USING MOBILE PROBE DATA TO ESTIMATE TRAFFIC STATES
AND THE EFFECTS OF PERIMETER CONTROL ON URBAN NETWORKS

A Thesis in
Civil Engineering

by

Andrew S. Nagle

© 2014 Andrew S. Nagle

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2014
The thesis of Andrew S. Nagle was reviewed and approved* by the following:

Vikash V. Gayah  
Assistant Professor of Civil and Environmental Engineering  
Thesis Advisor

Eric T. Donnell  
Associate Professor of Civil and Environmental Engineering

Paul P. Jovanis  
Professor of Civil and Environmental Engineering

Peggy A. Johnson  
Professor of Civil and Environmental Engineering  
Head of the Department of Civil and Environmental Engineering

*Signatures are on file in the Graduate School
ABSTRACT

Aggregate models of traffic on urban networks have been studied for decades extending back to the 1960s and various theories have been developed to describe networks as a function of measurable vehicular properties. Earlier models found that a monotonically decreasing relationship existed between the average speed and flow in the network. However, a monotonically decreasing relationship is only practical under free-flow or light conditions; therefore, these earlier models were only able to describe the uncongested regime. Recent work has shown that the average flow and density of a network are related and can dynamically describe both uncongested and congested conditions in an urban network. This relationship, known as the Macroscopic Fundamental Diagram (MFD), can be used to describe large-scale network dynamics and develop network-wide control strategies to improve efficiency.

This work examined the accuracy of combining data from fixed sensors and mobile probe vehicles to estimate traffic states in real-time. Formulae were developed to estimate the average flow, density, speed, accumulation, and exit flow, along with the uncertainty of these estimates. The methodology was tested using data obtained from several micro-simulated networks to replicate realistic urban transportation networks. Tests using the micro-simulated models showed that estimations were within 10% of the true value for average flow, density, accumulation, and exit flow when 20% of the circulating vehicles served as probes and average speed estimations were shown to be even more accurate. In addition, the mobile probe estimation methodology was used to inform a network-wide perimeter metering strategy on an idealized grid network. These
results were compared to metering based on actual traffic states, and the results illustrated that metering can significantly reduce total travel times compared to the non-metering case if 25% of the vehicles in the network serve as probe vehicles. In general, this work shows that the estimation method may be used to inform network-wide metering strategies, which improve network stability and reduce congestion without additional costly infrastructure improvements.
# TABLE OF CONTENTS

List of Figures .................................................................................................................... vii

List of Tables ....................................................................................................................... viii

Acknowledgements ............................................................................................................... ix

Chapter 1 Introduction ........................................................................................................ 1

Chapter 2 Literature Review ............................................................................................... 5

  2.1. Background on macroscopic models of urban networks ........................................... 5
      2.1.1. Early theories of aggregate level traffic relationships at the city-scale ............ 5
      2.1.2. The existence of the macroscopic fundamental diagram on urban networks ....... 7
      2.1.3. Network-wide control strategies and the macroscopic fundamental diagram ....... 12

Chapter 3 Estimation Methodology .................................................................................... 17

  3.1. Metrics of interest ...................................................................................................... 17
  3.2. Estimating the metrics of interest using probe data ................................................. 18
  3.3. Estimating the probe penetration rate ...................................................................... 20
  3.4. Accuracy of the estimates ....................................................................................... 21
      3.4.1. Accuracy of the probe penetration rate .............................................................. 21
      3.4.2. Accuracy of the metrics of interest ................................................................. 22
      3.4.3. Applications using fractional error ................................................................. 24

Chapter 4 Tests Using Micro-Simulation ........................................................................... 26

  4.1. Idealized network ..................................................................................................... 26
  4.2. Orlando network ....................................................................................................... 28
  4.3. Accuracy of estimations ........................................................................................... 31
      4.3.1. Accuracy of probe penetration rate ................................................................. 31
      4.3.2. Comparing probe vehicle data to all vehicle data ........................................... 33
      4.3.3. Estimating metrics of interest ....................................................................... 34
      4.3.4. Accuracy of variance estimations ................................................................... 38
      4.3.5. Estimating the macroscopic fundamental diagram ......................................... 42
      4.3.6. Quantifying the accuracy of MFD estimations ................................................. 47
      4.3.7. Applying estimation methods using fractional error ....................................... 50
  4.4. General remarks ....................................................................................................... 51
Chapter 5 Network-wide Perimeter Control Strategy .................................................. 53

5.1. Perimeter control strategy .................................................................................. 53
5.2. Traffic state estimations .................................................................................... 55
   5.2.1. Link method ................................................................................................. 56
   5.2.2. Mobile probe method .................................................................................. 57
5.3. Simulated network ............................................................................................... 58
5.4. Results .............................................................................................................. 60
   5.4.1. Non-metered scenarios for link and probe methods ................................. 62
   5.4.2. Metered scenarios for link and probe methods ......................................... 64

Chapter 6 Conclusions ............................................................................................ 69

References .............................................................................................................. 73
LIST OF FIGURES

Figure 2-1. Weighted average flow vs. unweighted average occupancy in Yokohama, Japan (source: Geroliminis and Daganzo, 2008). ........................................ 9

Figure 2-2. Hysteresis loop during the morning peak period in Toulouse, France (source: Buisson and Ladier, 2009). ................................................................. 11

Figure 4-1. A portion of the 16x16 idealized AIMSUN simulated network. .......... 27

Figure 4-2. Idealized network demand as a function of hour of day. ..................... 28

Figure 4-3. Downtown network of Orlando, FL (source: Gayah and Dixit, 2013). .... 30

Figure 4-4. Theoretical and actual variance of the estimated probe penetration rate. ....................................................................................................................... 32

Figure 4-5. Box plots showing the accuracy of estimates for average: a) flow; b) density; c) speed; d) accumulation; and e) exit flow for the idealized network. ................................................................. 36

Figure 4-6. Box plots showing the accuracy of estimates for average: a) flow; b) density; and c) speed for the Orlando network..................................................... 37

Figure 4-7. Comparison of theoretical and actual variances for average: a) flow; b) density; c) speed; d) accumulation; and e) exit flow for the idealized network. .......................................................................................... 40

Figure 4-8. Comparison of theoretical and actual variances for average: a) flow; b) density; and c) speed for the Orlando network.............................................. 41

Figure 4-9. Actual and estimated MFDs for probe penetration rates of: a) 5%; b) 15%; c) 30%; d) 40%; and e) 50% for the idealized network. ................................. 44

Figure 4-10. Actual and estimated MFDs for probe penetration rates of: a) 5%; b) 15%; c) 30%; d) 40%; and e) 50% for the Orlando network. ................................. 46

Figure 4-11. Visual inspection of the increase in MFD accuracy as the probe penetration rate increases................................................................. 49

Figure 5-1. Network-wide metering strategy. ......................................................... 54

Figure 5-2. Accounting for estimation uncertainty................................................. 58

Figure 5-3. Metering strategy demand distribution................................................. 59

Figure 5-4. Travel time savings............................................................................... 62
Figure 5-5. MFD of idealized network using a) link estimation method and b) probe estimation method.
........................................................................................................63

Figure 5-6. Travel time savings for a) link strategy and b) probe strategy for different reduced percentages.
........................................................................................................65

LIST OF TABLES

Table 4-1. Results of the statistical t-tests. ..................................................................................34

Table 4-2. Root mean square error of flow, density, and simultaneous flow and density for various probe penetration rates.........................................................48

Table 5-1. Travel time savings for the full and reduced metering strategies...........66
ACKNOWLEDGEMENTS

I want to thank my thesis advisor, Dr. Vikash V. Gayah, for his continual guidance, kindness, and insightfulness. The countless evenings sitting in front of his dry-erase board paved the way for my research and provided a medium to share our passions in life. I also want to thank my committee members, Dr. Eric T. Donnell, for inspiring my interest in the transportation field, and Dr. Paul P. Jovanis, for his thoughtful discussions and invaluable input both inside and outside the classroom. I am also forever thankful for Julianne McCobin, for her unparalleled support and friendship. Finally, I want to thank my parents for the encouragement to pursue my dreams.

This research was supported by the Mid-Atlantic Universities Transportation Center and the Federal Highway Administration Dwight D. Eisenhower Fellowship program.
Chapter 1

Introduction

Traffic congestion is a severe problem in urban transportation networks across the world due to its negative effects on safety, reliability, and the environment. Congestion in 2012 caused urban Americans to travel 5.5 billion hours more, purchase an extra 2.9 billion gallons of fuel for a total cost of $121 billion, and release 56 billion pounds of additional carbon dioxide greenhouse gas into the atmosphere (Schrank et al., 2012). Transportation agencies have attempted to mitigate this problem for decades by constructing new infrastructure to provide more capacity to meet increasing demands. However, expanding existing or constructing new infrastructure in urban areas is infeasible due to limited space and high costs. It has also been shown that road improvements that attempt to mitigate congestion can actually attract traffic from other routes and encourage longer and more frequent trips—a problem commonly known as generated traffic or induced travel (Litman, 2001). This has led to pushes in alternative methods to reduce congestion by utilizing existing infrastructure.

A common method is the use of Intelligent Transportation Systems (ITS) to gain useful information from drivers’ vehicles to monitor traffic conditions. If conditions are accurately measured in real-time, data can be transmitted via communication systems, and ideal control strategies may be implemented. Well-established technologies have existed for several decades that collect data such as vehicle counts, occupancies, travel times, delays, and speeds that are processed and communicated to transportation users via
dynamic message signs, traffic signal systems, and road weather information systems. The most common methods are embedded sensors in the infrastructure such as inductive loop detectors or over-roadway sensors such as video image processors. Although these technologies are common practice and are extremely useful to detect vehicles at intersections or to determine freeway travel speeds, they are infrastructure intensive, costly to construct and maintain, and not very effective when vehicles are stopped. The lattermost of these problems tends to arise in urban networks where loop detectors are placed near intersections. Often in these situations, queues spill back over the detector, which causes inaccurate measurements (Buisson and Ladier, 2009; Geroliminis and Sun, 2011b; Courbon and Leclercq, 2011). In addition, these technologies are typically managed on smaller scales—i.e., along a specific coordinated corridor or section of freeway—and are not used to describe network-wide macroscopic properties.

Recent advancements in technologies such as Global Positioning System (GPS) devices provide a new type of sensor that eases the collection of traffic data. These technologies do not require additional infrastructure, greatly reduce the problems previously discussed, and are becoming increasingly common in fleet vehicles and private automobiles. The GPS devices collect vehicle information such as time, speed, and acceleration that can be aggregated over time to calculate network properties for real-time traffic prediction. In addition, the Federal Highway Administration (FHWA) is leading research on the Connected Vehicle program (U.S. Department of Transportation, 2013) that utilizes advanced wireless communications and smart infrastructure that can be directly used to collect vehicle data.
Most research in the field of mobile probe data focuses primarily on estimating traffic properties on arterial roadways (Srinivasan and Jovanis, 1996) or speeds on freeways (Hofleitner et al., 2012) and do not consider urban-scale networks. However, recently Daganzo (2007) proposed that a unique and reproducible relationship exists between the average flow and density of vehicles in a small urban neighborhood or network. This relationship, now commonly known as the Macroscopic or Network Fundamental Diagram (MFD or NFD), provides a feasible option to utilize GPS data to predict urban network conditions in real-time. The MFD can be used to describe the dynamics of urban traffic networks that unveil interesting phenomena and control policies. For example, it was demonstrated using this macroscopic model that urban networks have an innate tendency towards gridlock when congested, and an optimal control scheme was developed to improve network efficiency by metering vehicle entry into the network (Daganzo, 2007; Geroliminis and Daganzo, 2008).

The MFD must be estimated before it can be applied to describe the behavior of a network and implement control. Estimation of the MFD is a highly data-intensive process, and only few studies have been able to derive MFDs from empirical data (Geroliminis and Daganzo, 2008; Buisson and Ladier, 2009). These successful studies relied heavily on loop detector data, whose placement have been shown to significantly affect macroscopic relationships (Buisson and Ladier, 2009; Geroliminis and Sun, 2011b; Courbon and Leclereq, 2011). Although Keyvan-Ekbatani et al. (2012) showed that limited detector information is sufficient to implement gating control on a specific network, the placement of detectors may provide inaccurate MFDs. Since network control strategies such as pricing are highly sensitive, more accurate MFDs are required.
Gayah and Dixit (2013) developed a method to estimate average network densities in real-time by combining probe vehicle data with a known MFD. However, it requires that the MFD is known, it cannot predict traffic states in light flow conditions, and it cannot be used if the MFD exhibits any hysteresis. As an alternative approach, this current work provides a general method to accurately estimate average network flow, density, speed, accumulation, and exit flow under both uncongested free-flow states and congested states without prior knowledge of the MFD. This method allows for GPS data to be used to derive the MFD and monitor traffic conditions in real-time. With accurate state estimations, the information may be used to inform vehicle re-routing schemes and network-wide traffic control strategies to improve network stability and performance.

The objectives of this work are to first determine a methodology that combines mobile probe vehicle data with macroscopic traffic models for network-wide traffic state estimations. Next, the uncertainties of the traffic state estimations are to be quantified. To confirm the accuracy of the estimations and uncertainties, an idealized grid network and a more realistic simulated model of downtown Orlando, FL are used. Additionally, macroscopic traffic control strategies that utilize mobile probe data to improve overall network efficiency are developed. These strategies are assessed using the idealized micro-simulation model by comparing the effectiveness of the mobile probe data estimations with data from all vehicles in the network. Lastly, the practicality of the traffic state estimations is evaluated.
Chapter 2

Literature Review

This chapter provides background information on the foundation of aggregate models of urban traffic, the existence of the MFD on urban networks, recent work that describes properties of the MFD, the characteristics of networks that affect the aggregate flow-density relationship, and network-wide control strategies based on the MFD.

2.1. Background on macroscopic models of urban networks

Theories have been developed in the last four decades that describe traffic dynamics at an aggregate level. These macroscopic models tried to predict network properties as a function of measured parameters using empirical data. The first section of this chapter focuses on theories and shortcomings of early macroscopic models, and the second section discusses the development of a new macroscopic model that is able to adequately describe traffic dynamics on an aggregate level. Lastly, several network-wide control strategies based on the MFD are discussed.

2.1.1. Early theories of aggregate level traffic relationships at the city-scale

The earliest of the aggregate traffic models attempted to determine the maximum flow of vehicles that can enter an urban city as a function of the capacity of the city roadways, the nature of the city, and the types of trips to and from the network (Smeed,
1966). For several following years, more complex models were developed that related the average speed and average flow in a city network. A study conducted by Thomson (1967) used extensive data collected from central London for many years and found that there was seemingly a linearly-decreasing relationship between speed and flow. This model was expanded by Wardrop (1968), who derived a relationship between the average speed of a trip to central London and the flow of traffic on the network. This model also included characteristics of the roadway, such as the number and spacing of intersections and type of traffic controls on the network. Zahavi (1972) developed a model to estimate the relationship between traffic flow, density, and weighted space mean speed from data collected in two different cities. Zahavi proposed that speeds were inversely related to flows. All of these earlier works proposed that the relationship between average vehicle speed and average network flow is monotonically decreasing. This relationship means that the models are only valuable during free-flow or light traffic conditions, whereas congested conditions in a city could not be described.

A more realistic two-fluid model of town traffic was developed by Herman and Prigogine (1979). The model proposed that the speed in an urban network was related to the fraction of stopped cars at any given time. This provided a physically realistic model that related the speed and density of an urban traffic network. The model was then empirically verified by Herman and Ardekani (1984) with traffic experiments in Austin, Texas. The data supported the two-fluid model and illustrated that the average speed in a network is proportional to the fraction of vehicles stopped, raised to some power. A preliminary analysis of aerial photographs determined a two-fluid model parameter that related the fraction of stopped vehicles to the jam or maximum density. Although this
provided significant advancements in urban aggregate traffic modeling, equilibrium conditions could only be described, and the model was unable to identify how traffic conditions changed over time.

2.1.2. The existence of the macroscopic fundamental diagram on urban networks

An aggregate model of traffic was introduced by Daganzo (2007) that dynamically described single neighborhoods and systems with interconnected neighborhoods. This model, unlike earlier traffic models, only required observable inputs and could be used without knowing origin-destination information or microscopic system dynamics. This model operated primarily on the premise that oversaturated networks behave chaotically. Daganzo (2007) proposed to alleviate this problem by modeling city traffic at an aggregate level and focusing on control policies that can perform well independent of complex input variables and the unpredictability of driver behavior. It was argued that a well-defined relationship between average network flow and density, now commonly referred to as the Macroscopic Fundamental Diagram (MFD), should arise if a network is uniformly loaded so that all links are similarly congested. The study found that urban networks have an innate tendency towards gridlock when congested, and an optimal control strategy was developed to improve network efficiency by metering vehicle entry into the network.

Daganzo and Geroliminis (2008) showed that an MFD relating the average flow and density in a network must exist on any street network with no turns regardless of whether intersections are controlled by fixed timed traffic signals. The study also
proposed that the relationship should arise for homogeneous networks with multiple routes because this allows drivers to alter their routes and distribute themselves evenly throughout the network.

The proliferation of ITS technologies and collection of traffic data in cities allowed Geroliminis and Daganzo (2008) to strongly suggest the existence of an MFD using empirical data. The study revealed that an MFD linking space-mean flow, density, and speed existed in Yokohama, Japan. The weighted average flow versus unweighted average occupancy is illustrated in Figure 2-1. Contrary to earlier aggregate models, the empirical speed-flow relationship was not monotonic, and the results were consistent across the morning and evening peak hours. Geroliminis and Daganzo (2008) proposed that neighborhoods should have a well-defined reproducible MFD, which can be used to improve accessibility by monitoring the city’s trip completion rate. This can be accomplished using strategies outlined in Daganzo (2007), such as pricing, metering, and other perimeter control strategies based on network density and speeds. If the MFD is known and the state of traffic can be continuously monitored, transportation agencies can determine whether the network is serving traffic at desired levels. Lastly, the study suggests that many crowded cities can benefit from these controls if vehicles are equipped with GPS devices that serve as city-wide probes.
Figure 2-1. Weighted average flow vs. unweighted average occupancy in Yokohama, Japan (source: Geroliminis and Daganzo, 2008).

Recent work shows that the average flow in a network is a function of the average density of the network and the spatial variability of the density. Mouzliman et al. (2010) tested how spatial variability of traffic densities can affect urban network performance using simulation. They found that the variance of density is a key variable that is necessary for the existence of an invariant MFD and to provide a well-defined macroscopic relationship, even if origin-destination flows are significantly different. The authors proposed that strategies suggested by Daganzo (2007) and Geroliminis and Daganzo (2008), such as restricting access to neighborhoods that exceed some critical density or restricting the inflow to congested areas, would enhance traffic performance and reduce the variability of vehicle densities. Geroliminis and Sun (2011b) also found that the spatial distribution of vehicle density in an urban network is a key component that affects the shape, scatter, and existence of a well-defined MFD.
The impacts of heterogeneous congestion distributions on network stability and bifurcations were further explained by Daganzo et al. (2011). When networks consist of multiple overlapping routes, the observed flows in the congested state for a given density are less than what would be predicted by a network with homogeneously congested links and better distributed routes. Daganzo et al. (2011) attributed these phenomena to asymmetric equilibrium patterns, and networks jammed at lower densities and had lower associated flows if traffic was evenly distributed. This phenomenon is especially damaging to larger-scaled networks once the network exceeds some critical density threshold. Therefore, it is imperative that urban networks do not exceed this threshold to ensure optimal network stability and efficiency.

Knoop et al. (2013) proposed a generalized macroscopic fundamental diagram (GMFD) to relate the average flow in a network as a function of the average density in a network and the spatial inhomogeneity of density across the entire network. Empirical data from the Amsterdam ring road illustrated a continuous function of accumulation and density inhomogeneity (Knoop and Hoogendoorn, 2012), which showed that the GMFD explains more of the spread in production than the MFD, especially near maximum production levels.

The impacts of homogeneity assumed by Geroliminis and Daganzo (2008) were explored using empirical data in Toulouse, France by Buisson and Ladier (2009). It was found that different roadway classifications and the location of fixed detectors with respect to traffic signals have a strong impact on the shape of the MFD. The researchers witnessed a hysteresis-like shape in the MFD where some data exhibited several flow values for a single density, as illustrated in Figure 2-2. This hysteresis phenomenon was
further explained by Gayah and Daganzo (2011), finding that hysteresis in the flow-density relationship is likely to arise when driver adaptivity is low. Higher levels of driver adaptivity resulted in more uniformly-congested links, which reduced the occurrence and magnitude of the hysteresis effect. Hysteresis and capacity phenomena were also characterized by Saberi and Mahmassani (2012) by examining archived loop detector data in freeway networks. It was observed that hysteresis phenomena did not follow a consistent repetitive pattern but instead varied across days and networks. These patterns depended on the shape, size, and spatial distribution of congestion.

Figure 2-2. Hysteresis loop during the morning peak period in Toulouse, France (source: Buisson and Ladier, 2009).

The MFD that relates the average flow and density of a network was also found to be related to the mean and day-to-day variation in travel time (Gayah et al., 2013).

Theoretical predictions and simulated data confirmed that counter-clockwise hysteresis
loops exist between the mean and variance of travel times under periods where clockwise hysteresis loops are present in the MFD. Mahmassani et al. (2012) connected the network-wide travel time reliability and the MFD by plotting simulated models of four networks’ flow-density relationships and the travel time reliability. A fundamental relationship between the average network flow, density, and network travel time reliability was found. The standard deviation of travel time increased linearly with an increased mean travel time, and the distance-weighted standard deviation of travel time rate increased monotonically with network density. This further illustrates the importance of maintaining network stability and homogeneous density patterns to increase network flows and reduce travel times within a network.

2.1.3. Network-wide control strategies and the macroscopic fundamental diagram

To mitigate the heterogeneity of congestion in urban networks and delay the onset of oversaturation, researchers have examined several control strategies to homogenize network density. Daganzo (2007) first suggested an optimal control strategy to improve network performance by metering vehicle entry into a network. Keyvan-Ekbatani et al. (2012) developed and implemented this control strategy, which concentrated on metering vehicles entering the network at the perimeter. A feedback-based control model that utilized the MFD was developed and tested in a microscopic simulation environment to protect the network from oversaturated conditions. The gating strategy led to significant improvements in network performance compared to a non-gating control case.
Additionally, adaptive signal control on the network scale shows promise in improving overall network performance. The effects of adaptive traffic signal control systems on the MFD were first studied by Zhang et al. (2013). The authors compared two realistic signal systems and observed well-defined stationary MFDs, which relied heavily on the type of adaptive control system and the heterogeneity of density. Adaptive traffic signals and their effects on the MFD were further explored by Gayah and Gao (2014). Abstractions of an idealized urban network were used to explore the impact of adaptive signals on a network’s ability to avoid inhomogeneous congestion conditions. The adaptive green signals helped mitigate the network’s unevenness when the network was moderately congested and delayed the occurrence of gridlock by providing higher average network flows. However, the adaptive control strategies had no effect on aggregate network behavior when the network was overly congested.

Dynamic routing control based on the macroscopic fundamental diagram can also spread congestion across a network and improve network efficiency. Daganzo et al. (2011) illustrated how, if drivers select their routes adaptively in response to traffic conditions, the critical density of a network considerably increases. This theory was expanded by Knoop et al. (2012) by implementing a routing control scheme based on a subnetwork fundamental diagram to spread congestion homogeneously throughout the network. Several routing strategies based on aggregate speed information in a network and accumulation in a subnetwork were used to determine how the MFD changes with control. Routing control based on full speed data and aggregate subnetwork data showed benefits to network performance, but the shape of the MFD was largely unchanged.
However, if the route paths were optimized within a subnetwork, the MFD was shown to change.

A calibrated model of the Chicago central business district (CBD) was used by Mahmassani et al. (2013) to study the effects of dynamic vehicle routing on a network’s gridlock properties. Dynamic control strategies mitigated congestion and improved accessibility and mobility by reducing the gridlock size and propagation speed, while increasing the gridlock recovery speed. Additionally, Gayah and Gao (2014) illustrated that adaptive driver routing can provide network stability in cases where adaptive signal control cannot. However, Leclercq and Geroliminis (2013) found that, although flow distributions among routes smoothly varied with respect to the total flow, discontinuities in networks provide challenges to developing a control strategy based on routing in the congested regime.

Furthermore, this macroscopic model can be used to develop pricing strategies that improve mobility and relieve congestion in cities. Geroliminis and Levinson (2009) used the MFD to describe recurring congestion in a network. A dynamic model was developed using a cordon-based congestion pricing scheme to reduce congestion during the morning commute. By applying an optimal toll at a bottleneck, delays were greatly reduced and the length of the rush hour was shortened, which provided travel and schedule savings. This model was extended by Gonzales and Daganzo (2012) to consider competing modes such as cars and transit. A model was developed that shows how a transit agency should operate and how to set tolls to minimize a system’s generalized cost. By providing public transit, both car and transit users experience a lower cost, and introducing public transit reduced the duration of the rush period and the overall cost.
The MFD can also be used to optimize the allocation of urban street space to transit services. Daganzo et al. (2012) constructed two parsimonious macroscopic models using an analytical method for transit and empirical data from Yokohama, Japan, to balance the costs of both modes. By using little observable data, such as the number of car or transit trips, average vehicular speed, and average trip length, optimal costs and space allocation for cars and transit can be determined. Furthermore, Zheng and Geroliminis (2013) explored how the distribution of multiple transportation models affected a network’s performance. A multimodal MFD for allocation of road space among models was developed to minimize total person hours traveled (PHT) in the network, given relevant data about a city. A two-region bi-modal case study was performed for a city center, which found that dynamically allocating space to transit minimized PHT by serving a higher number of passengers during the peak period.

Although the MFD has been increasingly studied in the past few years, very few effective estimation methodologies were presented for real-time traffic management. Keyvan-Ekbatani et al. (2013) found that the feedback-gating procedure developed by Keyvan-Ekbatani et al. (2012) can be used with only a subset of the network’s links, which greatly reduces the amount of data needed, providing a perimeter control scheme that can be implemented in real-time. However, using loop detector data may provide inaccurate MFDs. This may cause problems since some network-wide control strategies, such as pricing, are very sensitive to the MFD.

With the recent technological advancements and understanding of the MFD relationship, Gayah and Dixit (2013) conducted a study to provide a method to estimate average network densities in real-time by combining mobile vehicle probe data with a
known MFD of urban traffic. The method was tested using micro-simulation data and was shown to be reliable at predicting congestion and estimating densities near congested states. However, several limitations existed, including: since a known MFD was necessary, the method could not accurately predict traffic conditions in light traffic states, and it may not be used if the MFD exhibits any significant amount of hysteresis. Therefore, it is pertinent that more accurate methodologies are studied to provide estimations of average flows and densities that do not rely on the existence of an MFD and can be used in uncongested and congested states, which this work aims to achieve.
Chapter 3

Estimation Methodology

This chapter discusses the methodology used to estimate network-wide traffic metrics by introducing the traffic metrics of interest and the estimation procedure. First, each metric is defined, and the procedure to measure the metric is discussed. Then, assuming the probe penetration rate is known a priori, a method to estimate each metric using data from probe vehicles is developed. Since the probe penetration rate may not be known a priori, a method to estimate the fraction of probe vehicles in the network is explored. Lastly, using only data from probe vehicles, the uncertainties of the traffic metrics are derived.

3.1. Metrics of interest

A network’s MFD must be estimated by calculating the average flow, $q$ [veh/hr], and average density, $k$ [veh/mi], of vehicles in the network. These properties are defined using Edie’s generalized definitions of traffic (Edie, 1965) for a given analysis period:

$$q = \frac{d_T}{LT} = \frac{N\bar{d}}{LT}, \quad \text{and}$$

$$k = \frac{t_T}{LT} = \frac{N\bar{e}}{LT},$$

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

For networks with well-defined MFDs, the average speed serves as a proxy for the level of congestion or density (Daganzo et al., 2012), since each state on the MFD is represented by a unique value of average vehicle speed, \( v \) [mi/hr]. Thus, as per Edie, the average vehicle speed in the network is defined as:

\[
v = d_T / t_T = \bar{d} / \bar{t}.
\]  \( (3.3) \)

The Network Exit Function (NEF), which relates the average flow that vehicles are able to exit the network, \( f \) [veh/hr], to the average accumulation inside the network, \( n \) [veh], may be more convenient for modeling purposes. These two variables are defined as:

\[
f = qL / l = N\bar{d} / Tl = N_e / T \quad \text{and} \\
n = kL = N\bar{t} / T,
\]  \( (3.4) \) \( (3.5) \)

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 the average traffic conditions within a network.

### 3.2. Estimating the metrics of interest using probe data

The metrics of interest can be directly calculated if trajectory data are available for all vehicles that enter and exit the network, by providing the values of \( d_T, t_T, \) and \( N_e \). Unfortunately, trajectory data for all vehicles in the network is highly unlikely, and only a subset of vehicle trajectories is available. With only a fraction of the vehicles serving as
mobile probes, the metrics of interest cannot be calculated. However, it may be possible to estimate these metrics using only probe data if the probe penetration rate in the network, $\rho$, is known a priori. Thus, the number of probe vehicles in the network during any analysis period should be equal to a fixed proportion of the total number of vehicles in the network during that time interval, $N_p = N\rho$. Additionally, if the probe vehicles are uniformly distributed throughout space, the average distance traveled and average time spent in the network by the probe vehicles should accurately represent the average distance traveled and average time spent in the network by all vehicles, i.e., $\bar{d} = \bar{d}_p$ and $\bar{t} = \bar{t}_p$. By substituting these equalities into Equations 3.1-3.3 and 3.5, the average flow, density, speed, and accumulation in the network can be estimated as follows:

$$\hat{q} = \frac{N_p \bar{d}_p}{\rho LT} \quad (3.6)$$

$$\hat{k} = \frac{N_p \bar{t}_p}{\rho LT} \quad (3.7)$$

$$\hat{\vartheta} = \frac{\bar{d}_p}{\bar{t}_p} \quad (3.8)$$

$$\hat{n} = \frac{N_p \bar{t}_p}{\rho T} \quad (3.9)$$

Similarly, the total number of probe vehicles that exit the network during an analysis period should be proportional to the total number of vehicles that exit the network, $N_p^e = N^e \rho$. By substituting this into Equation 3.4, we obtain the estimate of the average exit flow:

$$\hat{f} = \frac{N_p^e}{\rho T} \quad (3.10)$$

Equations 3.6-3.10 rely on the quantities of $\bar{d}_p$, $\bar{t}_p$, $N_p$, and $N_p^e$, all of which are directly measureable in real-time from the trajectories of probe vehicles. Actually, detailed vehicle trajectories are not necessary. Instead, Gayah and Dixit (2013) suggest
these quantities can be obtained if probe vehicles report odometer readings at discrete time periods that represent the beginning and end of all analysis intervals and the odometer reading and time when the probe vehicles enter and exit the network. Therefore, if the probe penetration rate is known, we can estimate the metrics of interest, the MFD, and the NEF using only data from probe vehicles and Equations 3.6-3.10.

3.3. Estimating the probe penetration rate

The estimation methods presented in the previous section rely on a priori knowledge of the probe penetration rate. However, this value may not be known a priori and will most likely vary across time and space. For example, if fleet vehicles such as taxis are used as probe vehicles, a larger proportion of taxis may be necessary during peak periods than off-peak periods. Thus, the proportion of traffic that the probe vehicles represent should be expected to change significantly over time. Fortunately, it may be possible for analysts to combine data from probe vehicles and traditional fixed sensors to estimate the fraction of probe vehicles during any particular analysis period, as in Herrera et al. (2010).

To estimate $\rho$ using this method, detectors can be placed at locations throughout the network to count the total number of vehicles, $N^d$, that cross all detectors during an analysis period. Simultaneously, virtual detectors can be coded into the GPS devices where the physical detectors are located, and the number of probe vehicles crossing the detectors, $N^d_p$, can be counted during the same analysis period. The probe penetration rate during a specific time interval can be estimated by:
Although loop detector estimates near signalized intersections may be unreliable for certain metrics such as density or occupancy, vehicle counts are more reliable as long as the analysis period is greater than one cycle length. This provides a simple method to calculate $\rho$ using only data from probe vehicles and existing loop detector data.

3.4. Accuracy of the estimates

The uncertainties of the estimated network-wide metrics of interest are examined in this section. It will be shown that the uncertainties can often be estimated well using only probe data. First, the accuracy of the probe penetration rate is derived, and then formulae are derived that quantify the uncertainties of the estimated average network flow, density, speed, accumulation, and exit flow.

3.4.1. Accuracy of the probe penetration rate

First, consider the accuracy of the estimate of $\rho$. If a total of $N^d$ vehicles travel over loop detectors during some analysis period, it is expected that an average of $N^d_p = \rho N^d$ probe vehicles would travel over the loop detectors during the same analysis period. However, the number of probe vehicles that cross the detectors would have some variation due to randomness. By treating $N^d_p$ as a binomial random variable, where each vehicle crossing a detector has the same probability, $\rho$, of being a probe vehicle, it is found that $\text{var}(N^d_p) = \rho(1 - \rho)N^d$. Therefore, the variance of the estimator $\hat{\rho}$ is:

$$\hat{\rho} = \frac{N^d_p}{N^d}. \quad (3.11)$$
\[ \text{var}(\hat{p}) = \rho(1 - \rho)/N^d. \]  

(3.12)

Note that as \( N^d \) increases, Equation 3.12 tends toward zero. Thus, the accuracy of the estimation increases as more vehicles cross the detectors during an analysis period.

Two methods can achieve more accurate probe penetration rate estimates: increasing the number of detectors in the network, or more realistically, increasing the length of the analysis interval. Interestingly, Equation 3.12 can be used to determine the necessary number of sampled vehicles to attain a desired level of accuracy by assuming the worst case, \( \rho = 0.5 \), that produces the largest variance. For example, if an agency requires the estimate of \( \rho \) to be within 0.5\%, the detectors must sample 10,000 vehicles during an analysis period. Since the number of detectors in a network is typically known (in addition to the minimum expected flow at each), the length of the analysis period needed to achieve this level of accuracy can be calculated.

### 3.4.2. Accuracy of the metrics of interest

The accuracy of the average flow, density, speed, accumulation, and exit flow are now considered. To determine the uncertainty of the estimations, the variance of the estimated metrics of interest (Equations 3.6-3.10) are derived, e.g., \( \text{var}(\hat{q}) = \text{var}(N_{p\bar{d}}\rho/\rho LT) \). When calculating the variance for the average estimated flow, density, speed, accumulation, and exit flow, the following variables were considered to be random: the number of probe vehicles, average distance traveled, average time spent by each vehicle in the network, and the probe penetration rate. If a large enough analysis period is selected and a sufficient number of fixed detectors exist, \( \rho \) can be estimated
quite accurately. Therefore, $\rho$ is treated as a constant, and the uncertainties of the variables are expressed as a function of $\rho$. Since the individual probe vehicles are selected from the set of all vehicles without replacement, 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. By taking the variance of Equations 3.6-3.10, the uncertainty of each estimate is found:

$$v\hat{a}(\hat{q}) = q^2 \text{var}(d_i)(1 - \rho)^2/N_p \hat{a}_p^2 + q \text{var}(d_i)(1 - \rho)/\hat{a}_p \rho LT$$  \hspace{1cm} (3.13)

$$v\hat{a}(\hat{k}) = k^2 \text{var}(t_i)(1 - \rho)^2/N_p \hat{\tau}_p^2 + k \text{var}(t_i)(1 - \rho)/\hat{\tau}_p \rho LT$$  \hspace{1cm} (3.14)

$$v\hat{a}(\hat{v}) = v^2 \text{var}(d_i)(1 - \rho)/N_p \hat{\alpha}_p^2 - 2v^2 \text{cov}(\hat{\alpha}_p, \hat{\tau}_p)(1 - \rho)/\hat{\alpha}_p \hat{\tau}_p$$  \hspace{1cm} (3.15)

$$v\hat{a}(\hat{\nu}) = n^2 \text{var}(t_i)(1 - \rho)^2/N_p \hat{\nu}_p^2 + n \text{var}(t_i)(1 - \rho)/\hat{\nu}_p \rho T$$  \hspace{1cm} (3.16)

$$v\hat{a}(\hat{f}) = f(1 - \rho)/\rho$$  \hspace{1cm} (3.17)

Note that Equation 3.15 holds by applying a first-order Taylor Series approximation to Equation 3.8 before taking the variance. Also, since the quantities $q$, $k$, $v$, $n$, and $f$ are generally unknown, the estimates provided by Equations 3.6-3.10 can be substituted into Equations 3.13-3.17.
The variances of $\hat{q}$, $\hat{k}$, $\hat{n}$, and $\hat{f}$ can be calculated in real-time using measurable data from individual probe vehicles. However, the variance of $\hat{\theta}$ (Equation 3.15) contains the term $\text{cov}(\tilde{d}_p, \tilde{t}_p)$, which cannot be calculated with probe data in real-time. However, since $\tilde{d}_p$ and $\tilde{t}_p$ should have a strong positive correlation, the covariance term should be positive—i.e., 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:

$$v\text{ar}(\hat{\theta}) \geq v^2 \text{var}(d_i)(1 - \rho)/N_p\tilde{d}_p^2 + v^2(1 - \rho)\text{var}(t_i)/N_p\tilde{t}_p^2,$$

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

### 3.4.3. Applications using fractional error

Since the variance of the metrics of interest can be derived using Equations 3.13-3.17, statistical methods may be used to determine the fraction of probe vehicles necessary to accurately estimate the average network flow, density, speed, accumulation, and exit flow in real-time within a desired fractional error. We know that the confidence interval surrounding the mean of variable $x$ is $\mu_x \pm z\sigma_x$ and that the fractional error is bounded by $\varepsilon = z\sigma_x/\mu_x$, where $z$ is the critical value of a normal distribution associated with the desired level of confidence. If an agency wishes to estimate the probe penetration rate necessary to estimate $x$ within some desired fractional error, $\varepsilon$, Equations 3.13-3.17 can be substituted into the error expression to find the minimum required value of $\rho$. For example, the probe penetration rate needed to ensure
that the estimate of average exit flow is within $\varepsilon$ of the true value with 95% confidence is:

$$\rho \geq z^2/(\varepsilon^2 f T + z^2).$$

(3.19)

Equation 3.19 confirms expectations that the minimum penetration rate increases with the level of significance and decreases with the required fractional error. Also, as the time period increases, the minimum penetration rate decreases. Thus, fewer probe vehicles are necessary if data are collected over a longer time period. This method can be repeated for all metrics of interest.
Chapter 4

Tests Using Micro-Simulation

Data from two micro-simulation models are used to test the accuracy of the estimation method presented in Chapter 3. First, the two simulations networks are described. Then, the simulation results are compared to the analytical equations developed in Chapter 3.

4.1. Idealized network

A 16x16 square grid network of alternating one-way streets was developed using AIMSUN micro-simulation software to test these methodologies in an ideal setting. The network consisted of 544 links, each with a length of 400 ft. Detectors were placed on the approach link to every intersection, totaling 512 detectors, to replicate networks with actuated traffic signal control. All intersection approaches operated with fixed signal timings with the following parameters: $CL = 60$ sec, $g_i = 26$ sec, $y_i = 3$ sec, and $AR_i = 1$ sec, where $CL$ is the cycle length, $g_i$ is the designated green time on approach $i$, $y_i$ is the designated yellow time on approach $i$, and $AR_i$ is the designated all-red time on approach $i$. The offset for every traffic signal in the network was set to zero. A portion of the idealized grid network is illustrated in Figure 4-1.

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 uniformly distributed for
simplicity, but the magnitude of traffic demands was adjusted to simulate a typical rush period of three hours, with a clearly defined one-hour peak period. Figure 4-2 illustrates the network demand as a function of time. Vehicles were randomly assigned as a probe vehicle as they entered the network with some fixed probability, $\rho$. The value of $\rho$ was held constant during each simulation, but due to randomness the actual probe penetration rate in the network varied across analysis periods.

Figure 4-1. A portion of the 16x16 idealized AIMSUN simulated network.

The probe vehicles produced sets of vehicle trajectory data for many hours over several days. The data sets included information that is easily attainable from existing GPS technologies, such as vehicle ID, time, geographic position, and speed at discrete
0.75-second time intervals. For the tests and plots presented in Chapter 4, a single 5-minute analysis period was randomly selected to simplify the presentation of results. During this period, \( q = 462.4 \text{ veh/hr}, k = 30.0 \text{ veh/mi}, v = 15.4 \text{ mi/hr}, n = 1,235.8 \text{ veh}, \) and \( f = 19,920 \text{ veh/hr}. \)

![Figure 4-2. Idealized network demand as a function of hour of day.](image)

### 4.2. Orlando network

Additionally, a micro-simulation model of downtown Orlando, FL, was used to test the methodologies on a more realistic network. The Orlando network provided a non-idealized structure with differing traffic signal parameters such as cycle lengths, green times, and offsets. Additionally, the Orlando network had varying geometric
characteristics such as roadway types with different speed limits, link lengths, link widths, and number of lanes. This network was used to verify the generalized results found on the idealized network and to illustrate how the estimation methodology can be used on large urban networks regardless of their traffic signal parameters, geometric characteristics, and size.

The network was developed using VISSIM, a microscopic simulation software program, as part of a study performed by the Florida Department of Transportation (Dixit et al., 2009). The model was constructed using traffic data collected by the City of Orlando and further calibrated and verified using mobile chase car data and parameters of the two-fluid model (Herman and Prigogine, 1979; Ardekani and Herman, 1987). The calibrated simulation network thereby replicated the microscopic properties of the downtown Orlando transportation network. The network covered a roughly 1.7 mile x 1.7 mile area that contained approximately 110 signalized intersections. The study area of the Orlando simulated network is illustrated in Figure 4-3. Note that only the surface streets in Figure 4-3 were included in the model, whereas the two freeways were excluded from the simulation. The simulation represented the AM peak period with a three-hour rush interval. A more detailed description of this model can be found in Dixit et al. (2009) and Dixit et al. (2011).
Figure 4-3. Downtown network of Orlando, FL (source: Gayah and Dixit, 2013).

Similar to the idealized network, the Orlando simulation network produced sets of vehicle trajectory data for many hours over several days. The data sets consisted of information that is attainable using GPS technologies such as vehicle ID, time, geographic position, and speed at a frequency of three seconds. Although it is preferable to collect vehicle data at a finer resolution than produced, it was difficult given data size limitations. As will be shown, this larger time resolution does not significantly impact the results. Additionally, as vehicles entered the Orlando network, vehicles were assigned as
either a probe vehicle or a non-probe vehicle. The summation of the two vehicle types at any instant in time represents the total number of vehicles in the network.

Furthermore, a single five-minute time interval was selected to illustrate the results in Chapter 4, in order to examine the accuracy of the methods developed in Chapter 3. During this time interval, \( q = 385.6 \text{ veh/hr} \), \( k = 36.2 \text{ veh/mi} \), and \( v = 10.7 \text{ mi/hr} \).

### 4.3. Accuracy of estimations

The methods developed in Chapter 3 are now tested on the idealized and Orlando micro-simulated networks. First, the accuracy of estimating the probe penetration rate by combining trajectory data and fixed loop detectors is examined. Then, critical assumptions made in Chapter 3 are verified. Next, the methods to estimate the average network flow, density, speed, accumulation, and exit flow are tested using the idealized grid network and the downtown Orlando network. Finally, statistical methods are used to determine the fraction of probe vehicles necessary to estimate the metrics of interest within a desired fractional error.

#### 4.3.1. Accuracy of probe penetration rate

To verify the accuracy and assumptions of Equations 3.11 and 3.12, estimates of the probe penetration rate and the variance of the estimated probe penetration rate were computed using the proposed methodology. For each run, the aggregate count of all
vehicles in the network, $N^d$, and probe vehicles in the network, $N^d_p$, that crossed all detectors were taken to calculate $\hat{\rho}$. This process was repeated many times to compute the variance of $\hat{\rho}$. The following trends were found: as the number of total vehicles counted, $N^d$, increased, the variance of the estimated probe penetration rate decreased, and as the sampling interval, $T$, increased, the variance of the estimate also decreased.

The estimated variance predicted the actual variance quite well, as illustrated in Figure 4-4, for a given time period within the idealized network. Additionally, the estimates were consistent and contained little variation. Equation 3.12 accurately estimates the observed variance, and the variances are very low for the entire range of probe penetration rates. In fact, the highest standard deviation for $\rho = 0.5$ was roughly 0.01. These results verify the accuracy of estimating $\rho$ and justify the treatment of $\rho$ as a constant.

![Figure 4-4. Theoretical and actual variance of the estimated probe penetration rate.](image-url)
4.3.2. Comparing probe vehicle data to all vehicle data

A critical assumption in the method to estimate the metrics of interest was that the probe vehicles accurately represent the average distance and time traveled by all vehicles ($\bar{d} = \bar{d}_p$ and $\bar{t} = \bar{t}_p$). The idealized and Orlando simulated networks were used to verify that the assumptions hold when all O-D pairs have the same fraction of probe vehicles. A range of $\rho = \{0.05, 0.10, 0.15, \ldots, 0.75\}$ was used to determine whether the two values are significantly different for varying levels of probe percentages. The values of $\bar{d}$, $\bar{d}_p$, $\bar{t}$, and $\bar{t}_p$ for both networks were calculated during the five-minute interval for many simulation runs, and one-sample t-tests were performed on the data sets with the following hypotheses:

<table>
<thead>
<tr>
<th>Test for distance</th>
<th>Test for time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{1,0}: \bar{d} - \bar{d}_p = 0$</td>
<td>$H_{2,0}: \bar{t} - \bar{t}_p = 0$</td>
</tr>
<tr>
<td>$H_{1,a}: \bar{d} - \bar{d}_p \neq 0$</td>
<td>$H_{2,a}: \bar{t} - \bar{t}_p \neq 0$</td>
</tr>
</tbody>
</table>

An alpha value of 0.05 was selected for statistical significance. Therefore, the null hypothesis is rejected if the p-value for any probe percentage is less than 0.05. The results of the t-tests are shown in Table 4-1. The p-values for all probe percentages ranging from 5% to 75% are greater than the alpha threshold of 0.05 for both the idealized and Orlando networks. Therefore, we fail to reject the null hypothesis for all probe percentages for the distance and time traveled in the networks. Thus, for all values of $\rho$, the average distance traveled and average time spent in the network were found to be statistically equivalent.
Table 4-1. Results of the statistical t-tests.

<table>
<thead>
<tr>
<th>Probe %</th>
<th>N</th>
<th>Idealized Network</th>
<th></th>
<th></th>
<th></th>
<th>Orlando Network</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>d, distance</td>
<td>t, time</td>
<td></td>
<td></td>
<td>d, distance</td>
<td>t, time</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T-stat</td>
<td>P-value</td>
<td>T-stat</td>
<td>P-value</td>
<td>T-stat</td>
<td>P-value</td>
<td>T-stat</td>
</tr>
<tr>
<td>5</td>
<td>132</td>
<td>-0.97</td>
<td>0.33</td>
<td>-1.08</td>
<td>0.28</td>
<td>105</td>
<td>0.80</td>
<td>0.43</td>
</tr>
<tr>
<td>10</td>
<td>325</td>
<td>-0.46</td>
<td>0.64</td>
<td>-0.23</td>
<td>0.82</td>
<td>272</td>
<td>1.26</td>
<td>0.21</td>
</tr>
<tr>
<td>15</td>
<td>408</td>
<td>-0.11</td>
<td>0.91</td>
<td>-0.21</td>
<td>0.84</td>
<td>336</td>
<td>0.07</td>
<td>0.94</td>
</tr>
<tr>
<td>20</td>
<td>586</td>
<td>-0.65</td>
<td>0.51</td>
<td>-0.58</td>
<td>0.56</td>
<td>495</td>
<td>0.37</td>
<td>0.71</td>
</tr>
<tr>
<td>25</td>
<td>715</td>
<td>1.50</td>
<td>0.13</td>
<td>1.41</td>
<td>0.16</td>
<td>619</td>
<td>0.05</td>
<td>0.96</td>
</tr>
<tr>
<td>30</td>
<td>857</td>
<td>-1.43</td>
<td>0.15</td>
<td>-1.34</td>
<td>0.18</td>
<td>723</td>
<td>-0.23</td>
<td>0.82</td>
</tr>
<tr>
<td>35</td>
<td>1,029</td>
<td>-0.91</td>
<td>0.36</td>
<td>-1.31</td>
<td>0.19</td>
<td>879</td>
<td>-1.18</td>
<td>0.24</td>
</tr>
<tr>
<td>40</td>
<td>1,160</td>
<td>0.36</td>
<td>0.72</td>
<td>-0.03</td>
<td>0.97</td>
<td>989</td>
<td>0.05</td>
<td>0.96</td>
</tr>
<tr>
<td>45</td>
<td>1,252</td>
<td>-0.37</td>
<td>0.71</td>
<td>-0.17</td>
<td>0.86</td>
<td>1,050</td>
<td>-0.24</td>
<td>0.81</td>
</tr>
<tr>
<td>50</td>
<td>1,451</td>
<td>-0.72</td>
<td>0.47</td>
<td>-0.90</td>
<td>0.37</td>
<td>1,235</td>
<td>-1.40</td>
<td>0.16</td>
</tr>
<tr>
<td>55</td>
<td>1,591</td>
<td>-1.14</td>
<td>0.25</td>
<td>-1.15</td>
<td>0.25</td>
<td>1,351</td>
<td>-1.25</td>
<td>0.21</td>
</tr>
<tr>
<td>60</td>
<td>1,717</td>
<td>0.62</td>
<td>0.54</td>
<td>0.67</td>
<td>0.50</td>
<td>1,477</td>
<td>0.66</td>
<td>0.51</td>
</tr>
<tr>
<td>65</td>
<td>1,866</td>
<td>-0.06</td>
<td>0.95</td>
<td>0.13</td>
<td>0.90</td>
<td>1,586</td>
<td>1.14</td>
<td>0.25</td>
</tr>
<tr>
<td>70</td>
<td>2,040</td>
<td>-0.47</td>
<td>0.64</td>
<td>-0.60</td>
<td>0.55</td>
<td>1,745</td>
<td>0.15</td>
<td>0.88</td>
</tr>
<tr>
<td>75</td>
<td>2,184</td>
<td>-0.26</td>
<td>0.79</td>
<td>-0.24</td>
<td>0.81</td>
<td>1,872</td>
<td>-0.48</td>
<td>0.63</td>
</tr>
<tr>
<td>100</td>
<td>2,870</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2,455</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.3.3. Estimating metrics of interest

Now, the accuracy of using Equations 3.6-3.10 to estimate traffic conditions are examined, and Equations 3.13-3.18 are verified to determine how well they predict the uncertainties of the estimates. Many simulation instances were performed for the range of $\rho = \{0.05, 0.10, 0.15, \ldots, 0.75\}$, and the estimates of $\hat{q}$, $\hat{k}$, $\hat{v}$, $\hat{n}$, and $\hat{f}$ were calculated for each instance for the idealized and Orlando networks. The actual values were simultaneously calculated using Equations 3.1-3.5 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 to measure the estimation accuracy.
4.3.3.1. Idealized network

Box plots of the results are provided in Figure 4-5. Values near 1.0 indicate that the estimated value is near the actual value. The bottom and top of the boxes represent the lower and upper quartiles, respectively, and the band in the middle denotes the median value. Dotted lines above and below the boxes represent data that are not regarded as outliers, whereas outliers are designated with a plus sign. A tolerance of ±10% of the actual value was drawn as a horizontal line on the box plot to further illustrate the accuracy. Note that a tolerance of ±3% was used for the plot of average speed (Figure 4-5c). As illustrated in Figure 4-5, the majority of non-outliers for average flow, density, accumulation, and exit flow are within 10% of the actual values for probe penetration rates of about 20%.

The individual estimates of speed are shown to be much more accurate, with estimates within 3% of the actual values when the probe penetration rate is roughly 10%. The higher accuracy of average speed can be explained by examining Equations 3.6-3.8. The method to estimate average network 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. However, in addition to the average distance and time, the average flow and density estimates also rely on the number of probe vehicles in the network. This introduces additional sources of error to the average flow and density estimates. Thus, the estimates for average flow and density should be less accurate than the estimates of average speed, as illustrated in Figure 4-5.
Figure 4-5. Box plots showing the accuracy of estimates for average: a) flow; b) density; c) speed; d) accumulation; and e) exit flow for the idealized network.
4.3.3.2. Orlando network

The method used in Section 4.3.3.1 was replicated using data from the Orlando network. The estimated and actual values of average network flow, density, and speed were calculated and the ratios between the estimated and the actual values were calculated and plotted in Figure 4-6.

Figure 4-6. Box plots showing the accuracy of estimates for average: a) flow; b) density; and c) speed for the Orlando network.
A horizontal line with a tolerance of ±10% was used to illustrate the accuracy of the flow and density estimations for the Orlando network, whereas a tolerance of ±3% was used for speed estimations. Similar to the idealized network, the majority of the individual flow and density estimates were within 10% of the actual values for probe percentages of 20%. Speed estimations are again more accurate, with the majority of the estimates falling within 3% of the actual values with probe percentages between 10-15%. The accumulation and exit flow are omitted from Figure 4-6 for brevity; however, it is clear that the estimation accuracies hold for both the idealized and Orlando networks.

4.3.4. Accuracy of variance estimations

The variances of the estimates are also calculated and plotted against the estimated variances from Equations 3.13-3.18, as shown in Figure 4-7 (idealized network) and Figure 4-8 (Orlando network).

4.3.4.1. Idealized network

As expected, the estimates become more accurate (i.e., the variance of the estimates decrease) as \( \rho \) increases. Additionally, the magnitudes of the variances are low—estimating all traffic variables from probe data using this methodology is very accurate when the probe penetration rate is just less than 20%. To examine this further, the 5-minute time interval selected for illustration has the highest flow and density observed, hence Equations 3.13-3.14 suggest this time interval should have the highest uncertainty.
in the estimates. With this worst case scenario, flow estimates are within ±47.5 veh/hr, and density estimates are within ±3.1 veh/mi, when ρ = 0.15 at the 96% confidence interval. If ρ increases from 0.15 to 0.30, the errors decrease to about ±30.1 veh/hr and ±2.0 veh/mi, respectively. Speed estimates are even more accurate. These results are consistent with the results shown in Figures 4-5 and 4-6.

In addition, confidence intervals for the theoretical variances are computed. If x represents one of the metrics considered here, the confidence interval for the true variance of x, \( \text{var}(x) \), is:

\[
(n - 1)\text{var}(\hat{x})/\chi^2_{1-(\alpha/2)} < \text{var}(x) < (n - 1)\text{var}(\hat{x})/\chi^2_{\alpha/2},
\]

where \( \chi^2_{\alpha/2} \) is the Chi-squared value associated with \( n - 1 \) degrees of freedom and a significance value of \( \alpha \). The 95% confidence intervals for the metrics of interest are illustrated in Figure 4-7 as the shaded region between the gray dashed lines. It is clearly shown that nearly all of the estimated variances using Equations 3.13-3.17 fall within the 95% confidence interval, which confirms the accuracy of the variance formulae. Thus, the theoretically derived Equations 3.13-3.17 provide a very accurate indication of the uncertainties of the traffic state estimations. Unfortunately, the upper bound for the accuracy of speed estimations is not tight. However, the magnitude of the upper bound is small, so the upper bound provides a useful indication of the accuracy of speed measurements.
Figure 4-7. Comparison of theoretical and actual variances for average: a) flow; b) density; c) speed; d) accumulation; and e) exit flow for the idealized network.
4.3.4.2. Orlando network

The variance of the estimated flow, density, and speed for a single time period are plotted and compared to the estimated variances of the Orlando network in Figure 4-8. As illustrated, as the penetration rate increases, the variances of the estimated metrics decrease. Note that the variances are again quite low. For a penetration rate of 15%, the variance of estimated flow implies that roughly 96% of the estimates (within 2 standard deviations) are within ±30 veh/hr. Additionally, at the same penetration rate, 96% of the density estimates are within ±1.7 veh/mi.

Figure 4-8. Comparison of theoretical and actual variances for average: a) flow; b) density; and c) speed for the Orlando network.
The 95% confidence intervals for the theoretical variances are also plotted in Figure 4-8 as the shaded gray region between the dotted lines. The actual variance for each metric of interest falls within the 95% confidence interval for all probe percentages considered. Again, the upper bound for the variance of speed is not tight; however, for a penetration rate of 5%, 96% of the estimates fall within ±2.3 mi/hr. Thus, the variance of speed can give an indication of the accuracy of the speed estimates. Furthermore, these results show that the estimation methodologies are accurate on both the idealized and realistic simulated networks.

4.3.5. Estimating the macroscopic fundamental diagram

Chapter 4.3.3 and 4.3.4 demonstrates that the methods developed in this paper accurately estimate average network flows and densities in real-time using limited trajectory data from probe vehicles and loop detectors. Now, this method can be applied to directly estimate a network’s MFD. These methods are tested on the idealized and Orlando networks to illustrate the potential for estimating the MFD in real-time using simulated data.

4.3.5.1. Idealized network

An entire microscopic simulation of the idealized network was run, and the average flows and densities were calculated at discrete 5-minute intervals to obtain the network’s actual MFD. This is illustrated by the dark solid lines in Figures 4-9. A
clockwise hysteresis loop is clearly visible in this MFD, similar to the shapes observed using simulated and empirical data (Buisson and Ladier, 2009; Geroliminis and Sun, 2011b; Gayah and Dixit, 2013). Furthermore, the estimation methodology was performed for many instances over a three-hour simulation period for $\rho = \{0.05, 0.10, 0.15, \ldots, 0.75\}$. The estimated MFDs produced by this method are illustrated by light gray lines in Figure 4-9. As expected, the MFD estimates converge upon the actual MFD as $\rho$ increases.
Figure 4-9. Actual and estimated MFDs for probe penetration rates of: a) 5%; b) 15%; c) 30%; d) 40%; and e) 50% for the idealized network.
4.3.5.2. **Orlando network**

An entire simulation of the Orlando network was also run, and the average flows and densities were calculated at discrete five-minute intervals to obtain the network’s actual MFD, which is plotted as the dark solid lines in Figure 4-10. A clockwise hysteresis loop is also present in the Orlando network’s MFD. The estimation methodology was used for several instances over the entire simulation period for $\rho = \{0.05, 0.10, 0.15, \ldots, 0.75\}$. The estimated MFDs are illustrated by light gray lines in Figure 4-10. As expected, and consistent with the idealized network, the estimated MFDs converge upon the actual MFD as the penetration rate increases. Next, the accuracy of the MFD estimations are quantified.
Figure 4-10. Actual and estimated MFDs for probe penetration rates of: a) 5%; b) 15%; c) 30%; d) 40%; and e) 50% for the Orlando network.
4.3.6. Quantifying the accuracy of MFD estimations

Three metrics are used to quantify the fit of the MFD estimates to the true MFD. The first two are the root mean square error (RMSE) of the average flow and density estimates for each five-minute interval. The RMSE of average flow and density indicates how well the estimated MFD fits the actual flow and density of the network. The RMSE of average flow and density were calculated by

\[ RMSE(q) = \sqrt{\frac{\sum (\hat{q} - q)^2}{N}}, \]

\[ RMSE(k) = \sqrt{\frac{\sum (\hat{k} - k)^2}{N}}, \]

respectively, where \( N \) is the number of time intervals during the simulation.

Additionally, a third metric is proposed that quantifies how well the average flow and density are estimated simultaneously. To calculate this metric, the values of flow and density were converted to dimensionless values by dividing each by the network capacity, \( q_c \), and jam density, \( k_j \), respectively. The network capacity and jam density were calibrated using simulation and found to be \( q_c = 900 \text{ veh/hr} \) and \( k_j = 320 \text{ veh/mi} \). The combined RMSE was calculated as the square of the distance between the actual and estimated scaled flow and density points on the MFD as such:

\[ RMSE(q, k) = \sqrt{\frac{\sum [(\hat{q}/q_c - q/q_c)^2 + (\hat{k}/k_j - k/k_j)^2]}{N}}. \] (4.2)

The RMSEs of both networks were calculated and presented in Table 4-2 for each five-minute time interval for the same range of probe penetration rates.
Table 4-2. Root mean square error of flow, density, and simultaneous flow and density for various probe penetration rates.

<table>
<thead>
<tr>
<th>Probe %</th>
<th>Idealized Network</th>
<th>Orlando Network</th>
<th>Idealized Network</th>
<th>Orlando Network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE(q) (veh/hr)</td>
<td>RMSE(k) (veh/mi)</td>
<td>RMSE(q,k) (veh/hr)</td>
<td>RMSE(k) (veh/mi)</td>
</tr>
<tr>
<td>5</td>
<td>33.66</td>
<td>2.02</td>
<td>3.79E-02</td>
<td>21.00</td>
</tr>
<tr>
<td>10</td>
<td>23.09</td>
<td>1.39</td>
<td>2.60E-02</td>
<td>14.85</td>
</tr>
<tr>
<td>15</td>
<td>18.58</td>
<td>1.12</td>
<td>2.09E-02</td>
<td>11.60</td>
</tr>
<tr>
<td>20</td>
<td>15.34</td>
<td>0.94</td>
<td>1.73E-02</td>
<td>9.79</td>
</tr>
<tr>
<td>25</td>
<td>13.41</td>
<td>0.82</td>
<td>1.51E-02</td>
<td>8.34</td>
</tr>
<tr>
<td>30</td>
<td>11.74</td>
<td>0.71</td>
<td>1.32E-02</td>
<td>7.40</td>
</tr>
<tr>
<td>35</td>
<td>10.24</td>
<td>0.62</td>
<td>1.15E-02</td>
<td>6.64</td>
</tr>
<tr>
<td>40</td>
<td>9.46</td>
<td>0.57</td>
<td>1.07E-02</td>
<td>6.13</td>
</tr>
<tr>
<td>45</td>
<td>8.59</td>
<td>0.52</td>
<td>9.68E-03</td>
<td>5.33</td>
</tr>
<tr>
<td>50</td>
<td>7.58</td>
<td>0.46</td>
<td>8.55E-03</td>
<td>4.82</td>
</tr>
<tr>
<td>55</td>
<td>7.02</td>
<td>0.42</td>
<td>7.91E-03</td>
<td>4.36</td>
</tr>
<tr>
<td>60</td>
<td>6.22</td>
<td>0.38</td>
<td>7.01E-03</td>
<td>3.95</td>
</tr>
<tr>
<td>65</td>
<td>5.67</td>
<td>0.35</td>
<td>6.39E-03</td>
<td>3.58</td>
</tr>
<tr>
<td>70</td>
<td>5.04</td>
<td>0.30</td>
<td>5.68E-03</td>
<td>3.11</td>
</tr>
<tr>
<td>75</td>
<td>4.49</td>
<td>0.27</td>
<td>5.06E-03</td>
<td>2.76</td>
</tr>
</tbody>
</table>

The analytical comparison shows the increase in accuracy as the probe penetration rate increases. Table 4-2 shows that the difference between the actual and estimated average flow and density is 33.7 veh/hr and 2.0 veh/mi on average for the idealized network, respectively, with a probe penetration rate of 5%. With an increased probe penetration rate of 15%, the error is greatly reduced, with an average difference between actual and estimated flow and density of 18.6 veh/hr and 1.1 veh/mi. Although the $RMSE(q, k)$ values are not meaningful on their own, they can be used to compare the accuracy of the MFD estimation for various values of $\rho$. This is illustrated in Figure 4-11 by dividing the simultaneous RMSE for all probe penetration rates by the RMSE of the
highest probe penetration rate of 75%. The simultaneous RMSE is a factor of 7.5 times larger for probe penetration rates of 5%, whereas the simultaneous RMSE is only 4.1 times higher for probe penetration rates of 15%. The accuracy of the idealized network MFD estimations increase rapidly until $\rho = 0.2$, where the gains begin to decrease.

![Graph showing the RMSE for idealized and Orlando networks](image)

Figure 4-11. Visual inspection of the increase in MFD accuracy as the probe penetration rate increases.

The estimations of the Orlando network were slightly more accurate, with a difference of 21.0 veh/hr and 1.6 veh/mi on average for the average flow and density estimates with a penetration rate of 5%. As the probe penetration rate increases to 15%, the error is reduced to 11.6 veh/hr and 0.9 veh/mi on average. Again, although the $RMSE(q,k)$ values of the Orlando network are not meaningful on their own, the MFD estimations of the Orlando network improve with an increased probe penetration rate, as illustrated in Figure 4-11. Overall, the results show that the methodology presented in this
paper is fairly accurate at estimating the MFD using limited mobile probe data for the idealized and realistic network.

4.3.7. Applying estimation methods using fractional error

To illustrate the practicality of Equation 4.2, the analysis periods presented in Figures 4-5 and 4-6 were considered. Here, we determine the necessary probe penetration rate to estimate the mean of any metric of interest within some desired fractional error. Thus, Equations 3.13-3.17 are substituted into the error expression \( \varepsilon = z \sigma_x / \mu_x \) to find the minimum required value of \( \rho \).

For the idealized network, with 95% confidence and a maximum fractional error of 10%, the minimum probe penetration rate needed is 17% for average flow, 18% for average density, 16% for accumulation, and 19% for exit flow. The average network speed requires a minimum penetration rate of approximately 10% for estimates within a maximum fractional error of 3%.

Similarly for the Orlando network, with 95% confidence and a maximum fractional error of 10%, the minimum probe penetration rate needed is 18% for average flow and 17% for average density. A minimum penetration rate of approximately 11% is needed for the average network speed for estimates to be within a maximum fractional error of 3%. These results are consistent with the results of Figure 4-5 and 4-6, with the exception of speed.
4.4. General remarks

This chapter used probe vehicle trajectories from two simulated networks to examine the accuracy of the estimation methodology proposed in Chapter 3. The idealized network consisted of a 16x16 square grid with equal block lengths and homogeneous demands. In addition, a more realistic simulated network of Orlando, FL, was used to test the estimation methodology on a network with varying demand inputs and network structures. The simulation tests on both networks confirm that the methodology increases in accuracy with the number of circulating probe vehicles on the network. In general, estimates of flow, density, accumulation, and exit flow can be obtained within 10% of the true value when only 20% of the circulating vehicles are probes. Estimates of speed are even more accurate, with estimates within 3% of the true value when only 15% of the vehicles serve as probes. The necessary probe penetration rate to obtain accurate speed estimates are comparable to other studies that focus on estimating travel times or speeds along roadway segments. However, the other metrics of interest have more uncertainty, which require additional probe data.

The simulations also verify that the analytical formulae of estimation accuracy provide useful indications of the expected uncertainty. Network-wide estimations of average density have only been conducted recently by Gayah and Dixit (2013), who found that penetration rates of at least 7.5% produce accurate estimations. However, the method used by Gayah and Dixit (2013) relied on previous knowledge of the MFD and is only accurate when the network is congested. The methods presented in this work are accurate during the entire range of traffic states, including free-flow, congested, and
unstable periods that contain hysteresis phenomena. Therefore, the estimation procedure presented in this paper should be very robust and may be used to directly estimate a network’s MFD and NEF. These estimations should be more consistent and comparable across different networks as opposed to estimates that rely on loop detectors, since estimations are highly sensitive depending on the placement of detectors.
Chapter 5

Network-wide Perimeter Control Strategy

This chapter uses two methods to estimate an idealized network’s MFD in real-time to inform a network-wide perimeter metering strategy during a typical morning rush period. First, the general metering strategy is introduced. Then, the two methods for traffic state estimations, which are used to inform the metering strategy, are discussed. Lastly, the idealized simulated network is described and used to examine the impacts of the perimeter network metering strategy using the two methods.

5.1. Perimeter control strategy

Recent work has shown that network-wide control strategies (Geroliminis and Levinson, 2009; Daganzo et al., 2012; Gonzales and Daganzo, 2012; Knoop et al., 2012; Keyvan-Ekbatani et al., 2012 and 2013) may benefit urban transportation networks. The strategies operate under the premise that networks have an innate tendency towards gridlock when congested (Daganzo, 2007); therefore, the control strategies should mitigate these effects by ensuring the networks do not reach the congested state. This paper expands upon this premise by metering vehicle entry into the network at the periphery of a network’s city center if the average city center density approaches some critical density, $k_c$. Two estimation methods are used to inform the perimeter control strategy: the first method uses data from loop detectors, and the second method uses data
from mobile probe vehicles. The goal is to improve the efficiency of the network by reducing total travel time and delay as an urban network becomes oversaturated.

The network-wide metering strategy based on the network’s traffic state is illustrated in Figure 5-1. The process is fairly simple and relies primarily on the average network density, $k$. At the beginning of each time period, $t$, the average density from the previous time period, $t-1$, is calculated. If the average density exceeds the critical density threshold, $k_c$, the green time, $g_i$, is reduced by 5 seconds, and the red time, $r_i$, is increased by 5 seconds for intersection approaches with through flows entering the city center (referred to as the subnetwork). If the average density during the previous time period does not exceed $k_c$, the maximum green time and minimum red time is adopted for all intersection approaches in the network. Note that all non-peripheral intersections always operate under fixed timings, so vehicles within the subnetwork (and the outer portion of the network) are uninhibited by the peripheral metering.

![Figure 5-1. Network-wide metering strategy.](image-url)
This paper explores two methods that estimate the network’s traffic state, which are used to inform the metering strategy. The first method uses data from loop detectors placed on links within the subnetwork, and the second method relies on trajectory data from vehicles within the subnetwork. The methods to estimate the network’s traffic state are described in the following section.

5.2. Traffic state estimations

The network’s traffic state can be described by the network’s MFD, which consists of the average network flow, \( q \), and the average network density, \( k \). Before metering is enabled, an accurate understanding of the current traffic state is necessary. This paper examines a single perimeter control strategy that operates with a scheme that is based on two different network-wide traffic state estimation methods. The current state of practice, and the first method considered, uses loop detector measurements to calculate the average flow and density, since loop detector technologies are well known and have existed for several decades. The second strategy uses data from mobile probe vehicles, which are becoming increasingly available with the advances in GPS devices and mobile computing. The two traffic state estimation methods are now described in further detail.
5.2.1. Link method

To estimate the network’s MFD using link data, the average flow and density are defined as follows:

\[ q = \frac{\sum_i q_i T L_i}{L T}, \quad \text{and} \]
\[ k = \frac{\sum_i [o_i N_i L_i T / 100 \lambda]}{L T}, \]

where \( q_i \) [veh/hr] is the measured loop detector flow on link \( i \), \( L_i \) [mi] is the length of link \( i \), \( T \) [hr] is the length of the analysis period, \( o_i \) [%] is the measured time-occupancy on link \( i \), \( N_i \) [lanes] is the number of lanes on link \( i \), and \( \lambda \) [mi] is the average vehicle length. To obtain reasonably accurate density estimations, loop detectors must be placed near the middle of the link (Papageorgiou and Vigos, 2008). Note that although all of the variables in Equations 5.1 and 5.2 can be calculated in real-time, the values do not represent the true average flow and density in the network. Instead, they provide an estimate of the MFD if detectors are placed on all links within the networks’ boundaries. However, the true MFD of the network cannot be calculated even if detectors are placed on all links within the network, since Edie’s generalized definitions require complete vehicle trajectories within the network to characterize the MFD.

If a reduced number of links are equipped with detectors, Equations 5.1 and 5.2 may still be able to describe the state of the network, as found by Keyvan-Ekbatani et al. (2013). Equations 5.1 and 5.2 are still used, but the detectors only cover a subset of links in the network. In this work, we explore informing the network-wide control strategy based on full link estimations, where 100% of the links in the network are equipped with
detectors, and reduced link estimations, where only a subset of the links in the network are equipped with detectors.

5.2.2. Mobile probe method

A detailed description of MFD estimation procedure using mobile probe data is located in Chapter 3 of this paper. The validity of the estimation procedure was verified in Chapter 4; therefore, the mobile probe estimation method will also be used to inform the network-wide perimeter control strategy. Equations 3.1 and 3.2 will be used to calculate the actual flow and density of the network in real-time to inform the metering strategy based on full vehicle trajectory data. However, in reality, trajectory data from all vehicles may not be available. Thus, the estimated average flow and density (Equations 3.6 and 3.7) will be tested to estimate the network’s traffic state and inform the metering strategy in real-time.

The uncertainty of the average network density estimations must be included to account for the real-time estimation error when using the limited mobile probe method. As illustrated in Figure 5-2, the actual average density of the network falls within $\hat{\kappa} \pm zSD$. Thus, the standard deviation of the density estimation can be calculated in real-time by taking the square root of Equation 3.14. This accounts for situations in which the estimation of density falls below the critical density but actual density exceeds the critical threshold.
5.3. Simulated network

A modified version of the idealized AIMSUN network described in Chapter 4.1 was used to test the metering strategy. The 16x16 grid network of alternating one-way streets consisted of 544 links, each with a length of 400 ft. In addition, an 8x8 square grid subnetwork was identified at the center of the overall network to represent a typical city center within an urban area. A total of 512 detectors were placed on the approach link to every intersection in the overall network, with a total of 144 detectors within the subnetwork. At the start of the simulation, the signal timings at all 256 intersections operated with the following parameters: $CL = 60$ sec, $g_l = 26$ sec, $y_l = 3$ sec, and $AR_l = 1$ sec, and the offset for every traffic signal in the network was set to zero. The 16
intersections at the periphery of the subnetwork operated with adaptive signal timings based on the scheme presented in Figure 5-1, whereas the remaining intersections operated with fixed signal timings with the initial traffic signal parameters. The cycle length remained fixed throughout the entire simulation for all intersections within the network.

Origins and destinations were calibrated to replicate a typical morning rush period, with the majority of the trips originating in the outer network and destined for the subnetwork. The demand distribution is exhibited in Figure 5-3. A loading period exists for the first hour of the simulation, before the demand increases in 10 minute intervals until the demand peaks for 20 minutes. Following the peak period, the demand relaxes and stabilizes 3 hours into the simulation. Therefore, the morning peak is simulated for a total of 4 hours, with a clearly defined 20 minute peak period. The simulated network was developed such that the 8x8 subnetwork was the focus of the metering strategy.

![Figure 5-3. Metering strategy demand distribution.](image)
Before the metering strategy can be implemented, the critical density threshold must be determined. Since the basis of the metering strategy revolves around this threshold, it is pertinent that the critical density is accurately calibrated. Since the link and probe state density estimation methods differed, the critical density had to be determined separately in the following manner. Several simulations were run without implementing the metering strategy to determine the reproducible MFD of the network for each method. The average density immediately before the point where the MFD was no longer reproducible (in the critical density regime) was selected as the critical density. A critical density of 40 veh/mi was found for the probe metering strategy, whereas a critical density of 45 veh/mi was used for the link metering strategy. The differing critical densities may be a result of queue spillovers, which typically cause the link method to overestimate the true density. Since the critical densities have been identified, the simulation can be used to estimate the traffic state in real-time and implement the control strategy based on the link and probe estimation methods.

5.4. Results

The idealized simulated network was used to test the following metering scenarios: metering based on full (100%) link data, reduced (75%, 50%, 25%, and 5%) link data, full (100%) probe data, and limited (75%, 50%, 25%, and 5%) probe data. Many simulation instances were run for each metering scenario, and the results were compared to the non-metered scenario. For all of the metering scenarios, the average network flows and densities were calculated in real-time using the methods described in
Chapter 5.2.1 or Chapter 5.2.2, depending on the scenario, which triggered the metering strategy illustrated in Figure 5-1.

The travel time savings within the simulation period, compared to the non-metering case, was used for the measure of effectiveness. The travel time savings was calculated as follows and is illustrated in Figure 5-4. For every simulation, the cumulative trips that depart for their destination (departure trips) and cumulative trips that exit the network (arrival trips) were output in 60 second intervals, equal to one traffic signal cycle length. When the cumulative departure trips are plotted versus time, this curve represents the network’s departure curve, \( D(t) \). Similarly, by plotting the cumulative arrival trips versus time, the curve represents the network’s arrival curve, \( A(t) \). The area between \( D(t) \) and \( A(t) \) is the total travel time for all vehicles in the network. The network’s departure curve, represented by the blue curve in Figure 5-4, is independent of the metering strategy. However, the purpose of the metering strategy is to maximize the number of completed trips as a function of time. Therefore, we expect the arrival curve of the metered case to have a higher trip completion rate compared to the non-metered case, as illustrated by the green and red curves in Figure 5-4. The travel time savings is the area between metered and non-metered arrival curves.
5.4.1. Non-metered scenarios for link and probe methods

First, 40 simulation instances of the non-metered link and probe scenarios were run to serve as the baseline of comparison for the metered strategies. The signal timings for all intersections within the entire network operated with fixed timings throughout the entire simulation. A subset of the 40 non-metered MFDs is plotted in Figure 5-5.

Figure 5-5a shows the MFD of the subnetwork with average flow and density estimates calculated using the full link data, whereas average flow and density estimates in Figure 5-5b were calculated using full probe data. As illustrated in Figure 5-5a, an increased amount of scatter occurs when $k = 45$ veh/mi for the link approach, whereas a similar amount of scatter occurs when $k = 40$ veh/mi, which illustrates the selected critical density thresholds for the two different measurement methods.
As the network reaches the congested state, the network becomes increasingly unstable and reduces the subnetwork’s average flow, as witnessed in the scatter. Notice how the density estimations using the link method are consistently higher than the probe estimations, possibly due to queue spillbacks causing the link method to artificially

Figure 5-5. MFD of idealized network using a) link estimation method and b) probe estimation method.
inflate the estimates. Therefore, the metering strategy aims to reduce this scatter by limiting the subnetwork’s density to below the critical threshold. Note that the recovery time period was omitted from Figure 5-5 for visibility purposes.

5.4.2. Metered scenarios for link and probe methods

In this section, the results are presented for the reduced link and probe metering strategies. For the reduced link estimation method, the subset of detectors used to calculate the average network density was selected at random at the start of the instance. The number of detectors selected depended on the scenario – i.e., for the 75% reduced link strategy, 75% of the detectors were randomly selected within the subnetwork. When the MFD is estimated, the MFD represents the fundamental diagram of the links with equipped detectors. Similarly, for the limited probe strategy, vehicles that entered the network were randomly assigned as a probe vehicle with a fixed probability, $\rho$. The value of $\rho$ was held constant during each simulation instance, but the actual probe penetration rate varied across analysis periods due to randomness.

The results of the 40 simulation instances for each strategy are exhibited in Figure 5-6, with the total time savings plotted versus the percentage of data used in the simulation. The blue dots in Figure 5-6 represent the travel time savings of a single simulation instance, the red square represents the mean travel time savings across all of the simulation instances for each percentage, the blue plus sign is the 95% confidence interval of the mean, and the red plus sign is the 95% confidence interval of the standard deviation.
Figure 5-6. Travel time savings for a) link strategy and b) probe strategy for different reduced percentages.
The results of the simulations are summarized in Table 5-1. If 100% of the links are equipped with detectors and provide estimations to inform the metering strategy, 245.4 veh-hrs are saved from vehicles traveling within the subnetwork on average. The 100% probe metering strategy provides even more travel time savings, as expected, since full vehicle trajectories calculate the true MFD of the subnetwork. In addition, the probe metering strategy provides more consistent results, with the standard error of the mean deviating 38 veh-hr from the mean travel time savings.

Table 5-1. Travel time savings for the full and reduced metering strategies.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Travel Time Savings (veh-hr)</th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
<th>SE</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100% Metered</td>
<td></td>
<td>245.41</td>
<td>-912.00</td>
<td>1435.00</td>
<td>71.06</td>
<td>438.04</td>
</tr>
<tr>
<td>75% Metered</td>
<td></td>
<td>265.07</td>
<td>-852.65</td>
<td>1430.43</td>
<td>78.28</td>
<td>482.54</td>
</tr>
<tr>
<td>50% Metered</td>
<td></td>
<td>187.45</td>
<td>-853.50</td>
<td>1336.25</td>
<td>68.40</td>
<td>421.63</td>
</tr>
<tr>
<td>25% Metered</td>
<td></td>
<td>158.66</td>
<td>-973.78</td>
<td>1567.88</td>
<td>82.47</td>
<td>508.40</td>
</tr>
<tr>
<td>5% Metered</td>
<td></td>
<td>-169.99</td>
<td>-1227.20</td>
<td>1279.70</td>
<td>100.53</td>
<td>619.69</td>
</tr>
<tr>
<td>Probe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100% Metered</td>
<td></td>
<td>315.58</td>
<td>-134.63</td>
<td>796.75</td>
<td>37.95</td>
<td>233.96</td>
</tr>
<tr>
<td>75% Metered</td>
<td></td>
<td>262.04</td>
<td>-272.87</td>
<td>808.18</td>
<td>37.74</td>
<td>232.62</td>
</tr>
<tr>
<td>50% Metered</td>
<td></td>
<td>255.94</td>
<td>-342.17</td>
<td>621.97</td>
<td>36.77</td>
<td>226.69</td>
</tr>
<tr>
<td>25% Metered</td>
<td></td>
<td>175.96</td>
<td>-499.20</td>
<td>569.33</td>
<td>39.25</td>
<td>241.98</td>
</tr>
<tr>
<td>5% Metered</td>
<td></td>
<td>-66.27</td>
<td>-826.80</td>
<td>519.27</td>
<td>58.00</td>
<td>357.51</td>
</tr>
</tbody>
</table>

The reduced link strategy provides lower travel time savings than the limited probe strategy for all percentages considered, except for 75%, where the savings are very similar. There appears to be a decrease in travel time savings as the percentage of link coverage decreases, as expected. If 25% of the links in the subnetwork are equipped with detectors, the average travel time savings is 158.7 veh-hr. If 75% of the links are
equipped with detectors, 265.1 veh-hr of travel time is saved on average. Although the links with equipped detectors were randomly selected for the reduced link strategy, metering performed well. This is promising since previous studies selected the equipped links with prior knowledge of the network and selected links that have a higher probability of representing the overall network conditions. The standard deviation of the mean travel time illustrates this randomness, since the standard deviation of the reduced link strategies are higher when compared to the full link strategy. When a reduced number of links are randomly selected to describe the overall network, it is possible that the links are not homogeneously congested, which causes inconsistencies in the density estimations.

The limited probe metering strategy is illustrated to be beneficial to overall subnetwork performance, even more so than the link strategy, with an average travel time savings of 176.0 veh-hr if 25% of the vehicles in the network serve as probes. The results are as expected for the limited probe metering strategies. As the probe penetration rate increases, the uncertainty of the estimations decrease, and the accuracy of the estimations increase. Therefore, the total time savings increases as the probe penetration rate increases. A significant increase in travel time savings is found as the penetration rate increases from 5% to 25%. The standard deviations of the mean travel time savings for probe penetration show that the benefits of the probe strategy are more consistent than the link strategy, since the density estimations using probe vehicles provide a larger coverage of the network—i.e., the probe vehicles better describe the conditions of the subnetwork as a whole and do not rely on point measurements.
Note that for both the reduced link and limited probe strategies, metering the perimeter of the subnetwork increases the total travel time experienced by all vehicles for several simulation instances, as exhibited by the negative values in Figure 5-6 and Table 5-1. This typically occurs when an increased amount of red time is assigned to the approaches entering the subnetwork, which cause queues to form and propagate backwards. If the queues extend and block upstream intersections, the queues disrupt vehicle flow and increase the time it takes vehicles in the outer network to reach their destinations. The results show that the probe strategy has fewer instances where the total travel time savings were negative. Nearly 37% of the link metering strategy instances actually worsened network efficiency, whereas 24% of the probe metering instances cause the total travel time to increase.

In general, these results show that the reduced link and limited probe metering strategies are beneficial for improving the stability of urban networks as they become congested. However, the probe metering strategy is shown to greatly improve network performance and reduce total travel times compared to the link metering strategy for the reduced data percentages considered. With as little as 25% of the vehicles in the network serving as probes, metering vehicle entry into the network by way of perimeter traffic signal control greatly reduces the total travel time in the network.
Chapter 6
Conclusions

This study provided a general method to estimate network-wide traffic conditions using data from mobile probe vehicles. Real-time estimations of traffic metrics, such as average flow, density, speed, exit flow, and accumulation, were calculated using the developed method. The method requires that the fraction of vehicles serving as probes is known, which can be calculated by combining data from probe vehicles and fixed detectors. Analytical formulae were also developed, which can estimate the uncertainty of the estimations using data from the probe vehicles themselves. Thus, the estimations and their uncertainties can be calculated simultaneously in real-time.

A simulated idealized grid network was used to verify the accuracy of the estimations and their uncertainties, which confirmed that the method increased in accuracy with the fraction of mobile probe vehicles in the network. Accurate estimations of the metrics of interest were obtained within 10% of the true value when as little as 20% of the circulating vehicles served as probes. Average speed estimations were even more accurate, since accurate estimations could be obtained within 3% of the true value when 15% of the circulating vehicles served as probes. A realistic simulated network of the downtown Orlando, FL network confirmed these results and illustrated that the estimation methods could be used on networks with varying link lengths, traffic signal parameters, demands, and network size. The results appeared to be accurate during the entire range of traffic states, including free-flow, congested, and unstable periods that
experience hysteresis phenomena. Therefore, the estimation procedure can be used to estimate a network’s Macroscopic Fundamental Diagram (MFD) and Network Exit Function (NEF).

A simple perimeter network-wide metering strategy was developed based on the mobile probe vehicle estimation procedure. The metering strategy operated under the premise that, as urban networks become congested, metering vehicle entry into the network will delay the onset of oversaturation and improve the overall performance of the network. An idealized grid network was used with a well-defined subnetwork that replicated a typical morning peak period within a city center. Once the subnetwork reached a pre-defined critical density threshold, adaptive signal control at the periphery of the subnetwork metered vehicle entry at the maximum allowable rate.

The probe metering strategy was compared to current practice MFD estimations that use link-based estimation methods. The link and probe metering strategies were compared to the non-metering scenario to quantify the effects of the metering strategy. The results showed that metering saved 158.7 veh-hr of travel time on average when the number of detectors covered at least 25% of the links in the subnetwork. Similarly, if at least 25% of the vehicles circulating in the network served as probes to estimate the traffic state, the metering strategy provided 176.0 veh-hr of travel time savings. The limited probe strategy provided higher travel time savings compared to the reduced link strategy and also provided more consistent results across simulation instances due to higher spatial coverage by the probe vehicles, which may account for inhomogeneous spatial distributions.
In general, the results of this work illustrate that network-wide traffic states can be estimated quite well in real-time using GPS data from mobile probe vehicles. In addition, the uncertainty of the estimations can be determined in real-time from the probe vehicles themselves. The probe estimation method is shown to be very promising for informing network-wide metering at the perimeter to provide travel time savings and increase trip completion rates during peak intervals. However, the estimation process may also be used to inform adaptive signal control, dynamic vehicle routing, pricing strategies, and the allocation of street space to improve network stability and performance.

**Future Work**

Although the results of the probe estimation methodology are promising, several limitations exist. It was assumed that probe vehicles were uniformly distributed throughout the network; however, in reality, portions of the network may have higher (or lower) probe penetration rates, which result in heterogeneous probe distributions. Further work should be conducted that characterizes the heterogeneous spatial distribution and quantifies how they may influence the uncertainty of the estimations. Additionally, more work is needed to quantify how uncertainties in the probe penetration rate estimations affect the overall metric estimates. Furthermore, detector and probe data may be combined in other ways for traffic state estimations. For example, average network density can be estimated indirectly by using detectors to calculate network flows and probes to estimate the average speed.
Additionally, the perimeter metering strategy shows promise but consists of a few limitations that should be explored in future work. First, an adaptive traffic signal control on the perimeter that utilizes common practice algorithms should be explored, such as the systems in Gayah and Gao (2014). Thus, instead of reducing the green time on the approaches by a constant integer, an appropriate reduction in green time may be calculated based on the link density. Additionally, future work should explore radial metering strategies that not only focus on the immediate periphery of a subnetwork but also slowly meter vehicle entry in increasing rates the closer to the subnetwork boundary. Lastly, the metering strategies based on the methods presented in this paper should be tested on a more realistic simulated network.
References


Dixit, V.V., Crowe, J. and Radwan, E. Evaluation and microscopic simulation of the Orlando downtown network. Final report submitted to the City of Orlando, Department of Transportation, 2009.


Gayah, V.V. and Dixit, V.V. Using mobile probe data and the macroscopic fundamental diagram to estimate network densities: tests using micro-simulation. in: Proceedings of the 92nd Annual Meeting of the Transportation Research Board, January 13-17, 2013, Washington, D.C.

Gayah, V.V., Dixit, V.V., and Guler, I.S. Relationship between mean and day-to-day variation in travel time in urban networks. EURO Journal on Transportation and Logistics, 2013. (In Press)


