are directly measureable in real-time from the trajectories of the probe vehicles, and p can be estimated by combining the

trajectory data with counts of all vehicles at detectors, as in (27).

Estimation of uncertainty

The network-wide traffic estimates obtained from a subset of link or probe data may not reflect actual traffic conditions. Thus, we refer to these estimates as exhibiting some amount of uncertainty. Here, we describe methods to estimate the magnitude of this uncertainty for both the link and probe methods. We focus only on density here since the perimeter control strategy relies only on density measurements.

Link method

The micro-simulation network was used to quantify uncertainty in density estimates predicted from the link measurements. The density was estimated using the link method by randomly sampling various proportions of links in the network many times during several simulation periods for which the actual density using all the link data was known. The variance of the density estimates obtained from the estimation procedure is plotted in Figure 3 as a function of the actual density in the network for all reduced link percentages considered.



FIGURE 3 Variance of density as a function of actual density for the link-based estimations

Figure 3 confirms that the uncertainty of the density estimates increases as fewer links are sampled; i.e., as the fraction of links used to estimate density decreases, the variance of the estimates increases. Furthermore, the uncertainty increases with density up until some point (at around k = 175 veh/mi) and then begins to fall again. This suggests that estimates become more inaccurate as the network becomes congested, but when the network is extremely congested the estimates start becoming more accurate. This is consistent with previous work that shows traffic networks become more unstable and unpredictable as they become congested. Polynomial curves were fit to the data for each reduced link percentage to estimate the variance of the link-based density estimations:

$$va \ (\hat{k}_{75\%}) = -0.0006k^2 + 0.2k \tag{5}$$

$$var(\hat{k}_{50\%}) = -0.0012k^2 + 0.4723k \tag{6}$$

$$var(\hat{k}_{40\%}) = -0.002k^2 + 0.7633k \tag{7}$$

$$var(\hat{k}_{30\%}) = -0.0038k^2 + 1.3453k$$
 (8)

$$var(\hat{k}_{25\%}) = -0.0046k^2 + 1.5999k \tag{9}$$

$$var(\hat{k}_{20\%}) = -0.0048k^2 + 2.0276k \tag{10}$$

$$var(\hat{k}_{15\%}) = -0.0066k^2 + 2.7832k$$
 (11)

$$var(\hat{k}_{10\%}) = -0.0163k^2 + 5.831k$$
 (12)

$$var(\hat{k}_{5\%}) = -0.0302k^2 + 10.901k \tag{13}$$

These equations can be used during simulation runs by substituting k with its estimate from the link data.

Probe method

The accuracy of the probe-based estimates for average density was explored previously in Nagle and Gayah (25). The uncertainty of these estimates was shown to be well-described by the following equation:

$$v\hat{a}r(\hat{k}) = k^2 var(t_i)(1-\rho)^2 / N_p^2 \bar{t}_p^2 + kvar(t_i)(1-\rho) / \bar{t}_p \rho LT + k^2 (1-\rho) / N_p,$$
(14)

where t_i [hr] is the time spent in the network by each probe vehicle. Again, k was substituted with its estimate from the probe data to calculate the variance in real-time as the simulation ran.

THE EFFICIENCY OF METERING AT THE PERIMETER

The idealized grid network was used to test the following scenarios: no control; perimeter control informed by full (100%) and reduced (75%, 50%, 40%, 30%, 25%, 20%, 15%, 10% and 5%) link data; and, perimeter control informed by full (100%) and limited (75%, 50%, 40%, 30%, 20%, 10%, 7.5%, 5%, 2.5% and 1%) probe data. About 30 simulation instances were run for each metering scenario, and the average delay savings per vehicle for the control scenarios compared to the no control scenario was used as the primary measure of effectiveness.

No control results

First, simulation instances of the no control scenario were run to serve as the baseline of comparison of the control strategy under the various scenarios. In this case, the signal timings for all intersections within the entire network operated with the same fixed timings for the entire simulation. A subset of the simulation instances were randomly selected to derive the network's MFD using link-based and probe-based methods assuming all data were available. These MFDs are presented in Figure 4. Both the loading (left-hand side) and recovery (right-hand side) periods are illustrated; notice that the recovery period exhibits a clear clockwise hysteresis loop that suggests poor performance as the network recovers from congestion. Overall, these MFDs imply that gating or metering control would be beneficial since it would help avoid congested states (the decreasing branch of the MFD) and the performance drop during congestion recovery.



FIGURE 4 Non-metered MFDs showing loading and recovery processes for: a) link-based estimations and b) probe-based estimations

Comparison of the Figures 4a and 4b reveals that densities using the link method are consistently higher than those using the probes. This is most likely due to queues from the signalized intersections spilling over to the detectors placed in the middle of each link, which causes the link method to artificially inflate the density estimates. However, the average flows are generally the same using both estimation procedures, as would be

expected. A critical density of $k_e = 50$ veh/mi was selected to implemented perimeter

flow control based on the probe-based MFD. The corresponding critical density using

the link-based method was found to be k_{e} = 58 veh/mi. These values (shown by vertical

arrows in Figure 4) represented the density at which increasing density resulted in lower average flows and less predictable overall network behavior. The perimeter flow control strategy aims to prevent the subnetwork from becoming congested by keeping its density below the critical threshold. Note that since the critical values were chosen to correspond to the exact same actual densities, the performance of the network should be the same when full link and probe data is used.

Perimeter control without uncertainty

This section presents the results for when reduced link- and limited probe-based traffic state estimates are used directly to inform the control scheme without explicitly considering the uncertainty exists for these estimates. For the reduced link-based estimations, a subset of detectors equal to the fraction of links considered in the estimation procedure were randomly selected at the start of each simulation instance and used to estimate the average subnetwork density. For the limited probe-based estimations, vehicles that entered the network were randomly designated as a probe

vehicle with a fixed probability, *p*, equal to the limited probe percentage considered.

However, the actual probe penetration rate varied across analysis periods and was estimated in the simulation using the 'virtual trip line' method, which combines probe data with existing loop detectors and was shown to be very accurate (*25*).

Figure 5 presents the average delay savings per vehicle vs. the amount of probe or link data used to inform the metering. The blue dots in Figure 5 represent the delay savings of a single simulation instance, the red squares represent the mean delay savings across all of the simulation instances for each percentage, the blue plus signs represent the 95% confidence interval of the mean, and the red plus signs represent the 95% confidence interval of the standard deviation. The mean values are connected by a solid black line to present the trends observed. The results of the simulations are also summarized in Table 1. The minimum, maximum and standard deviation of the delay savings is also provided to show how much these values can vary across individual runs. Notice that the delay savings are about the same across the two methods when 100% of the data is used, as expected. Overall, the delay savings under perfect information is about 49 seconds/vehicle.



FIGURE 5 Travel time savings without uncertainty for a) link-based estimations and b) probebased estimations

The results of the link-based method suggest that the average delay savings are remarkably consistent when subsets of links are selected randomly to inform the perimeter control. Even when just 5% of the links are selected to inform the control, the expected delay savings is statistically equal to what is expected when the full detector information (100% of detectors) is used. In fact, the expected savings is statistically equivalent for all link percentages considered (verified by the relevant statistical tests). Furthermore, the variability across simulation instances is more or less consistent across the link fractions considered. Overall, this confirms previous work that studied reduced MFDs to inform perimeter control, which found similar results (20). However, this previous study suggested that the subset of links used to create the reduced MFD must be selected carefully based on prevailing traffic patterns to be sure that it represents a true MFD of the network. The simulations here show that randomly sampling links in a network might be adequate if the demand pattern is relatively uniform. Since fairly uniform demand patterns are required for well-defined MFDs, these results are especially promising for link-based MFD control.

Scenario		Delay Savings						
		(sec/veh)						
		Mean	Min.	Max.	SD	Green Time Resets		
Link	100% Sampled	48.30	0.81	148.31	39.20	4.375		
	75% Sampled	47.18	-0.54	148.31	41.35	4.325		
	50% Sampled	47.59	-6.36	148.50	41.88	4.55		
	40% Sampled	42.35	-5.10	131.34	35.90	4.6		
	30% Sampled	46.46	0.50	144.76	40.32	4.35		
	25% Sampled	45.33	-7.85	128.73	38.91	4.675		
	20% Sampled	48.18	-4.00	149.67	42.60	4.825		
	15% Sampled	40.93	-11.31	154.31	38.83	4.775		
	10% Sampled	45.14	-52.35	164.93	43.09	4.8		
	5% Sampled	49.01	-0.05	130.13	36.60	4.85		
	100% Sampled	49.88	0.06	149.05	40.44	4.45		
Probe	75% Sampled	43.12	-8.10	141.00	40.64	4.425		
	50% Sampled	43.51	-10.53	137.85	40.54	4.525		
	40% Sampled	40.47	-12.27	125.03	38.74	4.575		
	30% Sampled	41.82	-17.66	130.15	41.29	4.55		
	20% Sampled	42.01	-12.38	139.89	41.36	4.65		
	10% Sampled	40.62	-13.37	131.90	40.66	4.875		
	7.5% Sampled	38.73	-25.75	131.85	41.34	4.875		
	5% Sampled	36.63	-12.39	149.78	39.25	5.4		
	2.5% Sampled	40.48	-20.69	157.26	48.60	5.475		
	1% Sampled	42.96	-43.80	166.59	53.95	6.8		

TABLE 1 Travel time savings summary without estimation uncertainty

The results of the probe-based method show very similar results to the link-based method. Primarily, the expected delay savings are statistically equal across all probe fractions considered. Thus, the probe-based method works as well when very few vehicles serve as mobile probes (as little as 1%) as when many vehicles serve as mobile probes (100%). We considered very high demands in which an average of 44,000 simulated vehicles used the network during the 4-hour simulation period. Therefore, the 1% probe vehicle case still uses information from about 440 probe vehicles, which is quite a large number. The variability of the average delay savings across individual outcomes is quite large when a small number of probe vehicles are used to inform the metering strategy. However, this variability appears to stabilize once about 7.5% of the vehicles serve as mobile probes.

It should be noted that both Table 1 and Figure 5 show negative delay savings for some simulation instances. Such negative values would occur if the control is triggered too early and thus is too restrictive (e.g., the estimated density is greater than the critical density so the control is triggered when the actual density is well below the critical threshold). This is a likely outcome if uncertain data is used. The minimum values experienced provide an indication of the worst-case outcome when applying control under imperfect information. Examination of Table 1 suggests that the magnitude of the worst-case outcomes decrease in severity as more information is used to inform metering (i.e., higher probe and link percentages result in minimum expected delay savings that are more positive). The minimum value for 5% of links sampled seems to go against this trend, but we suspect that this is more likely due to stochastic fluctuations in the simulations.

Perimeter control accounting for uncertainty

In this section, the uncertainty of density measurements was incorporated into the perimeter control scheme. For these tests, the gating/metering strategy was modified and triggered when there was a significant probability that the true density associated

with the estimated density exceeded k_{e} ; specifically, metering was triggered if the sum

of the estimated density and one standard deviation of the error of the estimate was

greater than k_{e} . Many simulation instances were run for both the link- and probe-based

metering strategies, and the travel time savings are illustrated in Figure 5 and summarized in Table 2. Note that the results for 100% probe or link data are the same as in Table 1 because when all probe or link data is used there is no uncertainty in the estimates.



FIGURE 6 Travel time savings with uncertainty for a) link-based estimations and b) probe-based estimations

Overall, the same general trends emerge when uncertainty is accounted for in the metering process. For both probe-based and link-based control, the average delay savings is remarkably consistent across all percentages of probe vehicles and links sampled, respectively. The link-based method provides delay savings with the same variability across simulation iterations for all link percentages sampled, while the probe-based method exhibits decreasing variability as probe percentages increase from 1% to 7.5%, after which the variability is more or less constant for the probe percentages sampled.

Interestingly, the control accounting for uncertainty provides slightly higher delays savings (on average) than the control that does not account for uncertainty. This occurs for both the probe-based and link-based strategies. The magnitude of the additional benefits observed when accounting for uncertainty generally increases as the probe or link fractions decreases. This makes sense as the level of uncertainty in the estimates increases as less information is used to inform the control. It should be noted that the differences in average delay savings between the uncertainty and no uncertainty case are not statistically significant. Still, this suggests that a perimeter flow control scheme

that explicitly accounts for uncertainty might be more beneficial and should be studied further.

To further explore the differences between including uncertainty in the estimations, the number of green time resets was compared to understand the erraticism of the two methods. A green time reset occurs whenever the density of the subnetwork falls below the critical density threshold following a cycle in which the density exceeded the critical threshold. The final column of Tables 1 and 2 illustrate the number of green time resets for the probe-based strategy with and without uncertainty. There is a clear trend in both cases that the number of resets decreases as more probe vehicles or links/detectors are used to inform the control. Thus, as more information is used, the metering scheme becomes less erratic. Comparing with and without uncertainty, there are fewer green time resets when uncertainty is not included in the metering process. This occurs because if uncertainty is not included, the metering consistently occurs at higher densities. Since the green times mostly change during small periods of time at the onset and recovery of congestion, the density may increase into the critical range and not bounce back out of the critical range, causing fewer green time resets.

Scenario		Delay savings					
		(sec/veh)					
		Mean	Min.	Max.	SD	Green Time Resets	
	100% Sampled	48.30	0.81	148.31	39.20	4.375	
	75% Sampled	50.10	7.06	135.81	39.56	4.625	
	50% Sampled	50.11	0.49	144.66	39.93	4.75	
	40% Sampled	49.54	-12.43	153.72	41.15	5.1	
Link	30% Sampled	49.93	3.14	143.89	40.89	4.925	
	25% Sampled	53.62	-11.58	159.17	41.51	4.925	
	20% Sampled	51.60	-1.82	140.86	40.64	4.925	
	15% Sampled	51.41	-4.24	157.29	40.38	5.375	
	10% Sampled	50.12	-14.90	155.43	42.76	6.675	
	5% Sampled	53.48	-8.33	155.17	42.74	8.125	
Probe	100% Sampled	49.88	0.06	149.05	40.44	4.45	
	75% Sampled	44.66	-10.17	141.00	40.45	4.5	
	50% Sampled	44.27	-17.01	130.84	40.03	4.425	
	40% Sampled	44.91	-16.07	136.41	39.54	4.625	
	30% Sampled	43.11	-14.35	136.88	39.77	4.625	
	20% Sampled	44.98	-13.14	129.95	38.34	4.975	
	10% Sampled	41.91	-18.37	126.19	40.15	5.2	
	7.5% Sampled	41.63	-19.75	133.91	40.21	5.1	
	5% Sampled	46.17	-7.43	138.02	41.84	5.45	
	2.5% Sampled	47.36	-15.21	140.70	44.79	5.9	
	1% Sampled	52.98	-52.57	170.05	62.38	11.275	

TABLE 2 Delay savings summary with estimation uncertainty

CONCLUDING REMARKS

A network-wide perimeter flow control strategy was tested using traffic state estimations based on point measurements from loop detectors and mobile measurements from probe vehicles. The control strategy operated under the premise that as urban networks become congested, restricting vehicle entry into the network improves the overall performance of the network by delaying the onset of oversaturation. An idealized grid network with a congested subnetwork was considered to represent a typical peak period of an urban downtown. When the subnetwork reached a pre-defined critical density threshold, an adaptive signal control strategy at its periphery was applied to restrict vehicle entry.

The link- and probe-based metering strategies were each compared to the non-metered scenario to quantify the effects of metering based on different traffic state estimations. Additionally, estimation uncertainty was included in the metering scheme to determine how accounting for estimation uncertainty affects the delay savings. The results showed that the delay savings achieved from both link-based and probe-based information is consistent. In general, both strategies seem to provide adequate information to inform the perimeter control. Furthermore, the results suggest that reduced information from limited detector or probe information is also remarkably consistent for the fractions of probe vehicles or links/detectors sampled in this work. Thus, it appears that randomly sampled links/detectors or probe vehicles provide adequate information to inform perimeter metering control on urban network with uniform demand patterns. The results also suggest that accounting for uncertainty in state network state estimations would provide higher benefits when implementing a perimeter control strategy.

This work assumed uniform demand patterns and homogeneous coverage of probe vehicles between OD pairs. Additional work should be performed to expand this work to consider how heterogeneous probe distributions may impact the metering scheme. For example, dynamic partitioning schemes might be one way to address this concern by identifying regions of the network in which probe penetration rate is approximately uniform and combining estimates from each of these regions. Still, probe- and linkbased estimation methods are shown to be very promising for informing perimeter metering during peak periods. These state estimation procedures may also be used to inform other network-wide control strategies that rely on real-time traffic state estimations and the MFD, such as adaptive signal control, dynamic vehicle routing, and pricing.

ACKNOWLEDGMENTS

This research was supported by the Mid-Atlantic Universities Transportation Center.

REFERENCES

[1] Daganzo, C. F. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, Vol. 41, No. 1, 2007, pp. 49-62.

[2] Geroliminis, N., and C. F. Daganzo. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, Vol. 42, No. 9, 2008, pp. 759-770.

[3] Buisson, C., and C. Ladier. Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2124, No. 1, 2009, pp. 127-136.

[4] Tsubota, T., A. Bhaskar, and E. Chung. Brisbane macroscopic fundamental diagram: empirical findings on network partitioning and incident detection. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. in press, 2014.

[5] Geroliminis, N., and J. Sun. Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Research Part B: Methodological*, Vol. 45, No. 3, 2011, pp. 605-617.

[6] Daganzo, C. F., V. V. Gayah, and E. J. Gonzales. Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. *Transportation Research Part B: Methodological*, Vol. 45, No. 1, 2011, pp. 278-288.

[7] Gayah, V. V., and C. F. Daganzo. Clockwise hysteresis loops in the macroscopic fundamental diagram: an effect of network instability. *Transportation Research Part B: Methodological*, Vol. 45, No. 4, 2011, pp. 643-655.

[8] Mahmassani, H. S., M. Saberi, and A. Zockaie. Urban network gridlock: Theory, characteristics, and dynamics. *Transportation Research Part C: Emerging Technologies*, Vol. 36, 2013, pp. 480-497.
[9] Saberi, M., H. S. Mahmassani, and T. C. Director. Empirical characterization and interpretation of hysteresis and capacity drop phenomena in freeway networks. *Transportation Research Record: Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, DC*, 2013.

[10] Saberi, M., A. Zockaie, and H. Mahmassani. Network capacity, traffic instability, and adaptive driving: Findings from simulated network experiments. In *The 1st European Symposium on Quantitative Methods in Transportation Systems, Lausanne, Switzerland*, 2012.

[11] Geroliminis, N., and D. M. Levinson. Cordon pricing consistent with the physics of overcrowding.In *Transportation and Traffic Theory 2009: Golden Jubilee*, Springer, 2009. pp. 219-240.

[12] Gonzales, E. J., and C. F. Daganzo. Morning commute with competing modes and distributed demand: User equilibrium, system optimum, and pricing. *Transportation Research Part B: Methodological,* Vol. 46, No. 10, 2012, pp. 1519-1534.

[13] Zheng, N., R. A. Waraich, K. W. Axhausen, and N. Geroliminis. A dynamic cordon pricing scheme combining the Macroscopic Fundamental Diagram and an agent-based traffic model. *Transportation Research Part A: Policy and Practice*, Vol. 46, No. 8, 2012, pp. 1291-1303.

[14] Knoop, V., S. Hoogendoorn, and J. W. Van Lint. Routing strategies based on macroscopic fundamental diagram. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2315, No. 1, 2012, pp. 1-10.

[15] Gayah, V. V., X. Gao, and A. S. Nagle. On the impacts of locally adaptive signal control on urban network stability and the Macroscopic Fundamental Diagram. *Transportation Research Part B: Methodological*, Vol. 70, No. 1, 2014, pp. 255-268.

[16] Zhang, L., T. M. Garoni, and J. de Gier. A comparative study of Macroscopic Fundamental Diagrams of arterial road networks governed by adaptive traffic signal systems. *Transportation Research Part B: Methodological*, Vol. 49, 2013, pp. 1-23.

[17] Keyvan-Ekbatani, M., M. Papageorgiou, and I. Papamichail. Perimeter Traffic Control via Remote Feedback Gating.In *Procedia-Social and Behavioral Sciences, No. 111*, 2014. pp. 645-653.

[18] Haddad, J., and N. Geroliminis. On the stability of traffic perimeter control in two-region urban cities. *Transportation Research Part B: Methodological*, Vol. 46, No. 9, 2012, pp. 1159-1176.

[19] Keyvan-Ekbatani, M., A. Kouvelas, I. Papamichail, and M. Papageorgiou. Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B: Methodological,* Vol. 46, No. 10, 2012, pp. 1393-1403.

[20] Keyvan-Ekbatani, M., M. Papageorgiou, and I. Papamichail. Urban congestion gating control based on reduced operational network fundamental diagrams. *Transportation Research Part C: Emerging Technologies*, Vol. 33, 2013, pp. 74-87.

[21] Aboudolas, K., and N. Geroliminis. Perimeter and boundary flow control in multi-reservoir heterogeneous networks. *Transportation Research Part B: Methodological*, Vol. 55, 2013, pp. 265-281.

[22] Keyvan-Ekbatani, M., M. Yildirimoglu, N. Geroliminis, and M. Papageorgiou. Traffic Signal Perimeter Control with Multiple Boundaries for Large Urban Networks. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. 2013.

[23] Courbon, T., and L. Leclercq. Cross-comparison of macroscopic fundamental diagram estimation methods. *Procedia-Social and Behavioral Sciences*, Vol. 20, 2011, pp. 417-426.

[24] Ortigosa, J., M. Menendez, and H. Tapia. Study on the number and location of measurement points for an MFD perimeter control scheme: a case study of Zurich. *EURO Journal on Transportation and Logistics*, 2013, pp. 1-22.

[25] Gayah, V. V., and V. V. Dixit. Using mobile probe data and the macroscopic fundamental diagram to estimate network densities: Tests using microsimulation. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2390, No. 1, 2013, pp. 76-86.

[26] Leclercq, L., N. Chiabaut, and B. Trinquier. Macroscopic Fundamental Diagrams: A crosscomparison of estimation methods. *Transportation Research Part B: Methodological*, Vol. 62, 2014, pp. 1-12.

[27] Nagle, A. S., and V. V. Gayah. The Accuracy of Network-wide Traffic State Estimations Using Mobile Probe Data. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2421, 2013, pp. 1-11.

[28] Papageorgiou, M., and G. Vigos. Relating time-occupancy measurements to space-occupancy and link vehicle-count. *Transportation Research Part C: Emerging Technologies*, Vol. 16, No. 1, 2008, pp. 1-17.

APPENDIX E

Design and evaluation of network control strategies using the Macroscopic Fundamental Diagram

This article may be cited as: Du, J., Rakha, H. and Gayah, V.V. (2015) Design and evaluation of network control strategies using the Macroscopic Fundamental Diagram. *18th Annual IEEE Conference on Intelligent Transportation Systems*, 15-18 September, Canary Islands, Spain.

INTRODUCTION

Efficient management and control of urban transport system is a challenge due to the complexity of the transportation network where travelers have multiple alternative route choices. Recently researchers verified (with simulation and empirical data) that a well-defined relationship exists between the average flow and density measured across an urban network [1-5]. This relationship, known as the Network or Macroscopic Fundamental Diagram (NFD or MFD) is very helpful for researchers and traffic management agencies to monitor the status of a traffic network, design efficient traffic control strategies, and measure the effectiveness of network efficiency improvement strategies [3].

Previous research has examined the different attributes and applications of the MFD. For example, Geroliminis and Daganzo [4] used real data in Yokohama to demonstrate the existence of an MFD linking space-mean flow, density, and speed. Geroliminis and Sun [6] studied the spatial variability of vehicle density and found that it affects the shape, the scatter, and the existence of an MFD. Buisson and Ladier [7] found that heterogeneity, such as differences between the surface and highway network and distances between the loop detector and traffic signal, has a strong impact on the shape of the MFD. Geroliminis and Levinson researched various strategies, including road pricing and allocating vehicles that enter the network, that can utilize MFDs in optimizing network controls [8]. Keyvan-Ekbatani et al. tested the application of gating measures to improve mobility in saturated traffic conditions by using the MFD to derive clear gating targets to maximize throughput in the protected network part [9-11]. Ramezani et al. [12] used a hierarchical perimeter flow controller to minimize the network delay and distribute the congestion more homogeneously such that the scattering of the MFD is lowered. Haddad et al. [13] formulated an asymmetric cell transmission model for a large-scale mixed traffic network where they partitioned the urban network into two regions, each has a well-defined MFD. They found that centralized control is more effective comparing to a simple freeway ramp metering with urban MPC controller. Haddad and Geroliminis [14] partitioned the network in two regions and used a state-feedback control strategy that maximizes the number of vehicles that complete their trips given the assumption that MFD for both regions are known.

Although it is convenient to use an MFD to describe the traffic status across a network and design traffic control strategies, the data needed to plot the MFD are not always readily available. The individual travel distance and travel time for every single vehicle in a network is generally impossible to acquire. With the rapid development of GPS technology, now it is relatively simple to track vehicles and record their kinematic data, and thus obtaining trajectory data from at least a portion of the vehicles traveling in the network is now feasible. Recent research has shown that the trajectories of a certain subset of vehicles traveling in the network can be used to satisfactorily estimate the overall MFD in a way that is both reliable and accurate when probe vehicles are more or less uniformly distributed across a network [15]. Unfortunately, the assumption of
uniform probe vehicle distributions in previous research might be not realistic in practice. GPS technologies are likely to be included in newer vehicles and certain origin and/or destinations are thus more likely to have higher penetration rates of these types of vehicles than others. Methods to estimate the MFD are needed that account for this heterogeneous distribution of probe vehicles across a network to estimate more reliable MFDs. Only with the reliable methodology of estimating MFDs using a limited amount of probe data, can using MFDs as a tool to monitor and control network congestion a feasible method.

Previous research explored the usage of MFD to examine the impacts of congestion improvement strategies. Zheng et al. [16] proposed a cordon-based pricing scheme in the city of Zurich. The MFD is used to determine the optimal tolls. The results show that tolls determined by MFD is effective in mitigating congestion by decreasing the travel time both inside and outside of the cordon. Knoop et al. [17] used simulation to test how four routing scenarios affect the shape of the MFD: fixed routing; speed-based routing; subnetwork speed-based routing; and subnetwork accumulation-based routing subdivided into four types. The average flow increasing can be as high as 46%. Zhang et al. [18] studied the MFD of arterial road networks governed by different types of adaptive traffic signal control under various boundary conditions. They found that the shape of MFD depends on the signal system and level of heterogeneity. Keyvan-Ekbatani et al. [19] showed that the reduced Network Fundamental Diagram (NFD) exhibits a critical range of traffic states. Gating controls may be based on a reduced amount of real-time measurements and a gating control can drastically improve the reliability of the traffic conditions.

In this study, we first test the feasibility of using limited probe information to estimate MFDs. Simulation results using INTEGRATION [20-27] demonstrate that using weighted average probe penetration rates have an obvious advantage over the arithmetic average probe penetration rate in estimating MFDs from a subset of probe vehicle information. We further demonstrate the usage of MFD as a monitoring and control tool to regulate the performance of transport network. MFDs for a network under several different control strategies, including signal regulation and re-routing that aim at alleviating congestion, were built and illustrated how the effectiveness of such strategies impact on the shape of the MFDs. The paper is organized as below: Following the introduction part is the description of methodology. After that, the results of the estimating MFD using a limited amount of probe data are discussed. Then, the varied magnitude and shape of MFD under different control strategies are compared. The last part of the paper is the conclusion and discussion.

BACKGROUND AND METHODOLOGY

Constructing MFD Using Probe Data

Given known trajectory data of each vehicle in the network, the network-wide average density and flow can be estimated using Equations (1) and (2) per Edie [28].

$$k = \frac{\sum_{1}^{l} (t_{i})}{L_{n} * T} \qquad (1)$$
$$q = \frac{\sum_{1}^{l} d_{i}}{L_{n} * T} \qquad (2)$$

Here **q** is the network flow rate (vehicles/hour); **k** is the network density (vehicles/mile) for one analysis period; **I** is the total number of trips recorded in that analysis period (e.g. 15 minutes); **t**_i and **d**_i are the travel time (seconds) and distance (miles), respectively, for trip **i**; **L**_n and **T** are the network length (lane-miles) and analysis period length (seconds), respectively.

Since it is not always typical to acquire such detailed data in a real-world network, [29] proposed a model using a certain percentage of probe data to estimate the overall MFD in a network based on Equations (1) and (2) to generate Equations (3) and (4).

$$\hat{k} = \frac{\sum_{1}^{p_{I}}(t_{p_{I}})}{\rho L_{n} * T}$$
(3)
$$\hat{q} = \frac{\sum_{1}^{p_{I}} d_{p_{I}}}{\rho L_{n} * T}$$
(4)

Here t_{pi} and d_{pi} are the travel time and distance for trip i made by the probe vehicles; ρ is

the average probe vehicle market penetration rate; and I_n and T are the network length and analysis period duration, respectively.

Although it is convenient to assume one uniform penetration rate for the entire network, in reality the probe penetration rate is not a fixed number. To overcome this limitation, we propose a method for estimating the probe market penetration rates using weighted average involving travel time and travel distance by each trip.

Equation (5) and (6) are two methods that both seek to address the issue of nonuniform probe penetration rates by estimating a weighted average probe penetration rate to describe the overall network penetration rate. We start by assuming that the

probe penetration rate for vehicles traveling between origin i and destination j, ρ_{ij} , is

fixed for some analysis period and is known a priori. The average travel distances and travel times by the probe vehicles traveling between the various OD pairs are

incorporated in the computation of the weighted average penetration rate. The first method uses individual vehicle travel times as the weighting metric, while the other uses individual vehicle travel distances.

$$\widehat{\rho_{t}} = \frac{\sum \rho_{ij} * \widehat{t_{ij}}}{\sum t_{ij}} \quad \forall \text{ origin } i \text{ to destination } j \quad (5)$$

$$\widehat{\rho_{d}} = \frac{\sum \rho_{ij} * \widehat{d_{ij}}}{\sum t_{ij}} \quad \forall \text{ origin } i \text{ to destination } j \quad (6)$$

Here f_{ij} and d_{ij} are the average travel time and distance from origin i to destination j,

respectively. Each of these weighted average probe penetration rates can then be substituted into Equations 3 and 4 to estimate the average density and flow in the network, receptively. Equations 5 and 6 imply that the longer a trip occupies the network, caused either by a longer travel distance or travel time, the more meaningful of a contribution it makes to the estimate of network density and flow.

Monitor and Control Network Congestion

Using MFD as a tool, the transportation managing agency can not only monitor the traffic condition of the network, but also adopt effective control strategies to avoid over congestion and distribute the congestion in certain areas more evenly across the network. The magnitude and shape of the MFD can illustrate the traffic condition of a network. Control strategies can be adopted to regulate the vehicles in the network by changing their speeds, routes, and entering and exiting rate to prevent the average network flow and density from exceeding the capacity of a network, which can be observed from the magnitude and shape of the MFDs. To understand how the control strategies affect the MFDs, the following strategies are tested in this paper:

- 1. Perimeter controls, where the green times of the signals is decreased for the entering traffic and increased for the existing traffic at the boundary of a congested region;
- 2. Adaptive signal controls, where the signal timings are updated and optimized every 300 seconds; and
- 3. Re-routing, where the vehicles in the network re-calculate their routes every 300 seconds.

The MFDs for the network under these control strategies are plotted against the base scenario and the effectiveness of these strategies are compared in terms of cumulative vehicle arrival rates, delays, travel times, and fuel consumption levels. The results are illustrated in the sections below.

MFD CONSTRUCTION AND CONTROL STRATEGIES

Only if MFD can be constructed with readily available data, can one uses it as a controlling and managing tool. Therefore, in this section, we first testify an algorithm that can estimate MFD using only a limited subset of probe data. After proving the feasibility of accurately estimating MFD, we further explored the effectiveness of congestion control strategies by plotting the MFDs under such strategies that are applied to the network at the onset of congestions. By comparing the magnitude and shape of the MFDs, the effectiveness of such strategies are compared.

An idealized 16×16 square grid network with alternating one-way streets is used for this study. In total, there are 544 links, each measuring 0.3 KM and with 2 lanes for each direction, and 64 zones (origin and destinations). The zones are at the borders of the network. A uniform demand is created for each origin-destination (OD) pair. During the 3-hour simulation, the first hour is a building-up (or loading) period with a rate of 6 vehicles/hour per OD pair. The second hour has an average trip rate of 17 trips, representing the congested time period. The third hour has an average trip rate back to 6 trips, the same as during the first hour, and represents the recovery period from congestion. Two scenarios were considered in which the penetration rates of mobile probe vehicles differed based on the origin-destination pair. The goal of these scenarios was to test the combination effects of high versus low penetration rates between different OD pairs. The scenarios include (as shown in Figure 1): (a) ODs from the gray quadrant of the grids to the black quadrant of the grids have a penetration rate of 0.8. All the other ODs have a penetration rate of 0.1.



Figure 8. Idealized Grid Network with Varied Demand

MFD CONSTRUCED USING PROBE DATA

To use MFD as a controlling tool, it should be proved that MFD can be constructed using limited probe data and the accuracy of such an MFD should be comparable to the MFD using the complete set of vehicle trajectory data in the network. To evaluate the

accuracy of the estimated MFD, three measures of effectiveness (MOEs) that incorporate the root mean square error (RMSE) of the average flow, density, and flow with density are used [15]. The constructed MFD using weighted penetration rates and the arithmetic mean penetration rate is illustrated in Figure 2. As can be seen, while comparing with the MFD build using the complete set of data, the MFD using weighted penetration rate is more accurate than the one using arithmetic mean. The MOEs in

Table 1 confirms this conclusion. The measures of errors are much smaller in both scenarios when the estimated penetration rates are weighted by travel time or travel distance.

These results demonstrate that the MFD can be accurately estimated using limited probe vehicle data. In fact, in a previous research effort, it was demonstrated that combining loop detector data and a small percentage of probe vehicle data, which are both readily available in the real world, the estimated MFD can be accurately constructed. The RMSE can be as small as 0.02 [30]. Consequently, the next step is to explore the usage of this tool in network congestion monitoring and control.



Figure 9. MFD Estimated Using Probe Data

	RMSE	(A)	(B)
Average	q (Veh/h)	41.9	66.83
	k (veh/mi)	2.02	4.84
	RMSE (q,k)	0.05	0.08
WT by Time	q (veh/h)	26.97	22.51
	k (veh/mi)	1.67	1.63
	RMSE (q,k)	0.03	0.03
WT by Distance	q (Veh/h)	18.56	22.16
	k (veh/mi)	2.32	1.82
	RMSE (q,k)	0.02	0.03

Table 8. RMSE for Scenarios A and B

Effects of different Network Control Strategies

To illustrate how the MFD can be used in network monitor and control, the demand and network are slightly changed in this section of the study. The demand from and to the area in the right upper quadrant of the network is set to be 50% higher than the rest of the network. The goal of designing the demand as this is to build a designated congested area in the network for testing the effectiveness of the control strategies exhibited below. The network is expanded to a network composed of 0.3 km long 2-lane links. The resulting network is designed to resemble a real typical downtown grid network. To use the MFD as a monitor and control tool, one needs to observe the shape of the MFD and adopts effective strategies before the MFD reaches a critical point, from where the whole network usually will deteriorate and the recovery of the network back to the original status typically will take a much longer time. Conveniently, in our simulation, we have set the demand levels clearly such that the onset of the congestion is known a priori. Therefore, in this study, we can skip the step to decide the point where the congestion starts and only concentrate on the impacts of the potential control strategies.

The MFD is plotted to show the effects of the following control strategies: 1). Perimeter control; 2). Adaptive signal control; and 3). Re-routing control.

Perimeter Control

In this scenario, the congested area is controlled by changing the entering and existing green time at the perimeter. Due to the imbalanced demands, it can be assumed that the congested area will be in the upright corner. Traffic signals at the boundary of the congested area (shown as the gray area in Figure 1) are set to decrease the green time in the direction of entering the area by varied amounts of time while increase the green

time in the direction of exiting the area by the same amount of time. Three different scenarios are tested for this strategy:

S1: The signals at the border of the gray area are changed such that entering green time will decrease by 1, 3, or 5 seconds while at the same time the exiting green time increases by 1, 3, or 5 seconds. The locations of the signals that adopt this modification are shown in Figure 1 as the green perimeters. This perimeter control starts from the onset of the congestion, e.g. the second hour of the simulation when the demand increases.

S2: In addition to the signal timing changing described in scenario 1, the signal timing upstream and downstream of the congested area and are along the heaviest traveled corridor is changed by the same amount. The locations of the additional signals involving green time adjustments are illustrated in red lines (shown in Figure 1). Similarly, the starting time of this change of signal timing is the second hour of the simulation.

S3: In addition to the signal timing changes described in scenario 2, the signals that are upstream and downstream of the congested area and are along the heaviest traveled corridor are set to be coordinated with the master signal at the border of the congested area (as shown in gray circles in Figure 1). Starting time of the control is the second hour of the simulation as illustrated above.

Scenario 1 is the scenario where the control strategy only aims at the boundary of the congested area. When the traffic demand increases and congestion starts, the adjustment of the green time decrease the amount of vehicles that entering the congested area to expedite the exiting of vehicles from the area. The drawback of this strategy is that the congestion will be shifted to upstream or downstream of the congested area. Therefore, the second scenario is designed to adjust the downstream and upstream flow rate of the congested area along the heaviest traveled corridor. To further explore the effects of perimeter control strategies in the first two scenario, the third strategy is where the traffic signals at the boundary of the congested area on the heavily traveled corridor are set to be the master signal where the signals upstream and downstream of them (on the red lines in Figure 1) are operating coordinately with them.

The results of scenario 1 is shown in Figure 3. Three different lengths of green time changes are: 5 seconds, 3 seconds, and 1 second. As can be seen, the 5 second green time change significantly worsen the congestion: the curve has a large hysteresis loop, indicating a slower recovery duration. The 1 second case is about the same as the base case while 3 seconds case decrease the condition a little.



Figure 10. Perimeter Control S1

The results of scenario 2 is shown in Figure 4. As can be seen that the control strategies all do not improve the congestion. All three different amount of green time changings have similar effects on the overall congestion. The MFDs of all the cases are overlapping with each other.

The results of scenario 3 is shown in Figure 5. As can be seen that with a coordinated master signal plan, the modification of green time with 1 or 3 seconds will help improve the congestion: both shifted the MFD to the left. Table 2 shows the MOEs of the results of the three cases. As can be seen that with a combination of green time adjustment (both at the perimeter and the upstream/downstream intersections along the congested corridors) and the coordinated traffic signal plans (with master signals located at the boundary of the congested area), the results improved significantly when the adjustment green time is 1 second. The total delay decreases by 12%. The average travel time, number of stops, fuel consumption and emissions all decreased compared to the base case. Although not as significant as the 1-second adjustment case, the 3-second adjustment also improves the congestion. The only exception is the 5-second adjustment, which basically maintains the results as the base scenario.



Figure 11. Perimeter Control S2



Figure 12. Perimeter Control S3

Table 9. MOEs for Perimeter Control

	Delay (s)	TT (s)	Stops	Fuel (L)	HC (g)	NOx (s)
Base	309	702	10.73	0.73	0.86	1.37
1s Case	273	665	10.43	0.71	0.85	1.35
	(-12%)	(-5%)	(-3%)	(-3%)	(-2%)	(-1%)
3s Case	288	688	10.53	0.72	0.85	1.36
	(-7%)	(-2%)	(-2%)	(-2%)	(-1%)	(-1%)
5s Case	313	725	10.63	0.73	0.86	1.36
	(1%)	(3%)	(-1%0	(0%)	(0%)	(0%)

Network-wide Signal Control

It was confirmed from the previous section that although the congested area is confined within the up-right quadrant of the network, modifying the signal timing at the boundary of the congested sub network only (scenario 1) is far less effective as the other two scenarios where more upstream and downstream signals are incorporated into the optimization process. Therefore, the second set of control strategies tested in this study is a network-wide signal optimization. The optimization process involves an update of the cycle length, phase split, and constrained offset signal at a frequency of 3 minutes, 5 minutes, 10 minutes, and 15 minutes. The offset optimization uses a hill-climb search technique to find the optimum offset timings that minimize a localized performance index that accounts for vehicle delay and stops.

Figure 6 illustrates the results of the three different optimization frequency. All cases are similar to each other. The 10 minute case is slightly worse than the others. All improving strategies helped with the congestion by decreasing the delay, number of

stops and fuel consumption and emissions. The corresponding statistics are listed in Table 3.



Figure 13. Effectiveness of Adaptive Signal Control

Table 10. MOEs for Adaptive Signal Control

	Delay (s)	TT (s)	Stops	Fuel (L)	HC (g)	NOx (g)
Base	309	702	10.73	0.73	0.86	1.37
3 min	203	595	9.03	0.66	0.80	1.27
	(-34%)	(-15%)	(-16%)	(-9%)	(-7%)	(-7%)
5 min	198	591	9.0	0.66	0.8015	1.27
	(-36%)	(-16%)	(-16%)	(-9%)	(-7%)	(-7%)
10 min	193	585	9.31	0.66	0.8156	1.30
	(-37%)	(-17%)	(-13%)	(-9%)	(-5%)	(-5%)
15 min	188	580	9.08	0.66	0.8074	1.28
	(-39%)	(-17%)	(-15%)	(-10%)	(-6%)	(-6%)

Network-wide Re-routing

The last control strategy tested in this study is to allow the vehicles to update their routes during their traveling. With the development of telecommunication technology and rapidly increasing usage smart phones and the associated traffic data interchange and communication platform, it is now very common to share and spread traffic condition information while traveling. Therefore, changing routes according to the prevailing traffic condition is becoming more accepted by travelers. This control strategy is based on the fact that travelers will share the traffic status information they are experiencing and at the same time utilizing data they obtained from other users in the network.

The updating frequency of routing is tested at 3, 5 minutes, 10 minutes, and 15 minutes. As can be seen from Figure 7 that, while there is no significant differences between the

different updating frequency settings, when the re-routing updating interval is set to 5 or 10 minutes, the results are slightly better than the other two. Re-routing strategies with all the different updating frequencies effectively improve the congestion and almost eliminate the hysteresis loop in the MFD by spreading the load over the network more evenly.

Combined Strategies

With the comparisons of the different settings above, a combination of all the strategies are used. In this combination of strategies, all the signals in the network are optimized at a frequency of 10 minutes, all vehicles in the network update their routings every 10 minutes, and at the same time, when the demand increases in the second hour of the simulation, the perimeter control is adopted where the entering traffic has a green time 1 second shorter while the exiting traffic has green time 1 second longer both at the perimeter of the congested area as well as the upstream and downstream of congested corridors. The intersections along the congested corridors are coordinated with the master signal at the boundary of the corresponding upstream or downstream. Interestingly, as can be seen from Figure 8, no significant improvement is accomplished using the combination of the strategies.







Figure 15. Effectiveness of Control Strategies

CONCLUSIONS AND DISCUSSIONS

The paper used a distance and time weighted average probe penetration rate to estimate the MFD and illustrated how the MFD can be used to monitor and evaluate the effectiveness of network-wide control strategies to alleviate congestion. Several control strategies were tested. The study conclusions are: 1) The MFD is effective in monitoring, assisting in the design of control strategies, and evaluating the effectiveness of those strategies. 2) The majority of the strategies tested show their effectiveness in improving the system-wide travel time, decreasing traffic delay, and reducing fuel consumption levels. The most effective strategy was the network-wide adaptive signal control, which decreased delays by up to 40%. Average fuel consumption levels decreased by up to 10%. 3) A simple addition of multiple control strategies may not be effective. As shown in the paper, the dynamic re-routing and adaptive signal control systems operated in isolation produce more significant network-wide MOE savings compared to combining the strategies. This deterioration in performance is attributed to the fact that the systems are not fully-integrated or collaborative. Consequently, additional research is required to develop a fully integrated and collaborative control system that combines dynamic routing and traffic signal control. The impact of this integrated system on various measures of effectiveness should be studied and quantified.

ACKNOWLEDGMENT

This research was supported by the Mid-Atlantic University Transportation Center (MAUTC).

REFERENCES

[1] C. F. Daganzo, "Urban gridlock: Macroscopic modeling and mitigation approaches," *Transportation Research Part B: Methodological*, vol. 41, pp. 49-62, 2007.

[2] C. F. Daganzo and N. Geroliminis, "An Analytical Approximation for the Macroscopic Fundamental Diagram of Urban Traffic," *Transportation Research Part B: Methodological*, vol. 42, pp. 771-781, 2008.

[3] N. Geroliminis and C. F. Daganzo, "Macroscopic modeling of traffic in cities," presented at the 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.

[4] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings," *Transportation Research Part B: Methodological*, vol. 42, pp. 759-770, 2008.

[5] N. Geroliminis and J. Sun, "Hysteresis phenomena of a Macroscopic Fundamental Diagram in freeway networks," *Transportation Research Part A: Policy and Practice*, vol. 45, pp. 966-979, 2011.

[6] N. Geroliminis and J. Sun, "Properties of a well-defined macroscopic fundamental diagram for urban traffic," *Transportation Research Part B: Methodological*, vol. 45, pp. 605-617, 2011.

[7] C. Buisson and C. Ladier, "Exploring the Impact of Homogeneity of Traffic Measurements on the Existence of Macroscopic Fundamental Diagrams," *Transportation Research Record,* vol. 2124, pp. 127-136, 2009.

[8] N. Geroliminis and D. Levinson, "Cordon Pricing Consistent with the Physics of
 Overcrowding," in *Transportation and Traffic Theory 2009*, ed: Springer Science, 2009, pp. 219-240.

[9] M. Keyvan-Ekbatani, A. Kouvelas, I. Papamichail, and M. Papageorgiou, "Exploiting the fundamental diagram of urban networks for feedback-based gating," *Transportation Research Part B: Methodological*, vol. 46, pp. 1393-1403, 2012.

[10] M. Keyvan-Ekbatani, M. Papageorgiou, and I. Papamichail, "Urban congestion gating control based on reduced operational network fundamental diagrams," *Transportation Research Part C*, vol. 33, pp. 74-87, 2013.

[11] M. Keyvan-Ekbatani, M. Papageorgiou, and I. Papamichail, "Perimeter traffic control via remote feedback gating," *Procedia – Social and Behavioral Sciences*, vol. **Vol.** 111, pp. pp. 645-653, 2014.

[12] M. Ramezani, J. Haddad, and N. Geroliminis, "Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control," *Transportation Research B*, vol. 74, pp. 1-19, 2015.

[13] J. Haddad, M. Ramezani, and N. Geroliminis, "Cooperative traffic control of a mixed network with two urban regions and a freeway," *Transportation Research B*, vol. 54, pp. 17-36, 2013.

[14] J. Haddad and N. Geroliminis, "On the stability of traffic perimeter control in two-region urban cities," *Transportation Research Part B*, vol. 46, pp. 1159-1176, 2013.

[15] A. Nagle and V. Gayah, "The accuracy of network-wide traffic state esimations using mobile probe data," *forthcoming, Transportation Research Record,* 2014.

[16] N. Zheng, R. A. Waraich, K. W. Axhausen, and N. Geroliminis, "A dynamic cordon pricing scheme combining the Macroscopic Fundamental Diagram and an agent-based traffic model," *Transportation Research Part A: Policy and Practice*, vol. 46, pp. 1291-1303, 2012.

[17] V. L. Knoop, S. P. Hoogendoorn, and J. W. C. V. Lint, "Routing Stretegies Based on Macroscopic Fundamental Diagram," *Transportation Research Record*, vol. 2315, pp. 1-10, 2012.

[18] L. Zhang, T. M. Garoni, and J. de Gier, "A comparative study of Macroscopic Fundamental Diagrams of arterial road networks governed by adaptive traffic signal systems," *Transportation Research Part B: Methodological*, vol. 49, pp. 1-23, 2013.

[19] M. Keyvan-Ekbatani, M. Papageorgiou, and I. Papamichail, "Urban congestion gating control based on reduced operational network fundamental diagrams," *Transportation Research Part C: Emerging Technologies*, vol. 33, pp. 74-87, 8// 2013.

[20] M. Van Aerde and H. Rakha, "INTEGRATION © Release 2.40 for Windows: User's Guide – Volume I: Fundamental Model Features," M. Van Aerde & Assoc., Ltd., Blacksburg2013.

[21] M. Van Aerde and H. Rakha, "INTEGRATION © Release 2.40 for Windows: User's Guide – Volume II: Advanced Model Features," M. Van Aerde & Assoc., Ltd., Blacksburg2013.

[22] E. Chamberlayne, H. Rakha, and D. Bish, "Modeling the Capacity drop Phenomenon at Freeway Bottlenecks using the INTEGRATION Software," *Transportation Letters: The International Journal of Transportation Research*, vol. 4, pp. 227-242, 2012.

[23] H. A. Rakha, K. Ahn, and K. Moran, "INTEGRATION Framework for Modeling Eco-routing Strategies: Logic and Preliminary Results," *International Journal of Transportation Science and Technology*, vol. 1, pp. 259-274, 2012.

[24] H. Rakha, P. Pasumarthy, and S. Adjerid, "The INTEGRATION framework for modeling longitudinal vehicle motion," in *TRANSTEC*, Athens, Greece, 2004.

[25] H. Rakha and Y. H. Zhang, "INTEGRATION 2.30 framework for modeling lane-changing behavior in weaving sections," *Traffic Flow Theory and Highway Capacity and Quality of Services 2004*, pp. 140-149, 2004.

[26] M. Van Aerde, B. Hellinga, M. Baker, and H. Rakha, "INTEGRATION: An overview of traffic simulation features," *Transportation Research Records*, 1996.

[27] H. Rakha and Y. Zhang, "INTEGRATION 2.30 framework for modeling lane-changing behavior in weaving sections," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1883, pp. 140-149, 2004.

[28] L. C. Edie, "Discussion of traffic stream measurements and definitions," in *the 2nd International Symposium on the Theory of Traffic Flow*, 1965.

[29] A. Nagle, Gayah, V., "THE ACCURACY OF NETWORK-WIDE TRAFFIC STATE ESTIMATIONS USING MOBILE PROBE DATA," *forthcoming, Transportation Research Record,* 2014.

[30] J. Du, H. Rakha, and V. Gayah, "Deriving Macroscopic Fundamental Diagrams from Probe Data: Issues and Proposed Solutions," *Submitted to Transportation Research Part C.*, 2015.