Data Driven Identification of Candidates for Operational Improvement

Final Report
October 2021

Center for Transportation Research and Education

IOWA STATE UNIVERSITY
Institute for Transportation

Sponsored by
Iowa Department of Transportation
(InTrans Project 19-716)

About the Center for Transportation Research and Education

The mission of the Center for Transportation Research and Education (CTRE) at Iowa State University is to conduct basic and applied transportation research to help our partners improve safety, facilitate traffic operations, and enhance the management of infrastructure assets.

About the Institute for Transportation

The mission of the Institute for Transportation (InTrans) at Iowa State University is to save lives and improve economic vitality through discovery, research innovation, outreach, and the implementation of bold ideas.

Iowa State University Nondiscrimination Statement

Iowa State University does not discriminate on the basis of race, color, age, ethnicity, religion, national origin, pregnancy, sexual orientation, gender identity, genetic information, sex, marital status, disability, or status as a US veteran. Inquiries regarding nondiscrimination policies may be directed to the Office of Equal Opportunity, 3410 Beardshear Hall, 515 Morrill Road, Ames, Iowa 50011, telephone: 515-294-7612, hotline: 515-294-1222, email: eooffice@iastate.edu.

Disclaimer Notice

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the sponsors.

The sponsors assume no liability for the contents or use of the information contained in this document. This report does not constitute a standard, specification, or regulation.

The sponsors do not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

Iowa DOT Statements

Federal and state laws prohibit employment and/or public accommodation discrimination on the basis of age, color, creed, disability, gender identity, national origin, pregnancy, race, religion, sex, sexual orientation or veteran's status. If you believe you have been discriminated against, please contact the Iowa Civil Rights Commission at 800-457-4416 or the Iowa Department of Transportation affirmative action officer. If you need accommodations because of a disability to access the Iowa Department of Transportation's services, contact the agency's affirmative action officer at 800-262-0003.

The preparation of this report was financed in part through funds provided by the Iowa Department of Transportation through its "Second Revised Agreement for the Management of Research Conducted by Iowa State University for the Iowa Department of Transportation" and its amendments.

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Iowa Department of Transportation.

Technical Report Documentation Page

1. Report No.	**2. Government Accession No.**	**3. Recipient's Catalog No.**	
InTrans Project 19-716			
4. Title and Subtitle Data Driven Identification of Candidates for Operational Improvement		**5. Report Date** October 2021	
		6. Performing Organization Code	
7. Author(s) Christopher M. Day (orcid.org/0000-0002-3536-7211), Tahsin Emtenan (orcid.org/0000-0002-7001-6724), Shoaib Mahmud (orcid.org/ 0000-0002-3514-2055), and Skylar Knickerbocker (orcid.org/0000-0002-0202-5872)		**8. Performing Organization Report No.** InTrans Project 19-716	
9. Performing Organization Name and Address Center for Transportation Research and Education Iowa State University 2711 South Loop Drive, Suite 4700 Ames, IA 50010-8664		**10. Work Unit No. (TRAIS)**	
		11. Contract or Grant No.	
12. Sponsoring Organization Name and Address Iowa Department of Transportation 800 Lincoln Way Ames, IA 50010		**13. Type of Report and Period Covered** Final Report	
		14. Sponsoring Agency Code 19-716 TSIP	
15. Supplementary Notes Visit https://intrans.iastate.edu/ for color pdfs of this and other research reports.			
17. Key Words ATSPMs—automated traffic signal performance measures—probe vehicle data—operational improvement—travel-time reliability		**18. Distribution Statement** No restrictions.	
19. Security Classification (of this report) Unclassified.	**20. Security Classification (of this page)** Unclassified.	**21. No. of Pages** 74	**22. Price** NA

16. Abstract

In Iowa, the safety improvement candidate list has been used for about 20 years to identify roadway locations having disproportionate crash rates or severities. Recent federal rulemaking on the travel-time reliability measurement requires a push toward greater accountability for operational performance. In response to this guidance, the present study explores the development of an operational improvement candidate list (OICL) by using data that were available at the time of the study.

A survey of the available data is presented. Two data sets were selected on the basis of their availability to facilitate quantitative research during the study period: automated traffic signal performance measures (ATSPMs) and probe vehicle data. Performance measures for corridor ranking were developed independently from the ATSPM and probe vehicle data, and the two data sets were ultimately combined to develop a methodology for creating an OICL. The methodology was confirmed using the Cedar Rapids, Iowa area as a case study. In addition, data from Dubuque, Iowa were used to compare the ATSPM and probe vehicle data for corridor performance measures, which found a correlation between average segment speeds and intersection performance measures. The report concludes with a discussion of the scalability of the method and potential future research.

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

Data Driven Identification of Candidates for Operational Improvement

Final Report
October 2021

Principal Investigator
Christopher M. Day, Affiliate Researcher
Center for Transportation Research and Education, Iowa State University

Co-Principal Investigator
Skylar Knickerbocker, Research Engineer
Center for Transportation Research and Education, Iowa State University

Research Assistants
Shoaib Mahmud and Tahsin Emtenan

Authors
Christopher M. Day, Tahsin Emtenan, Shoaib Mahmud, and Skylar Knickerbocker

Sponsored by
Iowa Department of Transportation

Preparation of this report was financed in part
through funds provided by the Iowa Department of Transportation
through its Research Management Agreement with the
Institute for Transportation
(InTrans Project 19-716)

A report from
Institute for Transportation
Iowa State University
2711 South Loop Drive, Suite 4700
Ames, IA 50010-8664
Phone: 515-294-8103 / Fax: 515-294-0467
www.intrans.iastate.edu

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The authors would like to thank the Iowa Department of Transportation for sponsoring this research. The authors would also like to acknowledge the technical advisory committee members for the input they provided over the course of this project.

1. INTRODUCTION

1.1 Motivation for the Study

The allocation of resources for highway improvements is an essential activity for transportation agencies. Numerous factors go into the decision-making process. One of these factors is the quality of operations on highway facilities: for example, whether adequate capacity exists on various parts of the highway network in light of current demands. This is a separate consideration from making improvements to address pavement deterioration or safety concerns. The Highway Capacity Manual (HCM) has for many years been one of the most commonly used methodologies for determining level of service (LOS) for sections of highway, using mainly traffic counts as its input data in combination with information about roadway geometric characteristics. The HCM process is often undertaken when planning a new facility, and although in theory the methodology could be used to estimate LOS across an agency, this is not done in practice. Rather, increases in traffic volume in combination with anecdotal knowledge or public reports of congestion tend to draw attention to certain facilities in need of operational improvement.

In Iowa, the safety improvement candidate list (SICL) has been used for about 20 years to identify roadway locations that have disproportionate numbers of crashes or crash severity (Hallmark et al. 2002). The creation of that list helps fulfill the Iowa Department of Transportation (DOT) mission to improve highway safety as well as meet a federal requirement to identify locations that have high crash rates. Until relatively recently, an analogous federal requirement to evaluate the operational characteristics of roadways did not exist. However, in 2012, the passage of the Moving Ahead for Progress in the 21st Century Act (MAP-21) introduced provisions for performance evaluations. The Federal Highway Administration (FHWA) established a series of performance measures that were first published in a 2014 proposed rule that obliged states to calculate specific performance measures using the National Performance Measures Research Data Set (NPMRDS), a set of five-minute average speeds obtained from a private sector vendor and provided to state DOTs (FHWA 2014). At the same time, the FHWA programs have increasingly emphasized performance-based management (Day et al. 2020, FHWA 2021).

In light of the increasing emphasis on performance-based management driven by federal initiatives, as well as to better allocate scarce resources to locations having the greatest need, this research seeks to develop an operational improvement candidate list (OICL), analogous to the SICL, using data available for creation of such a list. Given that the Iowa DOT already has mature practices for monitoring interstate highway mobility (Iowa DOT 2016), this research focuses on the evaluation of signalized, non-limited access highways. This report contains a review of potential data sources. Two data sources were selected based on their availability and applicability toward development of an OICL. One of these was probe vehicle segment speed data, which the Iowa DOT already procures primarily to support monitoring interstate performance but which also extends to non-freeway facilities. The other was high-resolution data used to support automated traffic signal performance measures (ATSPMs), which has emerged in recent years as a means of obtaining detailed operational data from signalized intersections.

1.2 Organization of the Report

This report has six chapters. Chapters 2–5 contain most of the technical content of the report and are organized as follows:

- Chapter 2 presents a review of existing data sets, which reviews legacy data sets, currently available data, and emerging data sets. The chapter provides some description and examples of segment speed data and high-resolution data, which were adopted for use in later analysis tasks.
- Chapter 3 uses segment speed data to conduct a performance comparison of 250 signalized corridors across the state of Iowa. Their performance in 2019 and 2020 were compared, and the corridors were ranked according to a performance index (PI) based on travel time and travel-time reliability.
- Chapter 4 uses high-resolution data for two studies. In the first study, data from 150 intersections in the Cedar Rapids, Iowa area were used to rank intersections according to a few different performance measures that investigate the quality of capacity (or “green time”) allocation at those intersections. In the second study, corridor segment speed data were compared with measurements of percent on green (POG) from high-resolution data at a few intersections in the Dubuque, Iowa area to determine whether the outcomes of the two data sets correlate. Such a comparison has not been done before.
- Chapter 5 presents the development of an OICL for the case study of Cedar Rapids, where there is extensive coverage of both the segment speed data and the high-resolution data. The corridor and intersection metrics presented in the previous chapters were combined to yield a composite metric that takes both progression and capacity utilization into account. This was used as the basis for developing an OICL for 21 corridors in the Cedar Rapids area.

1.3 Summary of Findings

The main findings of this study are summarized as follows:

- The study demonstrated the feasibility of using a combination of segment speed data and high-resolution data to establish an OICL. A preliminary analysis was undertaken to develop an OICL for corridors in the Cedar Rapids area.
- A ranking of 250 signalized corridors across the state was carried out using probe vehicle segment speed data.
- A ranking of 150 signalized intersections in Cedar Rapids was carried out using high-resolution controller event data.
- The first study directly comparing signal performance measures from high-resolution data (specifically the POG and volume-to-capacity [v/c] ratio) with segment speed data was carried out, finding that the two data sets exhibit correlation when models are adjusted by day-of-week and time-of-day variables.

2. REVIEW OF DATA SOURCES

2.1 Introduction

This chapter examines different data sources that could be applied to the problem of developing an OICL. A wide range of available data is examined, including older data sources that have been traditionally used for various traffic management applications, data sets that have emerged in the past 15 years and are currently in or approaching common and widespread use, and emerging data sets that are anticipated or just now coming to market at the time of the study.

2.2 Conventional Traffic Monitoring

Before reviewing more recent automated data sets with the potential for system-wide, automated traffic monitoring, it is worthwhile to briefly discuss the recent state of the practice. Historically, the most widely used type of data for traffic monitoring has consisted of traffic count and classification data. Such data are usually obtained from limited time periods by temporary placement of automatic counters or by manual counting (especially at intersections where pedestrians are also counted). Given the limited time frame and long time periods between updates, such data are most useful for planning applications. Permanent counting stations are used by most agencies at selected locations to establish a continuous count. These data are used to establish adjustment factors for converting short-term counts elsewhere into annual average daily traffic (AADT). Automatic vehicle classification is also sometimes possible. A similar type of application is weigh-in-motion recorders, which record vehicle count and classification as well as vehicle weight, often for overweight enforcement as well as to estimate total pavement loading. These permanent installations could potentially be used to obtain real-time counts; however, they are typically only available at a small portion of the total system and are more commonly used on freeways. For traffic operations, count data are frequently used to track patterns in traffic flows over time and to act when changes are observed (FHWA 2016, Turner et al. 2010).

A limitation of this existing practice is the fact that short-term counts collected at periodic intervals are generally unable to capture the day-to-day variation in conditions. Furthermore, count data alone do not necessarily reveal the quality of service on a roadway facility. The HCM and other methodologies can be used to estimate LOS, but given the input data are mostly representative of the expected demand of a typical day created from adjusting the count of a specific day by adjustment factors, the resulting outputs are themselves also a rather broad estimation. The effects are likely greater for signalized facilities where representative counts would be used in addition to representative signal timing. The process of actuation can frequently result in green times that are considerably different from those used to define the timing plan. The HCM recommends the use of actuated green times for such cases, but these are often difficult to obtain.

For freeway facilities, speed sensors are frequently employed for monitoring purposes. Several different detector technologies are used, including inductive loops in a speed trap arrangement or radar. This permits the real-time monitoring of speeds throughout the network. Many urban areas have had such monitoring capabilities in place since the 1990s, or earlier. Point speeds are less useful on surface streets given the presence of traffic control, which induces delays at intersections, meaning there is no one point along any roadway segment that can capture a representative speed that relates to what a typical average speed might be along a segment containing stops, signals, or other traffic control devices. There is also the burden of installing and maintaining the system of detectors. For a high-traffic urban area, the costs of excessive congestion seem to have justified installation of such monitoring systems in the past, but these capabilities are most often limited to the most critical portions of the network.

This, then, is the starting point for the present research study. Over the past decade or so, several new data sets have entered into common usage. These have been employed for a variety of uses that will be discussed in the following sections, but they have not previously been used in tandem to establish a methodology for determining portions of the system most in need of operational improvement.

2.3 Crowdsourced Probe Vehicle Data

Probe vehicles are those for which travel is recorded by an observer. For many years, floating car studies were used to generate this type of data, originally with the use of a stopwatch and the vehicle odometer, and more recently with the assistance of Global Positioning System (GPS) devices. Floating car studies are labor intensive, given someone must be paid to drive the vehicles along the roadway, and several trips are needed to obtain a final travel time at a high confidence level (Quiroga and Bullock 1998). To try to avoid bias, drivers are often instructed to pass as many vehicles as they are passed by when executing a floating car run, or to follow the average speed of the traffic around them according to their judgment. These studies can potentially be biased by driver judgment or driving characteristics (Turner et al. 1998).

Over the years, various methods have emerged to obtain naturalistic probe vehicle data through observations of road users. This shifting of the core activity in probe vehicle data collection from the agency to the public has been called crowdsourcing. In the past, license plate studies were used for this purpose, in which observers recorded license plates and the times of observation at multiple locations and calculated the time between matches to obtain travel time. More recently, the proliferation of smartphones, navigation devices, and other such equipment among the public has made it possible to obtain similar identifiers through wireless communication, such as Bluetooth MAC addresses (Wasson et al. 2008). Data that use a vehicle identifier are called automatic vehicle identification (AVI) data. Such data are now in rather widespread use, with many vendors selling equipment that can collect a variety of identifiers and report travel-time information to the end user. A limitation of these data types is that they require the installation and maintenance of equipment in the field to obtain it.

The ubiquity of GPS-equipped smartphones has also opened up a separate avenue of collecting data on vehicle movement, by continuously tracking device positions over time. Such data are called automatic vehicle location (AVL) data. As the vehicle moves through the network, an onboard device obtains its location from its GPS feature and records it with a timestamp. Many smartphone applications harvest this information in exchange for service provided to the user.

Some of these applications may immediately retrieve the location, whereas others may store it and wait to do so when it is convenient. In either case, position data of many vehicles are collected in a repository. These AVL data are extremely common today and are used to support many functions such as the traffic status layer supported in the maps available in most smartphones or on websites such as Google Maps.

Figure 1 provides an example of raw AVL data.

Figure 1. Vehicle records from a sample AVL data set

Every dot on the map is the location of one timestamped record of a vehicle position. The traces of the individual vehicles are not visible, because the dots are not connected, but it is clear that the dots tend to follow the road network. As the map shows, the location of the observations is sometimes imprecise. A map-matching process is required to snap the locations to roadways and establish the vehicle path.

Rather than using raw AVL data directly, most agencies instead use derivatives of these data. One common type of data is probe vehicle average segment speeds, which consists of the mean speed on a predefined segment within regular time intervals. The NPMRDS is an example probe vehicle data set of this type, consisting of average speeds recorded once for every five-minute interval for a set of predefined roadway segments. The segmentation scheme used for the NPMRDS is the traffic message channel (TMC) segments, which are generally rather long segments spanning, for example, from interchange to interchange on freeways. Data providers also supply such data commercially at a higher spatial and temporal resolution. The Iowa DOT, for example, purchases data from INRIX that have a time resolution of one minute, and segments using a different definition that has a better spatial resolution. These are called XD segments (a brand name) and are intended to be about 1 mile in length. While the longer TMC segments may span several miles between interchanges, the shorter and more consistent length XD segments would likely break up the same distance across several segments.

An example of average speed data is presented in Figure 2.

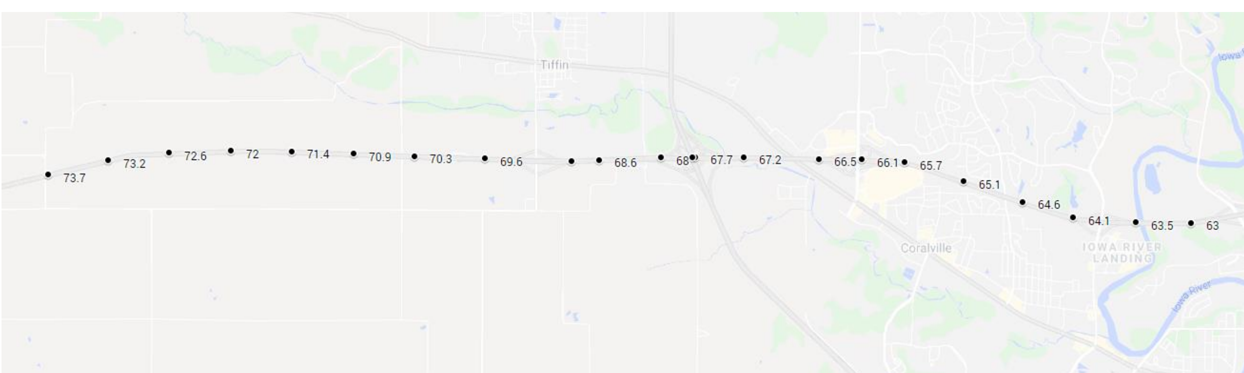
Time

Distance	4:00	4:01	4:02	4:03	4:04	4:05	4:06	4:07	4:08	4:09	4:10	4:11	4:12	4:13	4:14	4:15	4:16	4:17	4:18	4:19	4:20	4:21	4:22	4:23	4:24	4:25	4:26	4:27	4:28	4:29	4:30
63.0	66	65	65	65	65	65	64	64	64	64	65	65	65	65	65	65	68	69	70	70	70	70	70	70	69	69	69	69	68	68	68
63.5	65	65	65	65	65	65	65	65	65	65	65	65	65	63	63	63	64	66	66	68	68	68	68	69	69	69	68	68	68	68	67
64.1	66	66	66	66	66	66	66	65	65	65	65	65	66	66	66	66	66	65	65	65	65	65	67	67	69	69	69	69	67	67	67
64.6	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	66	67	68	68	68	68	67	67	66
65.1	68	67	67	67	67	66	65	65	65	65	65	65	66	66	66	66	65	64	64	64	64	64	65	65	65	65	65	66	66	66	66
65.7	67	67	67	67	67	65	65	65	65	65	65	65	65	66	66	65	65	63	63	63	63	63	63	63	64	62	62	62	61	61	61
66.1	66	66	66	66	67	63	63	63	65	65	65	65	65	66	66	66	65	65	63	63	63	61	61	60	61	60	61	61	58	58	58
66.5	67	66	66	66	67	64	65	65	66	66	66	66	63	66	65	66	65	63	62	62	62	62	62	62	63	63	62	62	61	61	61
67.2	66	66	66	66	66	66	66	66	67	67	67	67	64	67	65	65	65	62	63	63	63	61	62	62	61	62	62	61	61	61	61
67.7	65	65	65	65	65	65	65	65	67	67	64	64	62	59	57	58	58	58	58	58	57	57	57	57	55	55	55	56	60	60	60
67.7	65	65	65	65	65	65	65	65	67	67	62	60	58	57	55	55	55	59	59	58	58	58	58	58	57	57	57	57	58	58	58
68.0	60	60	60	60	60	60	60	60	62	62	58	57	52	50	48	42	42	40	40	40	42	42	42	42	43	43	44	45	47	47	48
68.6	60	60	60	60	60	60	60	60	63	63	58	57	52	52	46	46	46	45	45	45	47	47	47	47	50	49	49	51	52	50	50
68.8	57	57	56	57	56	56	56	55	56	55	55	55	52	52	52	51	50	48	48	48	47	47	47	47	47	47	48	48	49	49	49
69.6	52	52	52	52	52	51	50	50	50	50	50	50	48	48	48	48	48	49	50	50	50	50	50	50	48	48	48	48	50	48	49
70.3	52	52	52	52	52	52	49	49	49	49	49	49	47	47	47	47	47	47	48	48	50	49	49	49	48	48	48	48	48	48	48
70.9	52	52	52	52	52	52	50	50	50	50	50	49	47	48	47	47	47	48	48	48	48	48	48	48	47	47	47	47	48	48	47
71.4	52	52	52	52	52	52	51	49	49	49	49	49	48	48	48	48	48	48	48	50	50	50	50	50	48	48	48	48	48	48	48
72.0	47	47	47	47	47	48	47	47	42	42	42	42	41	41	43	43	44	44	44	45	45	45	45	45	45	46	46	47	46	46	45
72.6	45	45	45	44	45	45	46	45	41	40	40	39	38	37	39	39	40	40	39	39	40	40	40	40	42	40	40	40	41	41	40
73.2	54	52	52	52	52	52	52	52	50	50	50	50	49	47	48	48	49	49	48	47	47	46	46	46	46	45	45	45	45	45	45

Figure 2. Example average speed (mph) per segment

The data shows westbound travel through a work zone on I-80 near Coralville, Iowa during a Wednesday afternoon from 4:00–4:30 p.m. on October 7, 2020. The data shows evidence of slowed traffic on segments labeled 68.0 and 72.6, which are referenced to the start of I-80 on the eastern side of the state. The locations of the segments are shown in Base map image © 2021 Google Maps

Figure 3, which displays the locations of the segment endpoints.

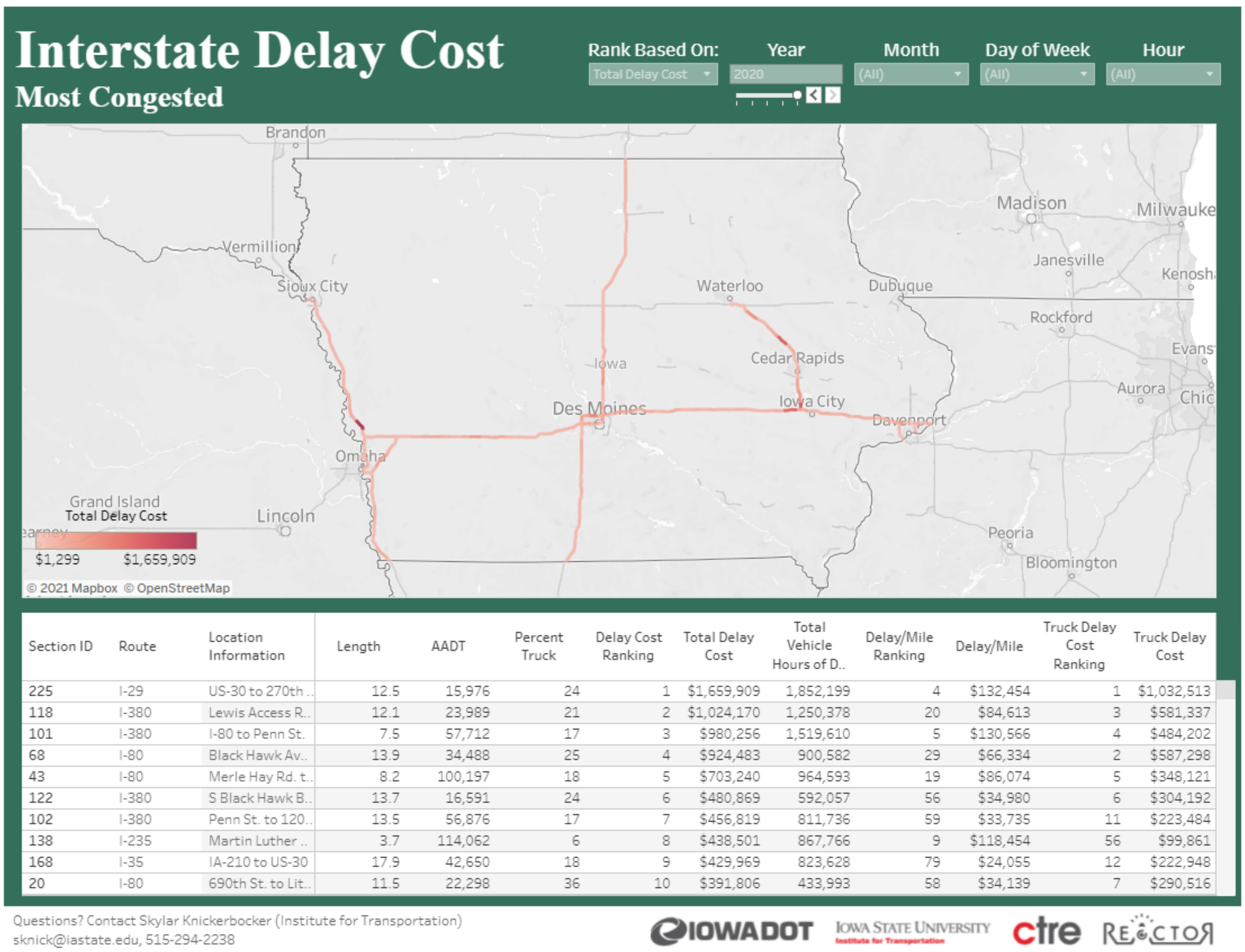


Base map image © 2021 Google Maps

Figure 3. Locations of segment endpoints

For example, segment 68.0 starts at the point labeled 68 in the map, and continues to 68.6, which shows that this segment is located right past the interchange of I-80 and I-380. The combination of westbound traffic on I-80 and merging traffic from I-380 is likely related to the speed reduction here. Meanwhile, segment 72.6 is on an isolated segment on the left side of the map, and the speed reduction is likely due to work zone activities.

These types of data have some utility for developing these relatively detailed views but are also useful for monitoring roadway performance at a high level. In Iowa and several other states, the data are used to create reports of roadway performance. An example is the interstate delay cost report shown in Figure 4.



Section ID	Route	Location Information	Length	AADT	Percent Truck	Delay Cost Ranking	Total Delay Cost	Total Vehicle Hours of D..	Delay/Mile Ranking	Delay/Mile	Truck Delay Cost Ranking	Truck Delay Cost
225	I-29	US-30 to 270th ..	12.5	15,976	24	1	$1,659,909	1,852,199	4	$132,454	1	$1,032,513
118	I-380	Lewis Access R..	12.1	23,989	21	2	$1,024,170	1,250,378	20	$84,613	3	$581,337
101	I-380	I-80 to Penn St.	7.5	57,712	17	3	$980,256	1,519,610	5	$130,566	4	$484,202
68	I-80	Black Hawk Av..	13.9	34,488	25	4	$924,483	900,582	29	$66,334	2	$587,298
43	I-80	Merle Hay Rd. t..	8.2	100,197	18	5	$703,240	964,593	19	$86,074	5	$348,121
122	I-380	S Black Hawk B..	13.7	16,591	24	6	$480,869	592,057	56	$34,980	6	$304,192
102	I-380	Penn St. to 120..	13.5	56,876	17	7	$456,819	811,736	59	$33,735	11	$223,484
138	I-235	Martin Luther ..	3.7	114,062	6	8	$438,501	867,766	9	$118,454	56	$99,861
168	I-35	IA-210 to US-30	17.9	42,650	18	9	$429,969	823,628	79	$24,055	12	$222,948
20	I-80	690th St. to Lit..	11.5	22,298	36	10	$391,806	433,993	58	$34,139	7	$290,516

https://reactor.ctre.iastate.edu/interstate-delay-cost-summary/

Figure 4. Interstate congestion report for 2020

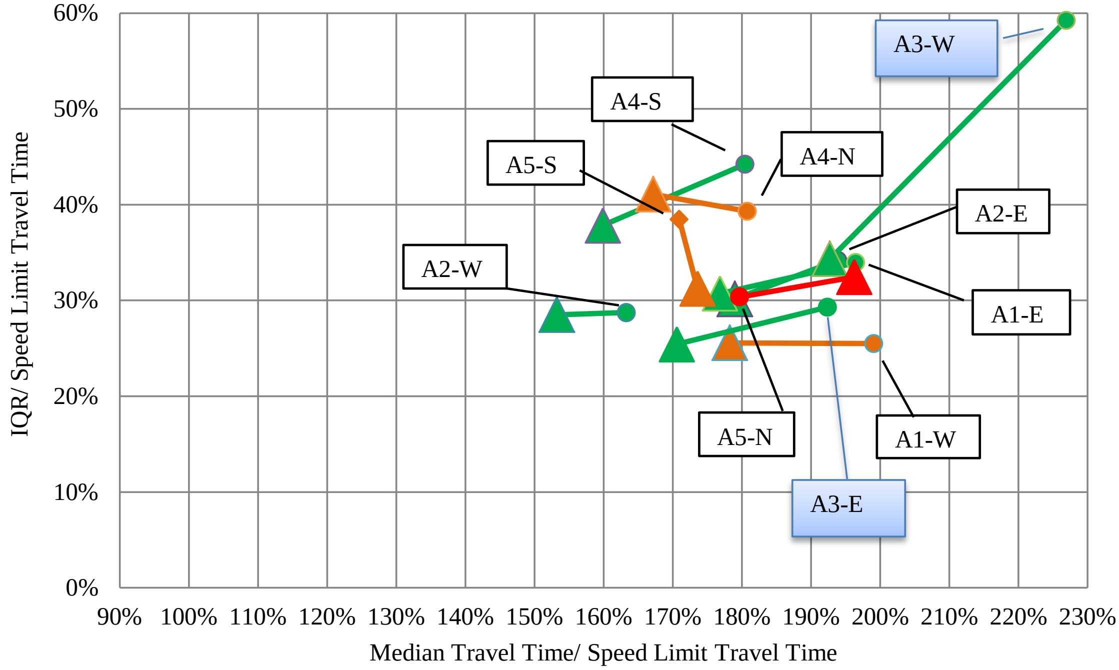
Figure 4 shows the estimated delay cost per section, with a top 10 list of sections with the highest delay. As the table shows, the sections are longer than the XD segments used to create the previous view. The section corresponding to the work zone area shown in the previous figures is the fourth item on the list (Section 68 on I-80). Probe vehicle data from various providers have been widely used for about 10 years by various agencies for similar applications (Day et al.

2016, Hainen and Dunn 2015, Iowa State University 2021, Michigan DOT 2020, Missouri DOT 2020, Pujara et al. 2019).

Most of these applications have focused on the performance of the interstate highway system. However, signalized arterials have also been analyzed through the use of probe vehicle data. A concern with the use of average speed per segment is whether traffic control would introduce dynamics that might be lost in the creation of an average. Researchers at the University of Maryland recommended that probe vehicle data be used for arterial highways with volumes exceeding 20,000 vehicles per day, a sparse density of traffic signals, low to moderate midblock friction, and with dominant through movements (Young 2014). Another study by researchers at the University of Washington found that travel times estimated from probe vehicle segment speed data had many more errors than those obtained from AVI data sources (Wang et al. 2014a). Still another study by researchers in Virginia compared probe vehicle data against Bluetooth AVI data, finding that the probe data are appropriate for evaluating long-term traffic state changes but are not useful for real-time applications (Hu et al. 2016a). Another study in Virginia applied probe vehicle data to an evaluation of adaptive signal control (Hu et al. 2016b).

In Indiana, probe vehicle data were applied to rank the performance of signalized arterials throughout the state using travel times measured from probe vehicle data (Day et al. 2015a). This was used to establish a performance measure that incorporated both delay, as measured using the travel time as a percentage of the speed limit travel time, and the standard deviation of the travel time (also normalized to the speed limit travel time). This was used to rank the corridor performance according to both the central tendency and amount of variability in travel characteristics. A similar analysis was done for the Pennsylvania DOT a few years later, which included an assessment of improvements to arterial corridors throughout the state (Mathew et al. 2017).

Figure 5 shows an example graphic from the Pennsylvania DOT study.



Mathew et al. 2017, Pennsylvania DOT

Figure 5. Before-after trends related to adaptive control system implementation for five arterial corridors in the Philadelphia area

Figure 5 shows the central tendency metric (median travel time) along the horizontal axis and the variability metric (interquartile range, IQR) along the vertical axis. Both the metrics are normalized to the speed limit travel time. For example, 100% along the horizontal axis means that the median travel time is the same as travel time at the speed limit, while 30% along the vertical axis means that the interquartile range (the difference between the 75th and 25th percentiles of travel time) is equal to 30% of the travel time at the speed limit. Each symbol represents the system state for a particular range of dates; the circular symbols represent a before period, and the triangular symbols represent an after period, with a line joining the two symbols that form a comparison pair.

There are two comparison pairs per corridor, because each corridor had two directions of travel that were analyzed separately. The green symbols indicate improved performance, the red symbols indicate worsened performance, and the orange symbols indicate that one axis saw improvement, whereas the other did not. The chart indicates that the overall system saw mostly net improvements for 6 out of 10 corridors/directions, while another 3 of them saw improvement in one dimension of performance, and while one other corridor saw worse performance. The public benefit of these changes was estimated at $24 million (Mathew et al. 2017).

This example is relevant to Iowa, because Pennsylvania takes a similar approach to management of signalized corridors, in that local agencies manage the signals on state highways. Pennsylvania made about $75 million in investments in signal systems through its Green Light-Go program from 2014–2019 (Pennsylvania DOT 2020). Probe vehicle data were already being purchased by the Pennsylvania DOT for other purposes. Cross-referencing the data set with a record of investments made under the statewide signals program permitted the benefits analysis to be accomplished.

In addition to average segment speeds provided by INRIX and HERE Technologies, there are some other data vendors that use AVL data to provide different kinds of traffic analytics. The Iowa DOT has purchased access to the StreetLight Data, Inc. analytics suite. Within this tool, the user is able to perform travel pattern analysis by manually defining gateways that delineate where vehicles would enter or exit a zone, and travel to and from other gateways can be identified from the data. The main use case for the data seems to be origin-destination (O-D) analysis, AADT estimation, and other count data. StreetLight Data volumes have been evaluated by several external studies for Minnesota, Oregon, Virginia, Louisiana, and Texas, with the following results:

- A 2020 study conducted by Texas Transportation Institute for the Minnesota DOT found an absolute error of 8%–10% for locations with AADT greater than 10,000, increasing to 42% for locations with AADT less than 1,000 (Turner et al. 2020).
- A 2019 study for Oregon DOT found 18% median absolute percent error comparing AADTs versus automatic traffic recorder data (Roll 2019).
- A 2020 study for Virginia DOT found mean absolute percent error (MAPE) of 7.1% for locations with AADT greater than 70,000, increasing to 18.2% for locations with AADT less than 10,000 (Yang et al. 2020).
- A 2020 study for Louisiana DOT reported a MAPE of 18.9% in StreetLight AADTs in comparison with permanent count stations (Codjoe et al. 2020). Other volume comparisons such as full-month and 24-hour counts had much higher errors.
- A 2020 study for Texas DOT focusing on the border region found a MAPE of 44.7% for locations with AADT greater than 10,000; a MAPE of 26.6% for AADT of 5,001–10,000; and MAPE of 33.6% for AADT less than 5,000 (Tsapakis et al. 2020).

The results of these AADT evaluations appear to vary considerably by location, and in most cases, the accuracy seems to improve for locations with higher volumes, although this is not always the case. AADTs alone are not sufficient to determine the quality of service on a roadway; it would be preferable to measure the roadway performance through vehicle travel times or other such measures. A search for results on this performance measure did not find any independent evaluations of travel time, although such data were mentioned as being available from StreetLight Data, Inc. The Virginia DOT report mentioned the placement of gateways around intersections to collect turning movement counts (Yang et al. 2020). However, the possibility of measuring delays or other such operational outcome measures was not explored in that study.

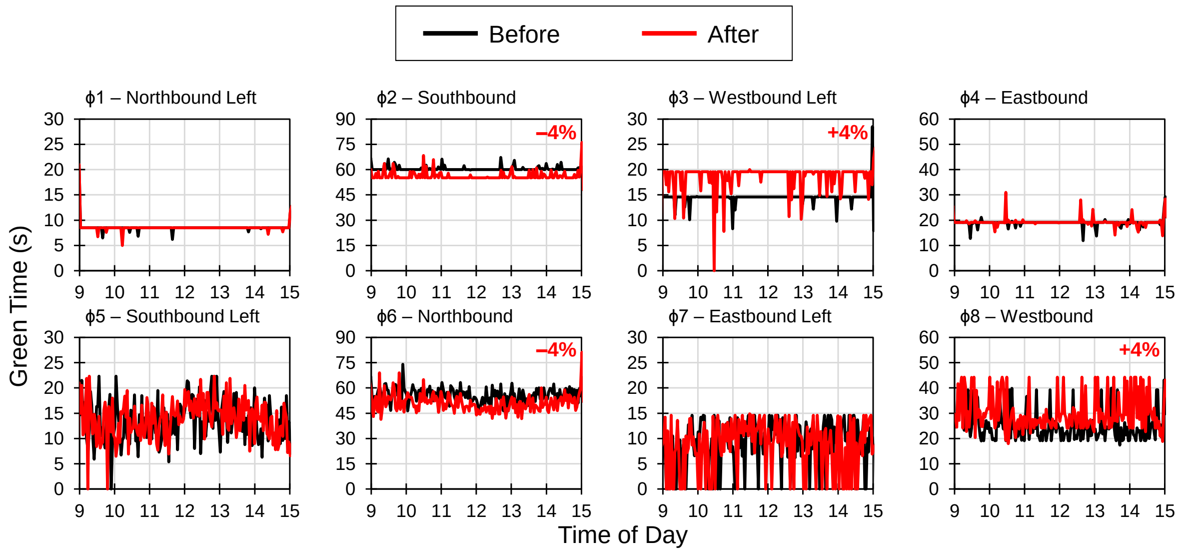
The raw version of AVL data remains an alternative data source and one that has a great deal of promise for evaluating details of intersection operations. This has been demonstrated in several studies in the past decade (Hofleitner et al. 2012, Wang et al. 2014b, Wolf et al. 2019, Wünsch et al. 2015). However, until very recently, data vendors had not begun marketing use of their data for this purpose. At the time of this present study, new data products are emerging that are being marketed as permitting a deeper analysis of signal operations. An example is the INRIX U.S. Signals Scorecard, which used a week of data to evaluate the performance of signals at 210,000 intersections in the US (INRIX n.d.). Such data sources are still in their infancy and have not yet been evaluated, but it seems that as the quantity of AVL data continues to increase, data for evaluating signalized intersections are likely to become increasingly available in the future.

2.4 Automated Traffic Signal Performance Measures

Traffic signals are present on many important facilities throughout any roadway network. While it would be difficult to estimate, it would not be too specious to guess that a majority of all trips by motor vehicle involve the passage through at least one traffic signal. They often control the ingress and egress from freeways, and signal timing is one of the primary determinants of the LOS on major surface streets of many cities and towns. Signalized intersections are frequently equipped with detection systems to provide for actuated control. For many years, traffic signals had very limited data collection capabilities. Many traffic controllers possessed native capabilities to record volume and occupancy data, typically collected in 5- or 15-minute intervals, to support traffic responsive operation, while some central system software for remotely programming traffic signals have included the capability of monitoring actuated green times.

Starting in the mid-2000s, researchers began collecting what has been called event-based data (Smaglik et al. 2007) or high-resolution data (Hu and Liu 2013), which consist of a record of the state changes in the detector inputs and the signal head outputs of a traffic signal. Such events include, for example, the time when a detector turns on or off (i.e., when a detector input state transitions from low to high or from high to low), and when a signal output changes from green to yellow to red, and so forth. High-resolution data can also capture some internal controller events that do not necessarily have a direct visual output, such as times of pattern changes or whether an actuated phase terminated because all its demand is served or because its maximum green time has expired. Numerous performance measures have been developed from these data that can be used to evaluate traffic signal operation (Day et al. 2014, 2015). This data source can yield some information about details of that operation that are not easy to be obtained by other means. The resulting performance measures are today known as ATSPMs. ATSPMs were selected as a focus technology in the fourth round of the FHWA’s Every Day Counts program and are today in use at an increasing number of intersections.

A very basic example of how ATSPMs may be used to examine details of signal operation is presented in Figure 6.



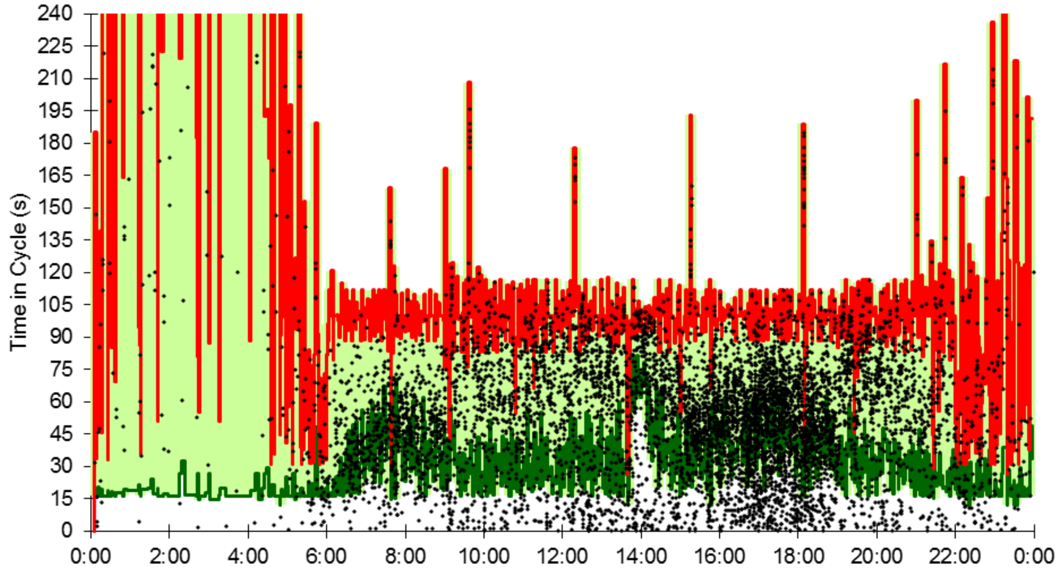
Day et al. 2015b, Purdue University

Figure 6. Green time during each signal cycle, arranged by phase, before and after a change to the splits at an actuated-coordinated intersection

Figure 6, which shows data from 9 a.m. to 3 p.m., contains eight charts arranged in the form of a phase diagram, which show how much green time was given to each phase in each cycle. In this example, 4% of the cycle was taken from phases 2 and 6 and given to phases 3 and 8. The shifting of the resulting green from the black before line to the red after line confirms the decrease in green times for phases 2 and 6, and the increases for phases 3 and 8. Because the signal is actuated-coordinated, one could speculate that when phase 3 does not use all of its programmed split, phase 4 that follows it could potentially access some of the yielded time. However, the chart shows virtually no change in the green time for phase 4, meaning that phase 3 utilized all of the time it received.

Additional metrics can be developed when measures of demand, such as count or occupancy, are taken into consideration. The availability and configuration of detection determines the type of metrics and level of detail that can be achieved from such an analysis. Where count data are available, the v/c ratio can be calculated (Smaglik et al. 2007), while stop bar occupancy data enables the number of split failures to be estimated from cycles in which a phase has a high level of occupancy during green and during a portion of the red time (Freije et al. 2014). When combined with overall measures of the intersection utilization, it is possible to determine where spare capacity exists and where there are opportunities to improve the signal timing by adjusting splits (Day et al. 2010a).

In addition to capacity allocation, another operational area critical to corridor performance is the quality of signal coordination. Where advance (setback) detectors are available, vehicle arrivals at the intersection can be measured and their times of arrival compared against the signal state to ascertain whether more vehicles are arriving in green or during red, which corresponds to good or poor coordination (Day et al. 2010b). An example of a detailed graphical performance measure for this purpose, called a coordination diagram, is presented in Figure 7.



Day et al. 2015b, Purdue University

Figure 7. Coordination diagram showing 24 hours of operation at a signalized intersection

The diagram shows arrivals (dots) during green (shaded region) for one signal approach. The basic idea is that every vehicle arrival is represented by a dot, and each arrival is charted according to the time of day it occurs (horizontal axis) and the time in cycle when it occurs (vertical axis). The signal state is superimposed on top of this, so that the green shaded regions represents when the signal is green, meaning that most of the dots coincident with this region arrived during green, and the others arrived during red. In brief, the more dots in green, the better the coordination. The performance can also be summarized using the POG as a single quantitative metric.

Many other details about the signal operation can be inferred from the chart. For example, the cycle length is indicated by the fluctuating red line in the figure, which shows the varying time between the beginning of red in each cycle and the previous beginning of red. During the late night and early morning periods, the cycle length is often very long, reflecting times when the signal rests in green on the major street while there is no side-street traffic. From 6:00 a.m. to 10:00 p.m., the signal is coordinated, and the cycle length varies from about 90–120 seconds, with minor variations due to the use of early yield. There are also individual cycles when the cycle length is unusually high or low, which is associated with preemption events that appear to cause the major street phase to be extended for a long period of time.

While these performance measure views present a highly microscopic view of the signal operation, it is also possible to use the data to develop corridor- and system-level views of performance through aggregation. This sort of aggregation can potentially be more helpful if it is undertaken from a perspective of identifying opportunities for improvement (Day et al. 2018). That is, rather than simply finding the average performance or the worst-performing movement in the system, it is possible to take into consideration whether enough flexibility exists at the intersection to allow for retiming.

An example of a system-level scorecard is presented in From Day as presented at Every Day Counts (EDC-4) workshop in Helena, Montana, August 2018

Figure 8.

Performance Information	Corridor Number							
	1	2	3	4	5	6	7	8
Number of Intersections Total	15	12	5	9	10	13	8	7
Number of Intersections Online	14	10	5	5	10	11	7	7
Percent Online	93%	83%	100%	56%	100%	85%	88%	100%
Communication Subscore	B	C	A	F	A	C	C	A
Number of Detectors	185	138	75	85	142	133	100	97
H1 Detectors	19	11	42	9	4	3	6	8
H1 Rate (% of detectors affected)	10	8	56	11	3	2	6	8
H1 Subscore	B	B	F	B	A	A	B	B
H2 Detectors	1	0	0	0	4	0	0	6
H2 Rate (% of detectors affected)	1	0	0	0	3	0	0	6
H2 Subscore	A	A	A	A	A	A	A	B
Number of Phases	382	242	129	183	199	253	209	148
H3 Phases	42	31	123	2	12	3	31	42
H3 Rate (% of phases affected)	11	13	95	1	6	1	15	28
H3 Subscore	B	B	F	A	B	A	B	C
H4 Ped Phases	0	0	0	0	0	0	0	0
H4 Rate (% of pedestrian phases affected)	0	0	0	0	0	0	0	0
H4 Subscore	A	A	A	A	A	A	A	A
Detection Subscore	B	B	F	B	B	A	B	C
Highest red light violation rate per 1000 vehicles	15.2	8.6	(a)	12.8	17.3	23.1	8.8	16.4
Safety Subscore	C	B	(a)	C	C	D	B	C
AM Peak capacity subscore	B	B	(a)	C	A	B	C	C
Midday capacity subscore	B	B	(a)	C	A	C	C	C
PM capacity subscore	C	B	(a)	C	B	C	C	D
Capacity Allocation Category Subscore	C	B	(a)	C	B	C	C	D
AM Peak progression subscore	C	B	(a)	C	C	(b)	B	C
Midday progression subscore	B	B	(a)	B	C	(b)	A	C
PM Peak progression subscore	B	B	(a)	B	C	(b)	A	C
Progression Category Subscore	C	B	(a)	C	C	(b)	B	C
Overall Corridor Score	**C**	**C**	**F**	**F**	**C**	**D**	**C**	**D**

From Day as presented at Every Day Counts (EDC-4) workshop in Helena, Montana, August 2018

Figure 8. Example of system level ranking

From Day as presented at Every Day Counts (EDC-4) workshop in Helena, Montana, August 2018

Figure 8 shows the results of a system-level analysis are presented for eight arterial corridors that have between 7 and 15 intersections. The overall score is shown at the bottom; this reflects the worst score out of five sub-scores. These sub-scores evaluate the corridor's performance according to the following criteria:

- Communication: Are intersections online and recording data?
- Detection: Are most of the detectors in working order? The following four heuristic scores are included that define different ways that detectors could fail:
  - H1: Detectors are not reporting any data.
  - H2: Detectors are reporting erroneous data (overcounting).

  - H3: The signal is constantly cycling during overnight hours when it should not be.
  - H4: Pedestrian button is stuck (pedestrian phase is constantly called).
- Safety: Are there a high number of red-light runners?
- Capacity: Are movements at intersections receiving adequate green time?
- Progression: Is traffic being progressed at coordinated intersections?

Thus, with the use of ATSPMs, it is possible to obtain a variety of data covering numerous maintenance and operations issues. In particular, the considerations of capacity and progression would be relevant to assessing the operational performance of a corridor. However, ATSPMs do require a certain amount of infrastructure to enable their use including the following:

- A means of data collection is required. Many recent signal controller models possess the capability of logging high-resolution data. There are also third-party devices that can be used with controllers that lack this capability.
- Detection at the intersection is needed to measure demand and utilization. In many cases, existing detection can be used for this purpose. However, intersections that do not have working detection or that have movements that do not include detection will be able to report a rather limited amount of information.
- Communication to the intersections is needed to automatically download the high-resolution data. Otherwise, it may be possible to store the data on site and retrieve it periodically, although some controllers have limited storage space.
- A system that processes the data or yields performance measures would also be necessary to be able to make better use of the data. One such system option is the open-source software released by the Utah DOT. Several vendors offer a product to do the same. At the time of this present study, there were still limited options for performing aggregation from detailed intersection- and movement-based performance up to system-wide views.

2.5 Connected Vehicle Data

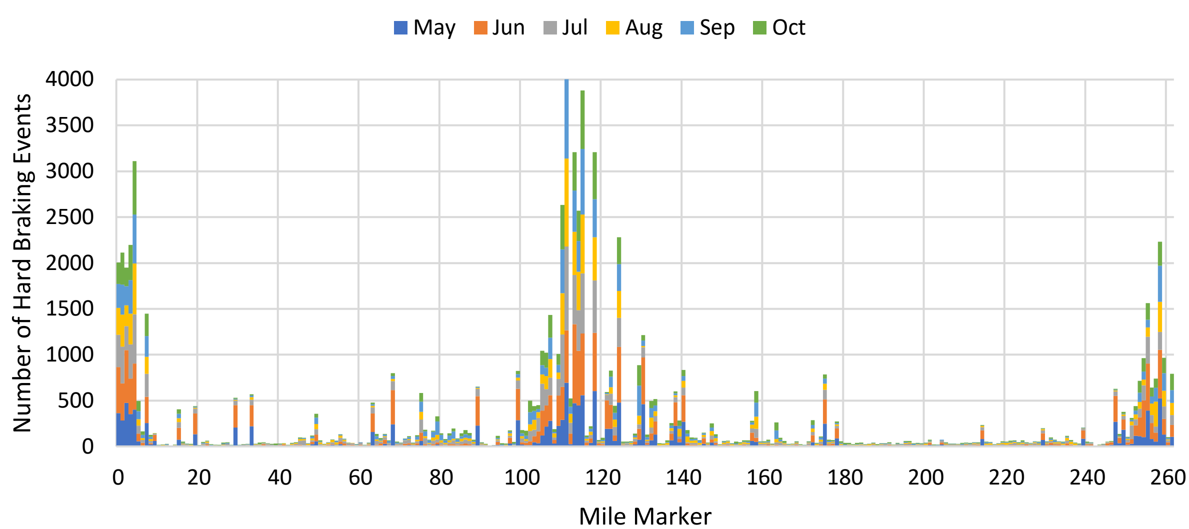
At the time of the study, new data sets are emerging that can provide a more detailed view of vehicle travel than existing data sets and a host of driving-related events. For the purpose of this report, these data will be called connected vehicle (CV) data, although that name has varied meanings depending on its context. The most basic definition of a CV is a vehicle that possesses some capability of automatic, two-way communication with other vehicles, infrastructure, and devices such as smartphones, in a manner that is integrated with the vehicle systems such that internal information available through those systems can be shared. This includes not only the vehicle position and speed as measured through GPS devices but other data from the various systems used in running the vehicle: steering, braking, turn signals, lights, windshield wipers, etc. This marks the main difference between CV data and the type of information obtainable through transported mobile devices, which are used to develop probe vehicle data described previously. CV data can provide the same position and speed information, as well as additional information about vehicle-system events.

There are two pathways for this type of data to be obtained by transportation system operators. One way to obtain CV data is for it to be received in communication between vehicle onboard units (OBUs) and roadside units (RSUs) installed as part of road infrastructure. One of the applications of this environment would be to broadcast the state of traffic signals through RSUs, in signal phase and timing (SPaT) messages. Vehicles would receive these messages and be informed about the current and future state of a traffic signal. Additionally, vehicles could communicate their position, speed, etc. through basic safety messages (BSMs) back to the RSU and to other vehicles. This could be used for applications such as collision avoidance, for example, by advising a driver to stop to avoid running a red light, or advising drivers on conflicting approaches that a vehicle appears to be about to run a red light.

To date, there have been various pilot deployments of RSUs, but they are very far from being ubiquitous. It is not immediately clear whether the data could have uses beyond real-time applications such as those described herein. For example, the U.S. DOT has stated that CV technology "does not involve exchanging or recording … vehicle movements" (U.S. DOT n.d.), which would seem to imply that any movement records would be limited to simple paths in the range of one RSU, e.g., the passage of vehicles through one intersection.

Another path for the data is through auto manufacturers and their partners, which receive the information not through RSUs but through cellular communication. Auto manufacturers collect these data, and in the past few years, a few companies have begun to sell both the raw data and derivatives of it. Again, the difference between these data and the probe vehicle data described previously is that CV data include information from vehicle systems, so information such as hard braking can be obtained in addition to positions and speeds. Because the data are not being collected by RSUs, concerns about managing the data process and dealing with privacy concerns are effectively transferred to the private sector.

An example of hard-braking events data is shown in Figure 9.



Mathew et al. 2020, Purdue University and Wejo Ltd. with data used under Creative Commons license.

Figure 9. Hard-braking events per mile on northbound I-65 in Indiana

Figure 9 shows the number of observed hard braking events on each mile of northbound I-65 in Indiana during six months in 2020. The data were obtained from an open-source data set consisting of extracted time and mile location of events (Mathew et al. 2020, 2021). The number of observed events seems to increase in the proximity of urban centers (Louisville near mile 0, Indianapolis between mile 100 and 120, and the greater Chicago area near mile 260). There appear to be other hot spots outside of those areas as well, likely near freeway entrances or exits. The results are rather intuitive, but this early view from the data set yields sensible results, and it is likely that more insights can be gained from the data in the future, with deeper analysis and linking to other traffic data sets.

2.6 Summary

This chapter examined data sources that could be used for development of an OICL. After briefly discussing conventional data sources such as traffic volumes or speed measurements by detectors, the chapter focused on two data sets that have some utility for OICL development, because they can capture some information about roadway operation. One of these is probe vehicle data, in particular AVL data that captures vehicle speeds. The Iowa DOT is already purchasing segment speed data from one provider, and these data are available for both interstate and non-interstate routes. Another data set is high-resolution data, which is associated with ATSPMs. At the time of the study, there were limited deployments of ATSPMs in Iowa, but it seems likely that such data will become more commonly available in the future. This data set permits the development of detailed metrics about signal operation. Finally, the chapter concludes with a discussion of emerging data sets that include both vehicle position and speed information as well as driving events such as hard braking.

3. EVALUATING CORRIDOR OPERATION WITH SEGMENT SPEED DATA

3.1 Introduction

The Iowa DOT currently purchases segment speed data from INRIX to support a variety of operational analyses. While much of the emphasis on mobility reporting has been focused on the performance of the interstate system, these data are also available for other types of roadways. This chapter explores the use of that data for evaluating the performance of signalized corridors in Iowa to support the development of an OICL.

3.2 Identifying Locations for Corridor Ranking

The first step in performing a ranking of signalized corridors in Iowa was to identify the locations of traffic signals in the state. To do so, the locations of signals in Open Street Maps (OSM) (www.openstreetmap.com) were used. Intersections with traffic signals are often (although not always) marked with signal icons. Roads with more than two lanes are sometimes represented in OSM with two separate one-directional roadway elements, and when these intersect other roadways, sometimes two or four signal icons are produced, as shown in Figure 10.

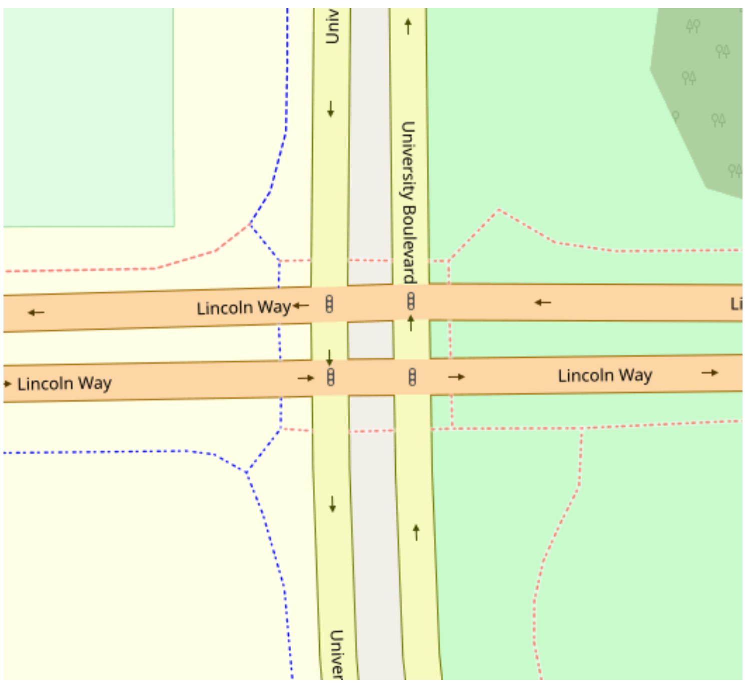
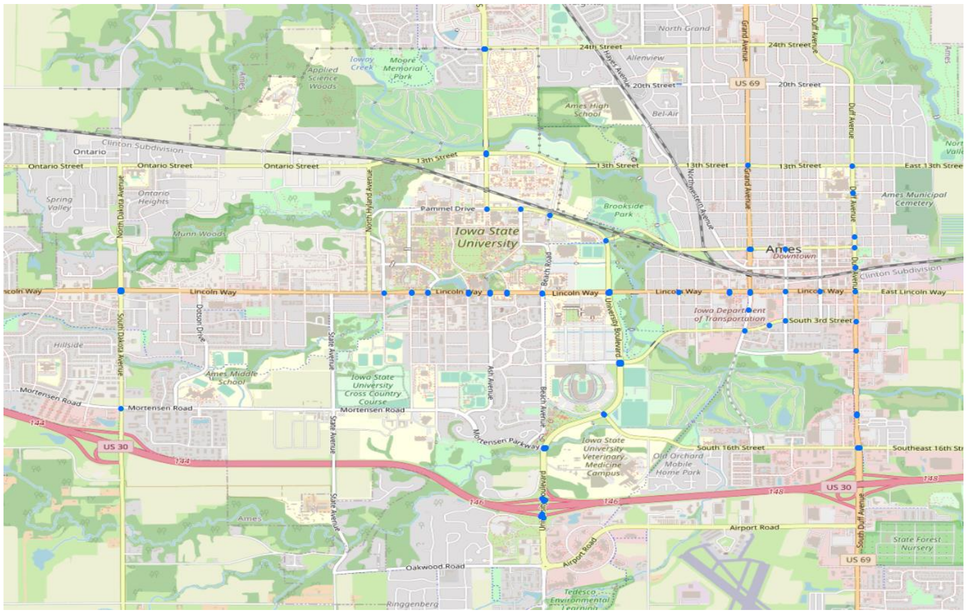


Figure 10. Example of an intersection with additional traffic signal icons

In some cases, even more signal icons may be generated, for example, in the case where a road has separate lane groups (e.g., channelized right turn lanes) that are represented in OSM as separate roadways. It is possible to obtain the coordinates of these icons by querying for these locations using an application programming interface (API), such as Overpass Turbo (overpass-turbo.eu).

Figure 11 shows the locations of signalized intersections recorded by OSM in Ames, Iowa.

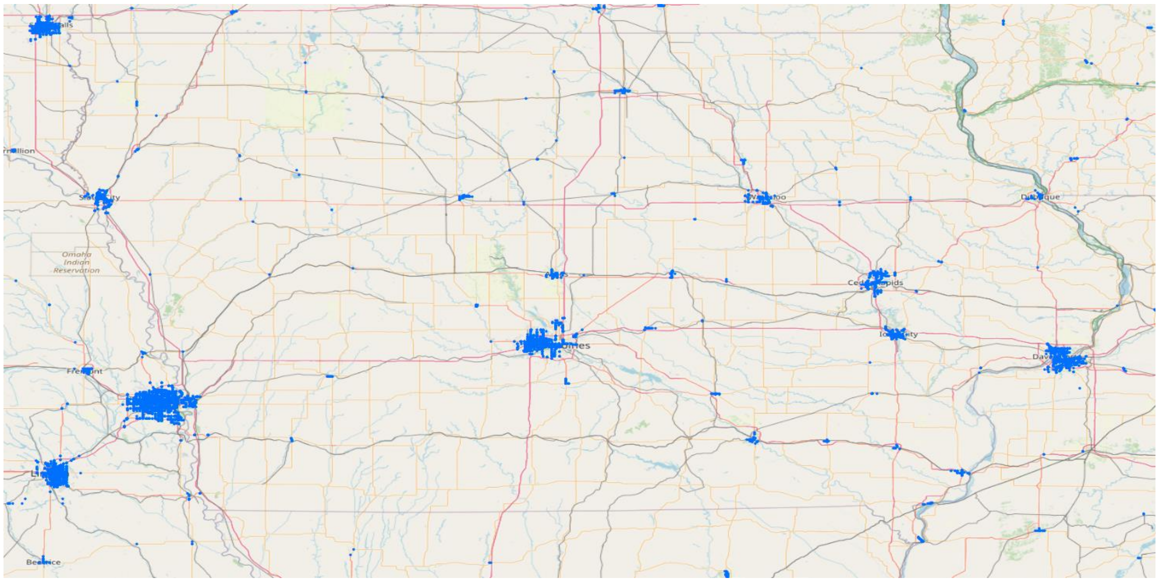


©OpenStreetMap contributors (www.openstreetmaps.org)

Figure 11. Locations of signalized intersections in Ames, Iowa, according to OSM

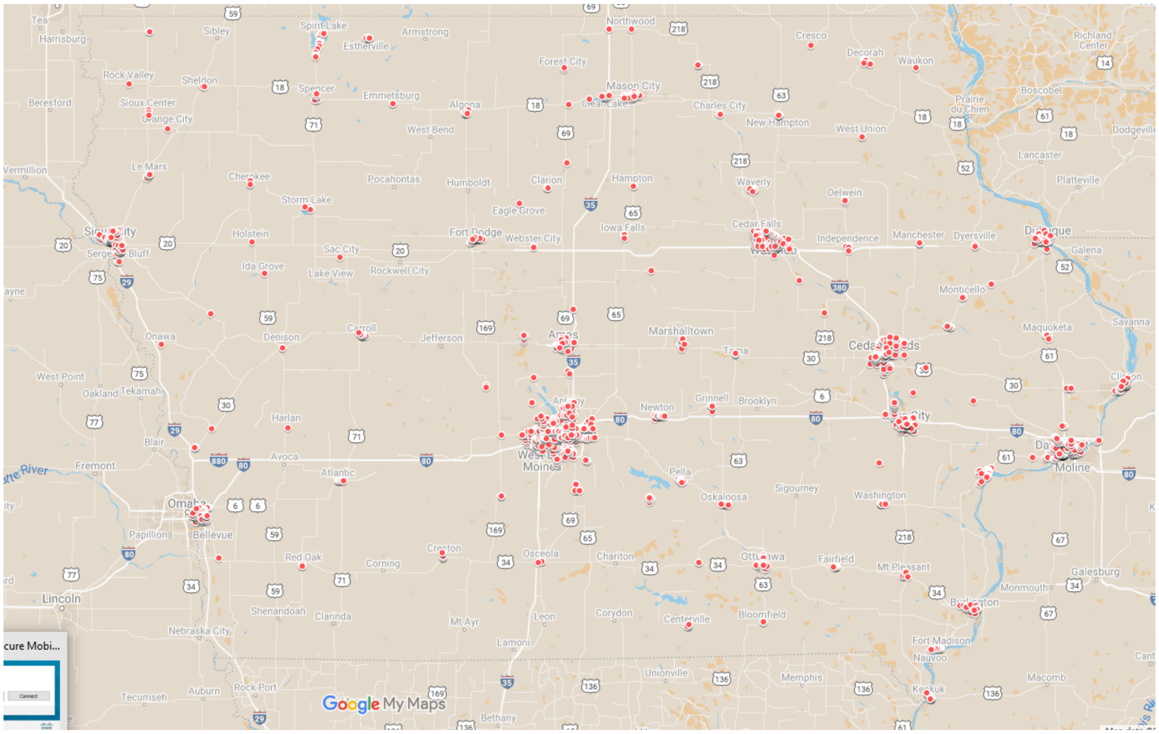
The OSM marking of signals is not perfect; there are several missing intersections, such as those on Grand Avenue north of 13th Street, and on Lincoln Way between North Dakota Avenue and Sheldon Avenue. There is also the issue of multiple signal icons located at the site of one signalized intersection. However, these data proved to be a useful starting point, with additional visual confirmation used to confirm signal locations and to add other intersections by a visual examination of major roads in urban areas, with the help of the OSM data.

Figure 12 shows the locations of signalized intersections from OSM across Iowa, while Figure 13 shows the confirmed locations based on visual comparison with satellite images on Google Maps. From this analysis, there are an estimated 2,300 signalized intersections within Iowa.



©OpenStreetMap contributors (www.openstreetmaps.org)

Figure 12. Locations of signalized intersections in Iowa, according to OSM



Base map image © 2021 Google Maps

Figure 13. Locations of signalized intersections in Iowa after visually confirming locations and identifying additional signals from satellite images

3.3 Obtaining and Processing Travel-Time Data

After defining corridors, the next step was to identify segments passing through the corridors. This was done by comparing the signal locations confirmed in the previous process against the INRIX XD segment shapefile. This was a manual selection process. Altogether, about 250 corridors were defined, as shown in Figure 14. Each corridor includes two directions of travel.

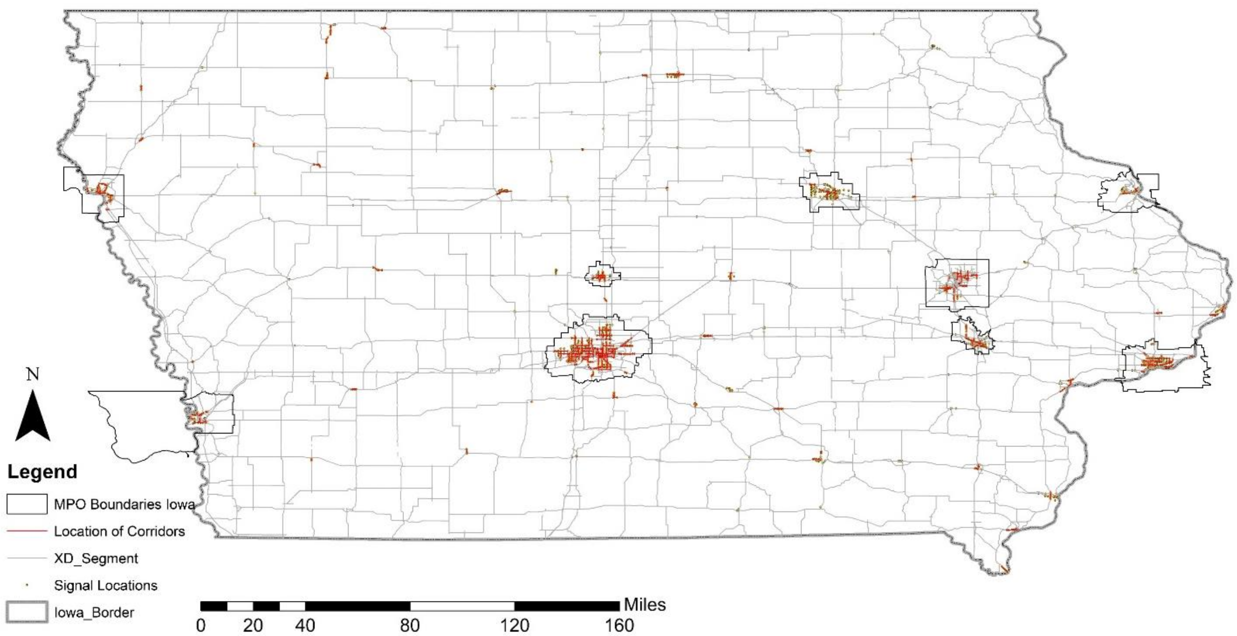


Figure 14. Distribution of signalized corridors in Iowa selected for ranking

Data were obtained from the data repository used by the Institute for Transportation (InTrans) Real-Time Analytics of Transportation Data Laboratory (Reactor Lab) for storing data. At the time of this analysis, the data were being migrated from an in-house server to a cloud server, and there were some time periods for which data were unavailable in 2019 and 2020. To mitigate the effects of missing data, the corridor ranking was performed using selected data on Wednesdays in June, September, and December of both 2019 and 2020. Those months had mostly complete data in both years.

An example of the travel-time data for one corridor is shown in Figure 15.

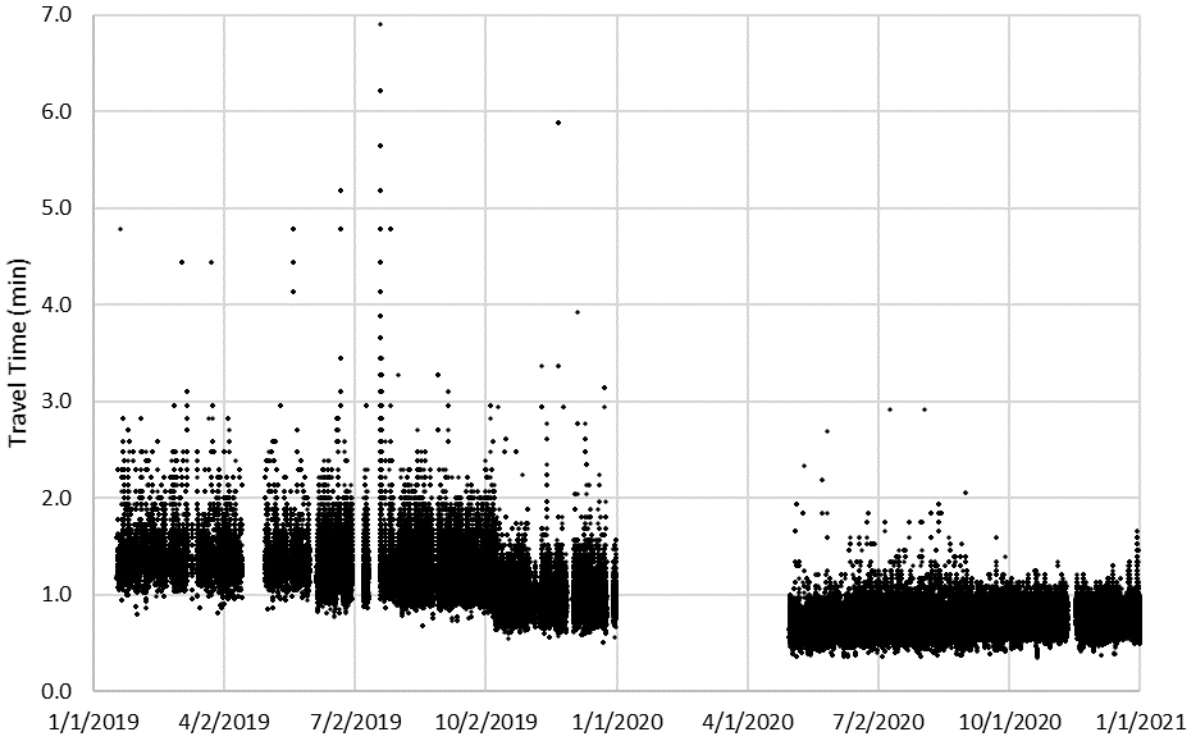


Figure 15. Travel times for an example corridor, including periods of unavailable data

Figure 15 includes only data for which a score value of 30 was available, indicating data from real-world observations. Despite the periods of missing data evident in this figure, there remains quite a bit of useful data remaining about the performance of this corridor. This chart contains approximately half a million travel-time observations from the two-year period. The impact of traffic reductions during the COVID-19 pandemic is also evident in this diagram, with considerably lower travel times appearing in 2020 as compared to 2019. Altogether, there were over 600 million individual travel-time observations for all corridors across both years, of which ultimately 3 million were used for the ranking analysis for the selected months, days of week, and times of day.

The character of the travel-time data for arterials is revealed by a closer examination of the raw data, as presented in the next few figures. Figure 16 shows a 24-hour view of the data along with counts of the number of observations.

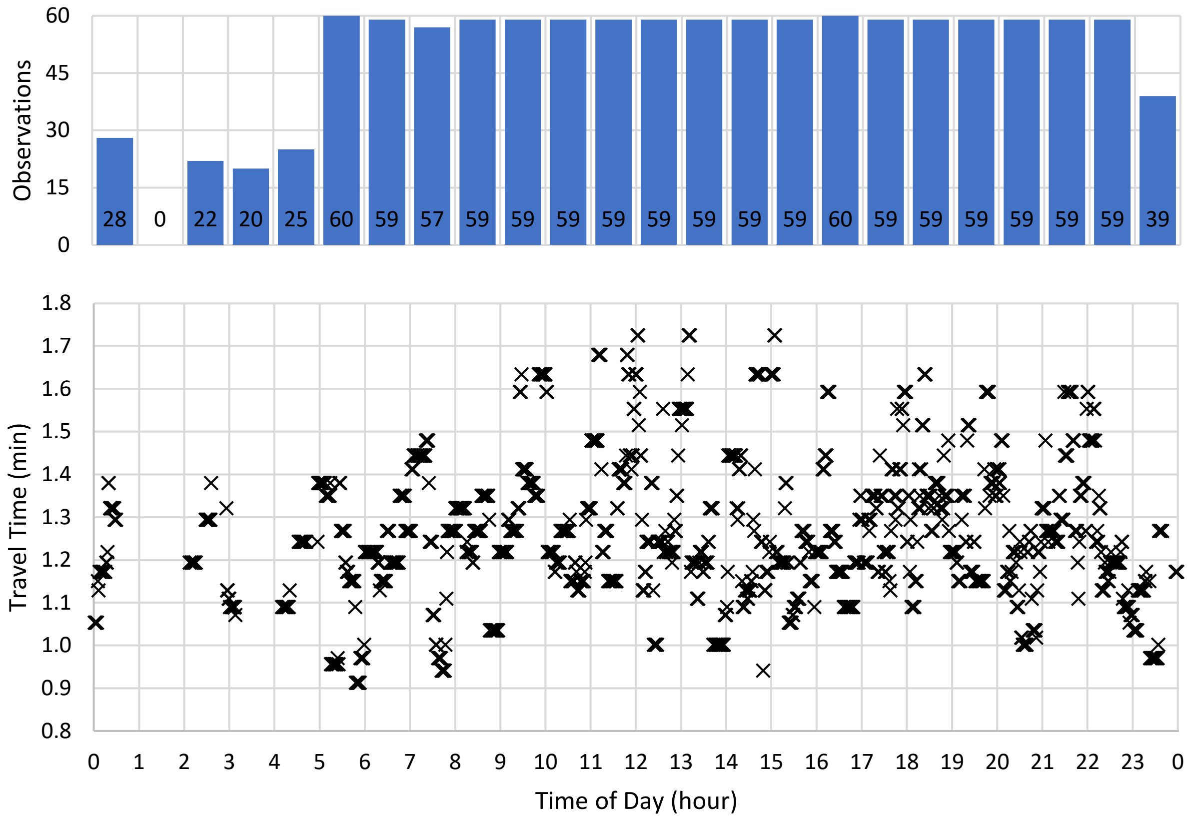


Figure 16. Data for June 12, 2019 and number of observations per hour

This analysis is similar to an analysis done several years ago on a different signalized corridor using INRIX data for a 24-hour period in 2013 (Day et al. 2015a), except that the data are now more complete, with most of the hours of the day reporting nearly complete coverage and only a few gaps in the middle of the night.

An even closer look at the data is presented in Figure 17, which shows the value for each minute during the noon hour for the same data shown previously.

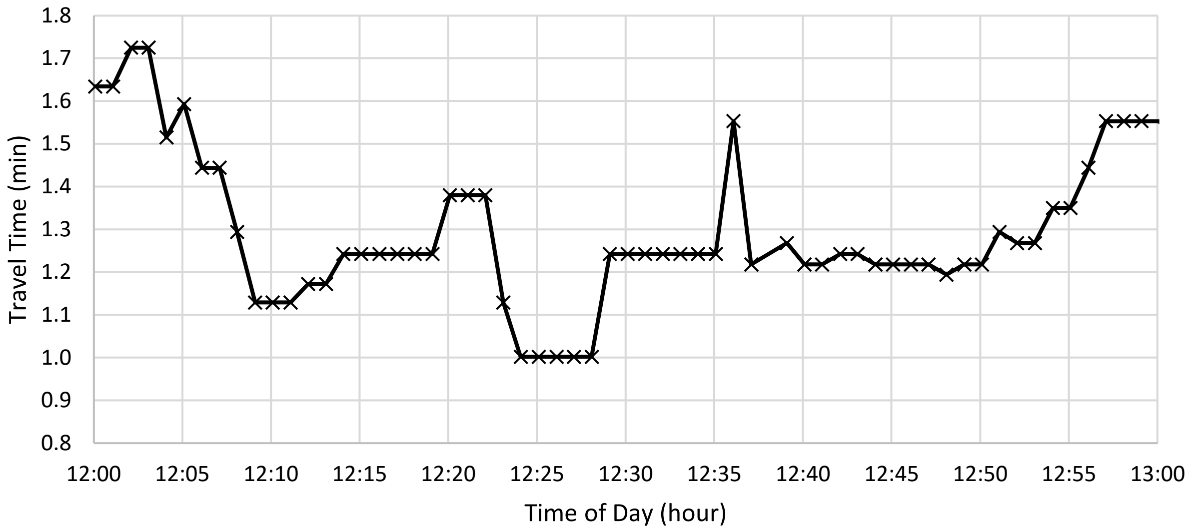


Figure 17. Minute-by-minute travel-time data

Note for Figure 16 and Figure 17 that data are shown for Eastbound 23rd Avenue in Council Bluffs, Iowa.

The travel time varied from 1.0 to 1.7 minutes, which is a lot of variation. This road segment is 0.7 miles long, so the corresponding speed range was 25–42 mph. This is likely due to the stop-and-go nature of interrupted flow facilities, which likely affects the observed speeds of the probe vehicles seen within each minute. The methods of the data provider, number of probes, and so forth are not public information, so the dynamics that yield this variation can only be speculated upon.

Travel time for a corridor was calculated as the sum of the travel time on all of the segments comprising each corridor. The 250 corridors varied from including a single segment to as many as 16 segments. A travel time was calculated for every one-minute interval using whatever data were recorded in that interval. This was feasible because the data set was relatively complete for the busy hours of the day, as illustrated in Figure 16. If one segment was missing a travel-time record for the minute, the missing record was substituted with travel time representing travel at the free-flow speed. Overall, about 60% of all the minute-segment pairs had real data from which travel-time estimates could be generated.

After obtaining a travel time for each corridor for each minute, the next step was to aggregate the travel times and produce metrics reflecting their average value and the reliability of the travel time. Standard deviation was used as a measure of variability/reliability. The aggregations were done across all of the data within one year for the months selected previously. Separate aggregations were done for the a.m. peak, midday, and p.m. peak. Thus, each corridor had six sets of values in the end for the three time-of-day periods and for the years 2019 and 2020.

Because each corridor has a different length and speed limit, its average and standard deviation values were normalized by using the travel time at the free-flow speed given respectively in the following equations:

$$x_n = \frac{x}{t_F} \tag{1}$$

$$s_n = \frac{s}{t_F} \tag{2}$$

where x and s are the average and standard deviation of travel time, respectively; x_n and s_n are the normalized average and standard deviation, respectively; and t_F is the travel time at the free-flow speed.

A PI was used to rank the corridor performance, incorporating the normalized average value x_n and the normalized standard deviation s_n. To bring the two metrics into a single quantitative measure, the Euclidean distance to a point $\{x_n, s_n\}$ was calculated by the following equation:

$$PI = \sqrt{x_n^2 + s_n^2} \tag{3}$$

In general, the lower the value of the PI, the better the performance of the corridor. The ideal value of the PI is 1.0, which represents travel at the speed limit with perfect reliability (i.e., no variation whatsoever). A separate PI value was calculated for each direction of travel on each corridor, for each time of day. The maximum value of the PI was used for the final value of the corridor used in the ranking.

3.4 Ranking of Signalized Corridors by Travel Time and Travel-Time Reliability

The results of the ranking analysis are shown by Figure 18.

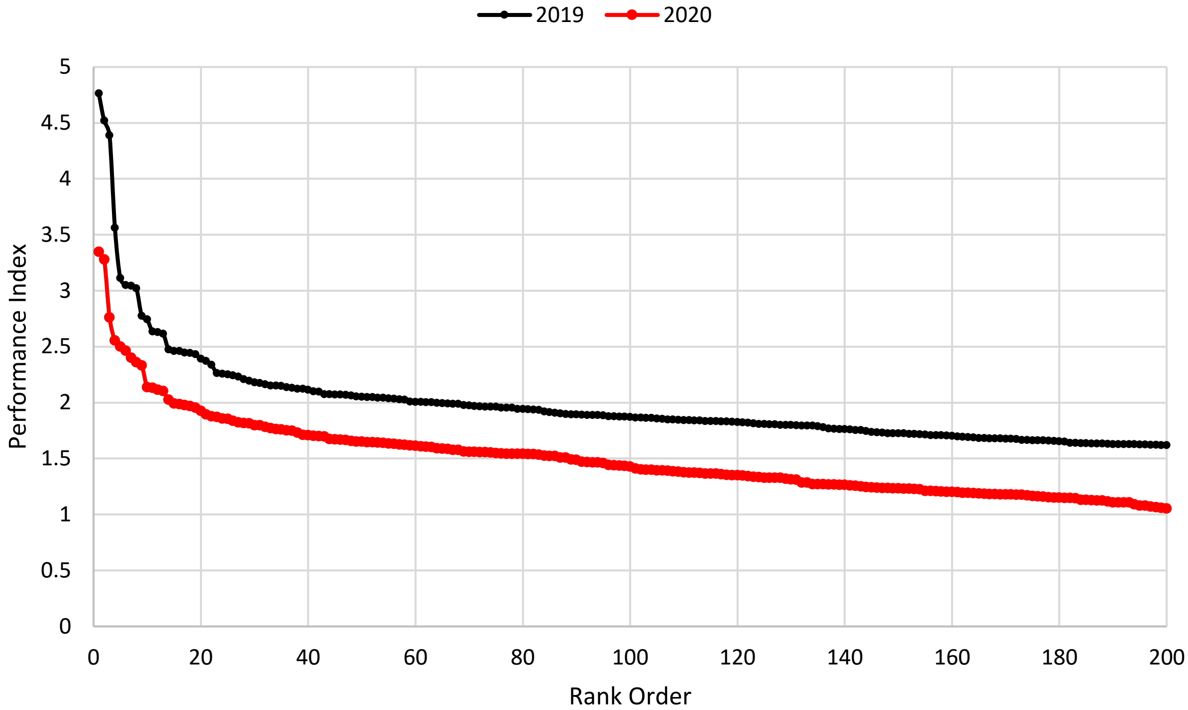


Figure 18. Pareto sort of signalized corridor performance

Figure 18 is a Pareto-sorted diagram showing the performance of the top 200 corridors for each year. As can be seen in the diagram, about 20 corridors out of the entire list appear to have a markedly higher PI than the others, as seen by the inflection point in the curve close to rank 20. It is also clear that PI was much lower in 2020 across the board. A breakdown of the PI by time of day is presented in Figure 19, which shows that similar trends are evident within each of the three time-of-day periods.

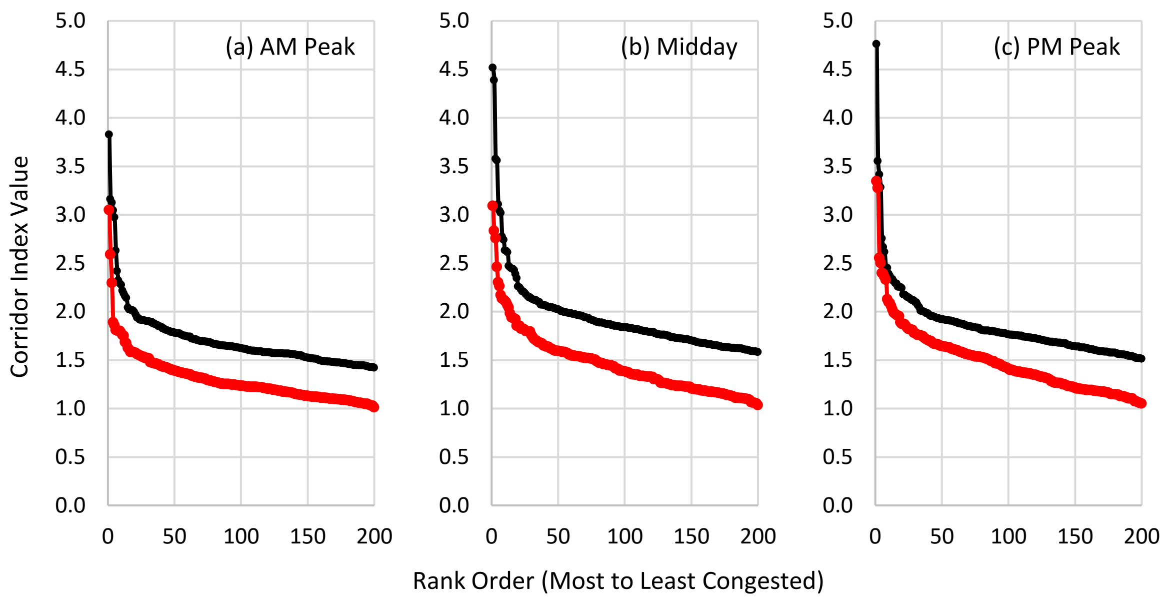


Figure 19. Pareto sort by time of day

Another view of the difference in the PI between 2019 and 2020 is presented in Figure 20.

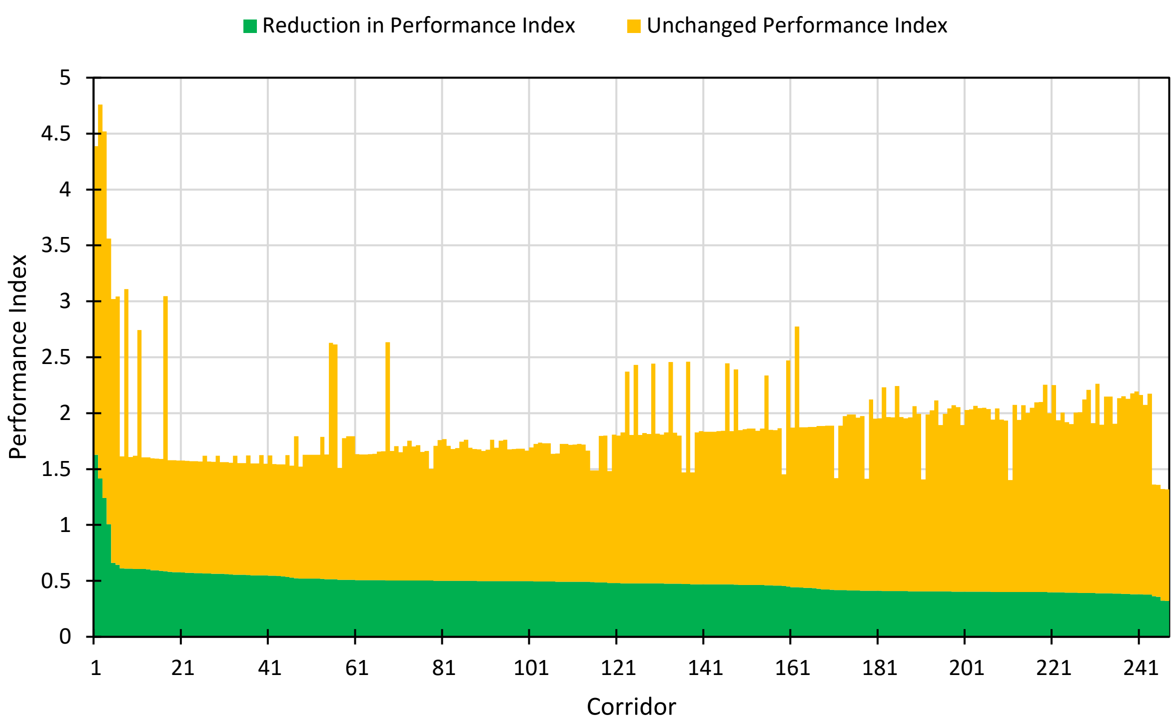


Figure 20. 2020 vs. 2019 performance: corridors with decreased PI (less congestion)

Figure 20 shows the amount of decrease in the PI, superimposed on top of the 2019 value of the PI. The data are sorted by the amount of decrease. Overall, the PI decreased by about 25%–33% as a result of reduced traffic volumes during the COVID-19 pandemic. Finally, the top 10 most congested locations in 2019 and 2020 are presented respectively in Table 1 and Table 2.

Table 1. Top 10 most congested arterial corridors in 2019

No.	Location	PI
1	7th Street, Des Moines	4.76
2	Washington Street, Waterloo	4.52
3	Wesley Parkway, Sioux City	4.39
4	North Dubuque Street, Iowa City	3.56
5	CR D-20, Iowa Falls	3.11
6	IA 130, Davenport	3.05
7	23rd Avenue, Council Bluffs	3.04
8	IA 1, Iowa City	3.02
9	3rd Street, Des Moines	2.77
10	East San Marnan Drive, Waterloo	2.74

Table 2. Top 10 most congested arterial corridors in 2020

No.	Location	PI
1	7th Street, Des Moines	3.35
2	North Dubuque Street, Iowa City	3.28
3	IA 130, Davenport	2.76
4	CR D-20, Iowa Falls	2.55
5	Gordon Drive, Sioux City	2.50
6	Southeast Corporate Woods Drive, Ankeny	2.46
7	Wesley Parkway, Sioux City	2.40
8	9th Avenue, Council Bluffs	2.36
9	Ingersoll Avenue, Des Moines	2.33
10	East Mullan Avenue, Waterloo	2.14

3.5 A More Detailed Look at Changes between 2019 and 2020

In 2020, the COVID-19 pandemic induced many changes in transportation activities around the world. In the US, many states enforced quarantines starting in March, and although many of those were lifted relatively quickly, many organizations encouraged employees to work from home in the following months. As a result, traffic volumes decreased substantially, with reductions in volumes of up to 60%–65% reported in some places (ITE n.d.). This situation offered an opportunity to use the corridor data developed for this chapter's analysis to visualize overall system performance. The charts presented previously show these results in the composite corridor PI, but the components of that PI are able to reveal additional information.

Table 3 presents a summary of the changes between 2019 and 2020 for the 250 corridors analyzed in this study, by time of day.

Table 3. Summary of performance outcomes

Condition	a.m.	Midday	p.m.
Better	154	131	127
Neutral	9	30	33
Worse	1	3	4
Insufficient data in 2020	37	37	37

Better indicates that both the mean travel time and standard deviation of travel time decreased, neutral means that one of these increased while the other decreased, and worse indicates that both increased. There were 37 corridors for which there were insufficient data from 2020 to perform a comparison. As Table 3 reveals, the vast majority of corridors saw improved performance, with lower travel times most likely due to the reduced traffic volumes. Detailed views of these data are given by Figure 21, Figure 22, and Figure 23 respectively for the a.m. peak, midday, and p.m. peak.

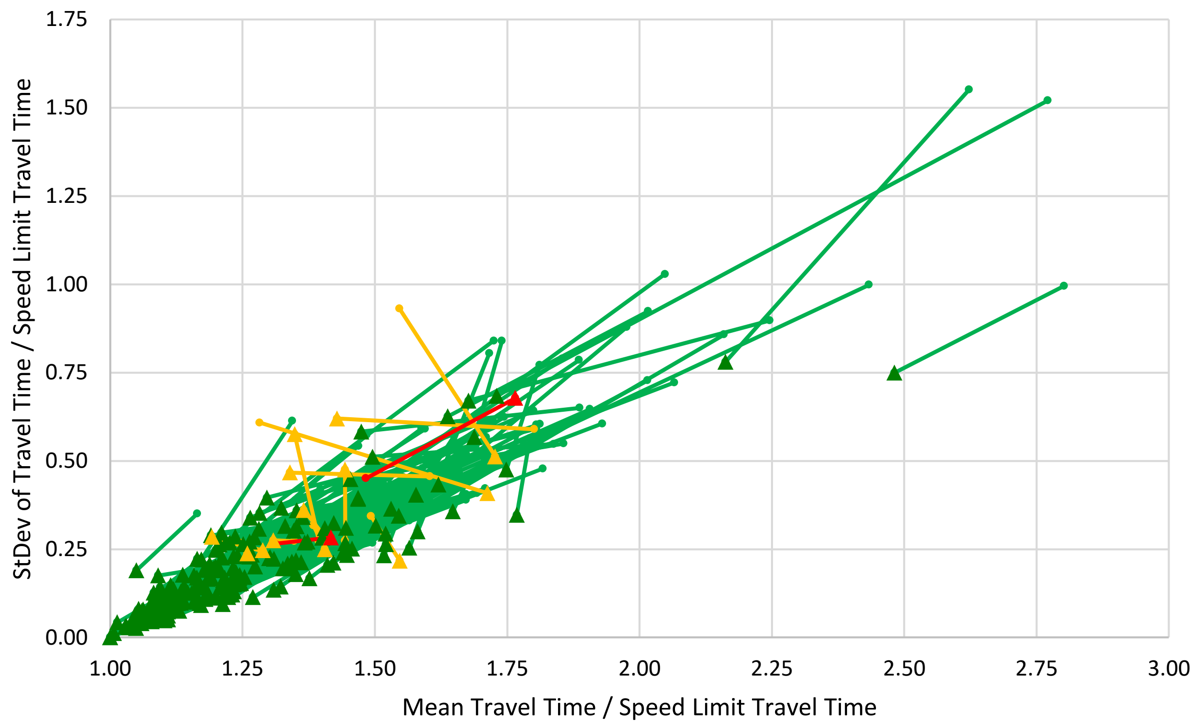


Figure 21. Normalized standard deviation versus normalized mean, a.m. peak

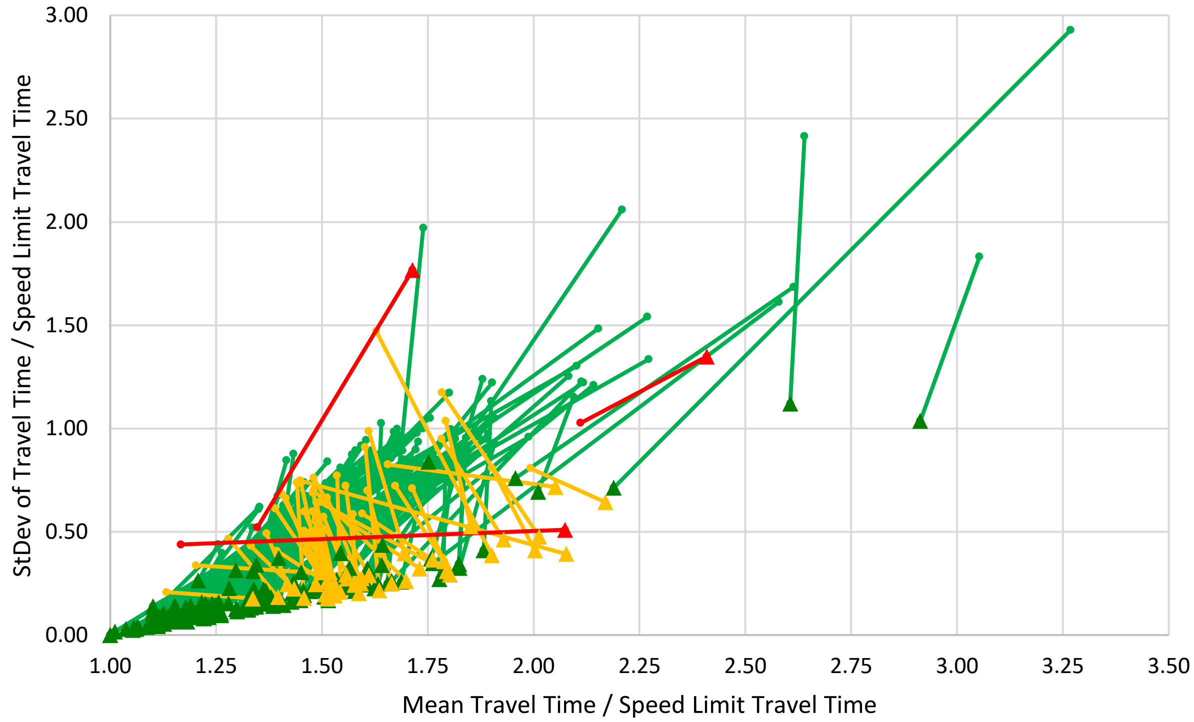


Figure 22. Normalized standard deviation versus normalized mean, midday

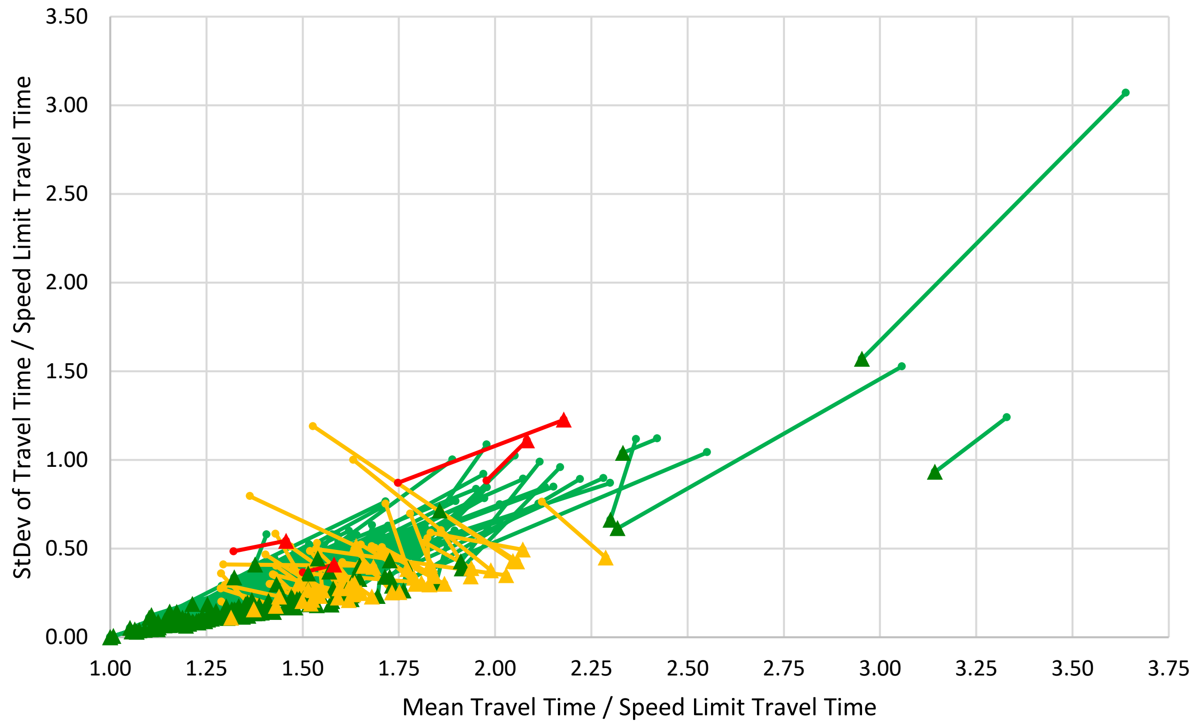


Figure 23. Normalized standard deviation versus normalized mean, p.m. peak

These figures show the movement of the system state from 2019 to 2020, with the triangular symbol representing the after case, with color coding of green for better, yellow for neutral, and red for worse performance. This shows not only the trend but the magnitude of the change. There are many corridors with very substantial decreases in both mean and standard deviation of the travel time. However, this is not universal. There are also quite a few corridors that had neutral outcomes; for most of these, the average travel times increased slightly, while the standard deviation of travel time decreased for many. Finally, there are a handful of corridors for which the performance worsened for a particular time of day. However, as earlier results showed, when combined into an overall index, there were no corridors that saw overall worse performance than 2019.

4. EVALUATING INTERSECTION OPERATION WITH HIGH-RESOLUTION DATA

4.1 Introduction

Whereas the segment speed data explored in the previous chapter are able to provide insights about corridor performance, the data do not provide much insight about the performance of crossing streets on those corridors, nor do they differentiate between different movements on signalized approaches. Thus, to investigate the operation of different movements at intersections, it is necessary to augment such data with another data set that has such capabilities. In this study, the high-resolution data used for ATSPMs are employed.

During the course of the OICL project, high-resolution data became available in the Iowa cities of Cedar Rapids and Dubuque. Cedar Rapids deployed data collection at many intersections throughout the city. During the time when data were being collected, detector configurations were being adjusted, complete documentation of the detection for the entire system was not immediately available, and dates when detection changed were not known for the entire time period. Although the detector assignments were not known, it was still possible to assess the quality of operation by examining the tendency of phases to use all of their assigned green time. Meanwhile, Dubuque deployed data collection at selected intersections, mostly along the US 20 corridor. These intersections had well-defined detection, with setback detection for measuring vehicle arrivals available at a few of the intersections.

These two data sets permitted two studies of high-resolution data for ranking intersection performance in Iowa. In the first study, high-resolution data from Cedar Rapids were used to rank locations according to movement and intersection capacity utilization, as well as estimated opportunities to improve by rebalancing splits. In the second study, high-resolution data from Dubuque were used to calculate measures of progression performance, and these were compared against the INRIX segment speed data to determine whether they are correlated.

4.2 Ranking of Intersections with ATSPM Data

To perform intersection ranking, high-resolution data from 143 intersections in Cedar Rapids were collected during the months of January to June of 2020. For this analysis, all of the Wednesdays in the data set were used. The high-resolution data contain a record of the reason for phase termination (e.g., why the end of green occurred). A signal phase roughly corresponds to a movement at an intersection (their association can be more nuanced, but for the purpose of this discussion, they will be considered roughly synonymous). There are four possible situations of a phase within every cycle:

- A phase may *gap out*, meaning that the green ends before the maximum amount of time has elapsed. This implies that there was a relatively low amount of demand in the previous cycle.
- A phase may *max out*, meaning that the green ends after it has been given the maximum amount of green time. This implies that there was a relatively high amount of demand in the previous cycle.

- A phase may *force off*, which has the same meaning as max out, except that the signal is coordinated, and the maximum green time is determined by the split of the phase rather than a maximum green parameter.
- A lack of any record of activity on the phase within a cycle indicates that the phase was skipped or omitted.

To assess the level of phase utilization, a basic comparison was made between the numbers of force-offs and max-outs (FOMOs) versus gap outs and phase omits/skips. Previous studies have shown that a comparison of phase termination distributions offers a way to obtain basic information about actuated signal operation without any knowledge of the detector configuration (Li et al. 2013). The first step was to identify all of the instances of each phase. A phase instance is defined as a pair of intervals delineated by the start and end of green, as shown in Figure 24.

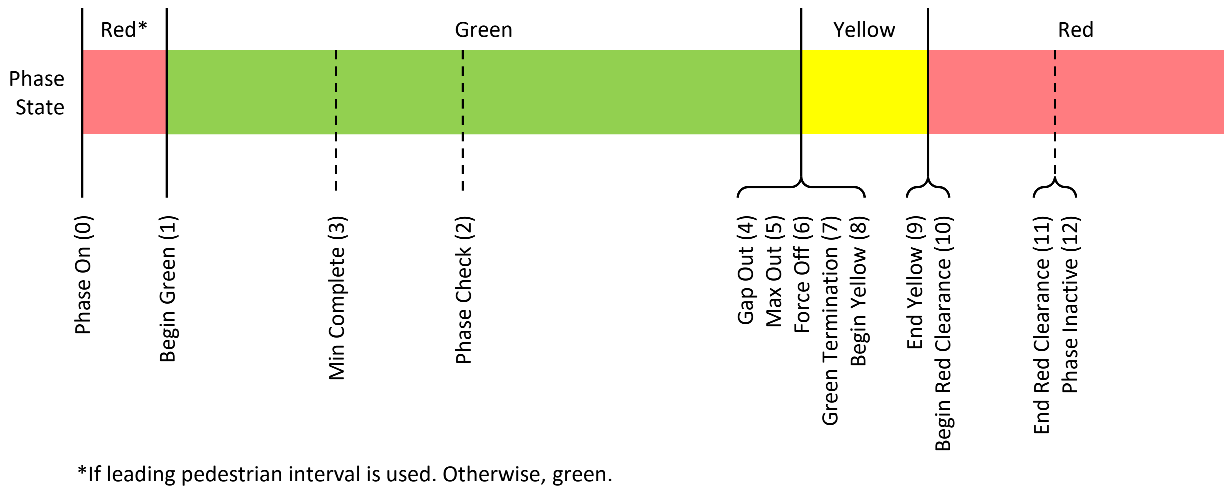


Figure 24. Association of phase state and events logged in high-resolution data

A red interval is assumed to lead to a green interval; this order is used because vehicles that arrive during red must wait until after the start of green to proceed through the intersection (Day et al. 2010a). Events 1 and 8 were used to identify the start and end of green. The yellow interval was included with the red interval for purposes of this analysis.

Next, the phase instances were matched to the termination events to identify whether the instance ended with a gap out, max out, or force off. In the Cedar Rapids data set, during all of the Wednesdays occurring across a six-month time frame, there were 13.7 million phase instances, of which 6.1 million ended in gap out, 1.5 million ended in max out, 6 million ended in force off, and about 50,000 ended without any phase termination code.

Omitted or skipped phases do not have any record, so to determine whether a phase is skipped, it is necessary to find the beginning and end of each cycle and find whether a beginning of a green event occurs within that interval for each phase. If so, that time can be used to find the relevant phase instance and termination code. If no beginning of green time is found, then the phase is either skipped or omitted.

To find cycle boundaries, the general structure of the phase and ring assignments must be known. For this study, it was assumed that all of the intersections in the Cedar Rapids area follow the dual-ring, eight-phase configuration. First, an ordered list of phase instance times was assembled. From this list, it was possible to see the order in which phases are served. Cycle boundaries were identified when the signal controller transitions from phases {1, 2, 5, 6} to phases {3, 4, 7, 8}. The reason for this is that the coordinated or major street through phases are usually served within group {1, 2, 5, 6}. Although that is not always true, this definition generally works well to find cycle boundaries. There is also a barrier crossing event, but in the researchers' experience, this event is less reliable.

Once the cycle times were known, it was possible to then match phase instances to cycles and find whether each phase was skipped/omitted or whether it gapped out, maxed out, or forced off in each cycle. Next, to simplify the remaining ranking tasks, the percentages of phase skips/omits, gap outs, and FOMOs were calculated by hour.

Figure 25 shows the distribution of omits (skip/omit), gap outs, and FOMOs for all the Wednesdays across the six-month period, for two example intersections exhibiting rather different behavior.

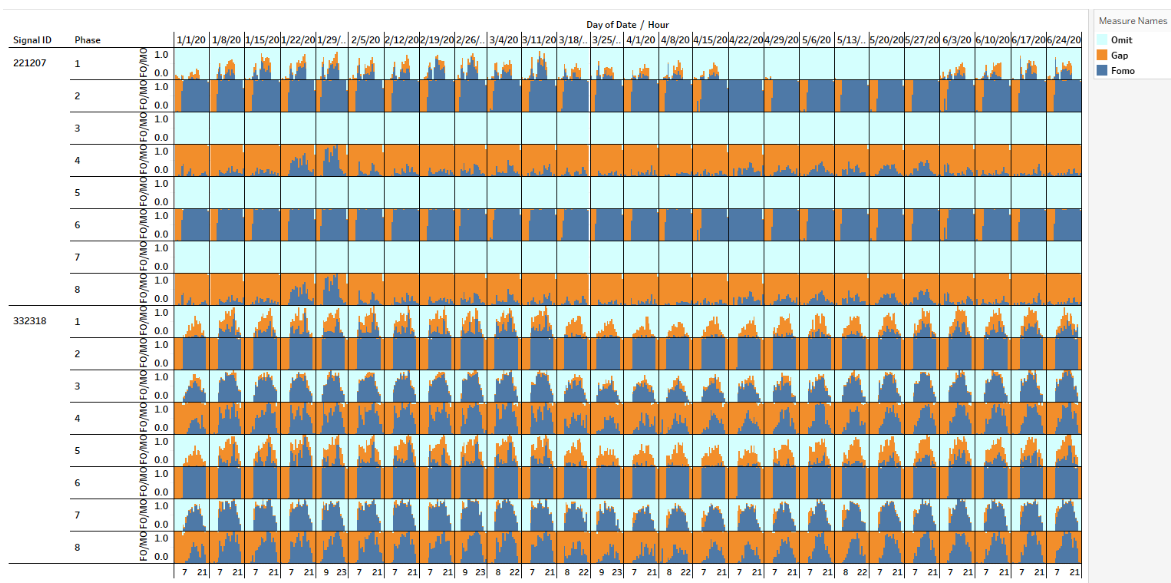


Figure 25. Distribution of phase terminations by date, and by phase, at two intersections

For Signal 221207, phases 3, 5, and 7 were always omitted, so these phases probably do not exist at this intersection. Phase 1 was occasionally served, but it was often omitted. Phases 4 and 8 frequently gapped out, although they had FOMO events at times, particularly in the middle of the day. Finally, phases 2 and 6 had rather typical coordinated phase behavior, with constant FOMO during the day—given coordinated phases typically serve the maximum amount of green time (except when early yield is used, which did not seem to be the case at this intersection)—and many gap-out events during non-coordinated, early morning operation. Signal 332318 exhibited similar dynamics, but all eight phases were in use.

The next step in the analysis was to determine which phases had useful data for analysis. Phases without any variation in their phase termination events are typically not very useful, because this lack of variation is due to some cause such as the following:

- Phases that are constantly omitted are not in use.
- Phases that constantly FOMO are operating in max recall, meaning they are given the maximum amount of green time in each cycle for some reason. One reason may be that the phases are coordinated, which as previously mentioned in the example is a typical configuration of coordinated phases. It is also possible that the phase might have a detector error placing a constant call. Finally, the phase might operate under max recall by design, e.g., at a pretimed or semi-actuated signal.

The phases that were absent for a particular intersection were excluded. To accomplish this, the data set was filtered to exclude all phases for which percentages of being omitted were equal to 100% throughout the entire data set. After excluding these not-in-use phases, the phases that were coordinated or under max recall were identified. The percentages of FOMO for such phases will mostly be higher, but not necessarily 100% at all times of day. The reason for this is that such phases are often operated differently during low-volume conditions. For example, in late night and early morning hours, intersections are often allowed to run in a free or fully actuated mode, and in that case, the phase that is ordinarily coordinated will be allowed to gap out when there is no demand. Figure 26 shows the phase termination chart for phase 2 at 10 different signals.

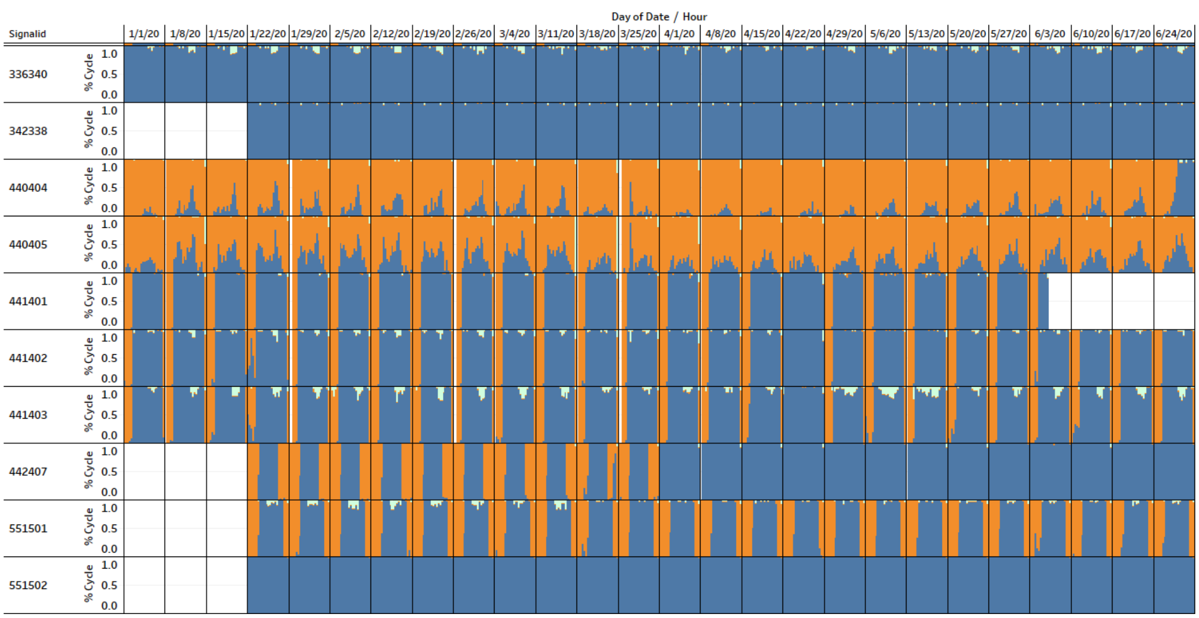


Figure 26. Examples of phase 2 termination for identifying coordinated phases

Signal 551502 appeared to have phase 2 in max recall, given it constantly FOMOs. Signal 441401 had typical coordinated behavior with FOMO during most of the day and gap out at night. Signals 440404 and 440405 appeared to be fully actuated at all times of day.

When the percentage of FOMO exceeded 80% for 12 or more consecutive hours, the phase was considered to be either coordinated or under max recall. These criteria appeared to effectively identify phases that should be excluded, based on visual analysis of the performance of the excluded phases.

Another possibility that needed to be considered was that phases might have encountered a detector failure at some time during the study period. In this case, it would have been expected that the phase would exhibit normal behavior but suddenly fall into a pattern of constant FOMO. In Figure 26, signal 442407 seemed to exhibit this type of behavior, with constant FOMO occurring starting from April 1, 2020. Because this phase was identified as a coordinated phase based on prior criteria, it was already excluded from the data set. To identify a phase affected by detector error, the research team searched for any phase for which FOMO was equal to 100% for more than 24 consecutive hours. However, the team did not identify any phases exhibiting such behavior that were not already excluded from the analysis in previous steps.

After excluding omitted phases, coordinated/max recall phases, and phases affected by detector errors, the researchers proceeded with the intersection ranking. A separate ranking was undertaken for three time-of-day periods: a.m. peak (6:00 a.m.–9:00 a.m.), midday (9:00 a.m.–3:00 p.m.), and p.m. peak (3:00 p.m.–7:00 p.m.). This analysis was possible for 130 intersections. Eighteen intersections had to be excluded from the ranking because they had no phases remaining to be analyzed after the previous phase exclusions.

Two criteria were employed for establishing criteria to rank the intersections. The first of these identified the worst-performing phase, whereas the second identified the overall capacity utilization. The intent was to locate problems using the first criterion and assess overall intersection utilization with the second criterion. Later, the two criteria were combined to identify opportunities for improvement by finding intersections with heavy utilization of any phase, but with low overall intersection utilization. The idea was inspired by a previously developed methodology that used v/c ratio and degree of intersection saturation (Day et al. 2010a), but given the detector assignments were not known for many of the intersections in this case, the concept was applied to the distributions of phase termination.

The first criterion was the average maximum percentage of FOMO of any phase at each intersection (i.e., the worst-performing phase). The FOMO percentages were calculated for each time-of-day period for each day in the data set. Next, the average rate of FOMO was calculated by taking the mean value across all of the dates for each phase. Finally, the maximum value was taken of all the phases at the intersection. The purpose of the maximum value was to allow any phase exhibiting capacity deficiencies to stand out rather than be lost in an intersection average.

The second criterion was the average number of phases at each intersection having FOMO percentages greater than 50% (i.e., overall intersection utilization). The FOMO percentages were again calculated for each time-of-day period for each day in the data set. Next, the research team determined the percentage of phases for which this value exceeded 50%. Finally, the average of this value was taken across the data in the data set to obtain the final number for ranking.

The outcomes of the intersection ranking are presented in the next few figures and tables. The first criterion found locations of the worst-performing phase in the system by time of day. Figure 27 shows a Pareto diagram of the ranking value by intersection, for each time of day separately.

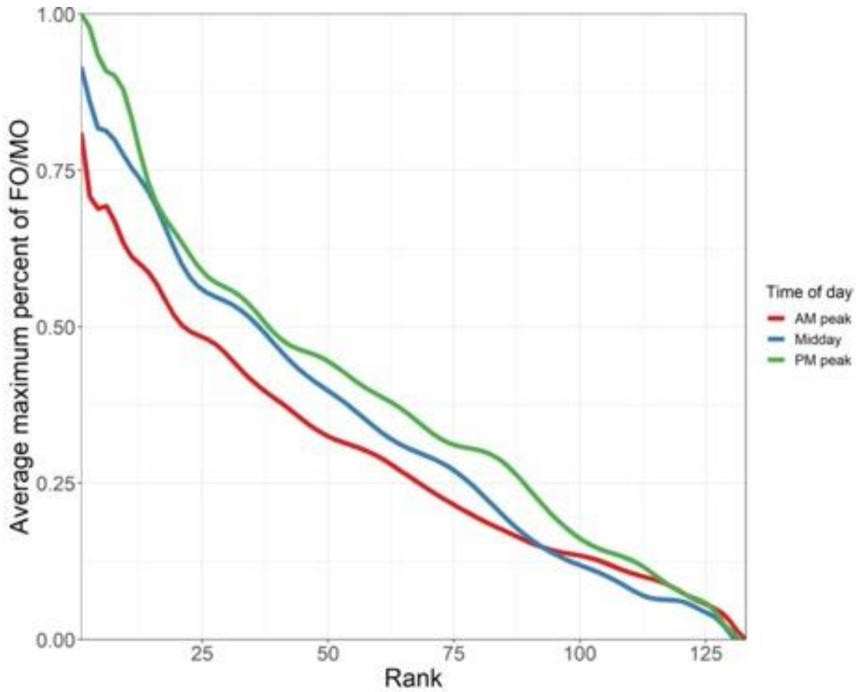


Figure 27. Pareto diagram of average maximum percent FOMO by intersection

A Pareto diagram is simply a ranked list of the resulting values for each data series, without any connection between the rank of the item among the different data series. This offers a way to view the distribution of values. Often, such charts exhibit a Pareto principle, also known as the 80/20 principle, wherein 80% of the activity is attributable to 20% of the population. In this case, there does seem to be a very slight inflection point around the upper quintile. That is, the slope is slightly steeper to the left of the vertical line at rank 25 (approximately the upper quintile). In addition, looking at the placement of the lines, it seems that the midday has more congestion than the a.m. peak, while the p.m. peak has still more congestion.

Table 4 provides a listing of the top five intersections by time of day according to the average maximum percent FOMO.

Table 4. Intersection ranking results based on average maximum percent FOMO

TOD	Intersection	Mean value	Rank
a.m. peak	Blairs Ferry Road NE & C Avenue NE	81.1%	1
	Collins Road NE & Lindale Drive NE	74.7%	2
	Collins Road NE & C Avenue NE	69.7%	3
	Edgewood Road NE & 42nd Street NE	69.2%	4
	29th Street NE & I-380	69.1%	5
Midday	Blairs Ferry Road NE & C Avenue NE	91.6%	1
	Collins Road NE & C Avenue NE	89.9%	2
	Collins Road NE & Twixt Town NE	84.4%	3
	Collins Road NE & Council Street NE	82.2%	4
	Collins Road NE & Lindale Drive NE	81.6%	5
p.m. peak	29th Street NE & Prairie Drive NE	100.0%	1
	1st Avenue E & Lindale/Home Depot	99.1%	2
	Blairs Ferry Road NE & C Avenue NE	96.5%	3
	Collins Road NE & C Avenue NE	94.7%	4
	Blairs Ferry Road NE & 10th Avenue NE	92.1%	5

The results from Table 4 were largely found to align well with experiences of the traffic engineering staff in Cedar Rapids, although a few intersections were unexpected. The 29th Street NE and I-380 a.m. peak was not expected and would warrant deeper investigation. The intersection of 29th Street NE and Prairie Drive was a marginally warranted signal. High values at the intersections along 29th Street were likely affected by area construction.

A Pareto diagram for the second criterion is shown in Figure 28.

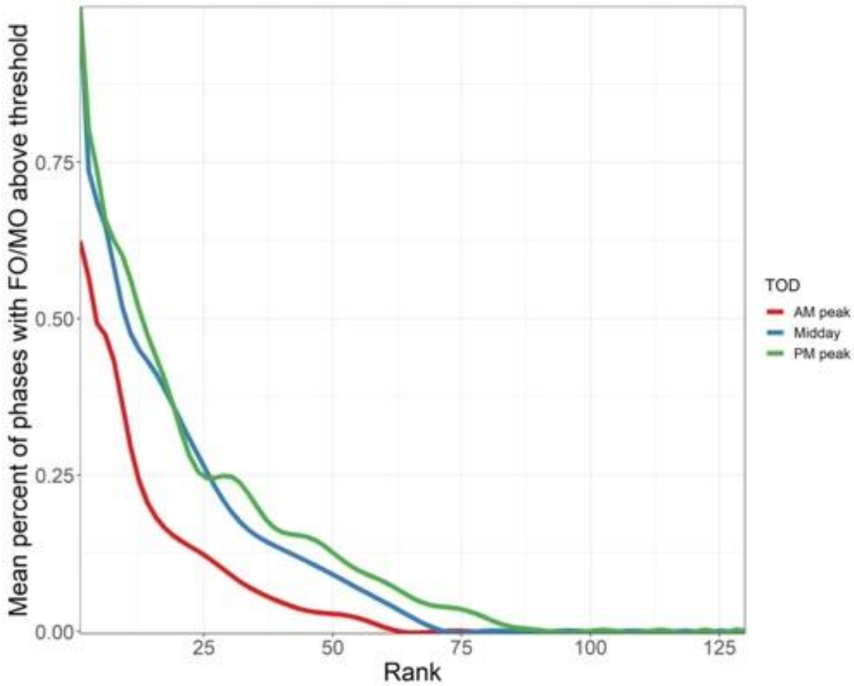


Figure 28. Pareto diagram of average number of phases with FOMO above 50%

Figure 28 represents the overall level of utilization of each intersection, as opposed to the worst-performing phase at the intersection. The diagram also appears to exhibit the Pareto principle with the slope of the line taking an upward turn to the left of the first quintile (around Rank 25). The top five intersections by time of day are shown in Table 5.

Table 5. Intersection ranking using average number of phases with FOMO above 50%

TOD	Intersection	Mean value	Rank
a.m. peak	Blairs Ferry Road NE & I-380	62.5%	1
	A-B Avenue NE & 1st Street NE	61.1%	2
	Collins Road NE & C Avenue NE	55.1%	3
	Wilson Avenue SW & I-380	50.0%	4
	29th Street NE & Prairie Drive NE	47.2%	5
Midday	A-B Avenue NE & 1st Street NE	100.0%	1
	Collins Road NE & Twixt Town NE	78.8%	2
	Collins Road NE & Lindale Drive NE	73.7%	3
	Edgewood Road NE & Highway 100	68.1%	4
	Collins Road NE & C Avenue NE	66.7%	5
p.m. peak	A-B Avenue NE & 1st Street NE	100.0%	1
	Collins Road NE & Lindale Drive NE	80.5%	2
	Collins Road NE & Twixt Town NE	79.7%	3
	Collins Road NE & C Avenue NE	77.7%	4
	Collins Road NE & Council Street NE	66.5%	5

Table 5 shows the busiest intersections at each time of day. There are some intersections with very high utilization; during the midday and a.m. peak, one of them has 100% utilization according to this criterion. The results shown in the table also largely agreed with experiences of traffic engineering staff in Cedar Rapids, with the exception of the 19th Street intersection which, as mentioned previously, was likely affected by construction during the study period. The A-B Avenue location is another anomaly, which represents a single controller operating two intersections that will be removed in the future.

After developing the rankings using the previously described criteria, the researchers identified the intersections where possibilities for improvement are likely to exist. In general, this is expected to be the case in situations where an intersection has a low overall utilization (the second criterion, mean percent of phases with FOMO above threshold)—meaning there is green time that can be redistributed—and at least one phase with high utilization (the first criterion, high average maximum percent FOMO)—meaning there is one phases that needs more green time.

Figure 29 shows a diagram of the two criteria.

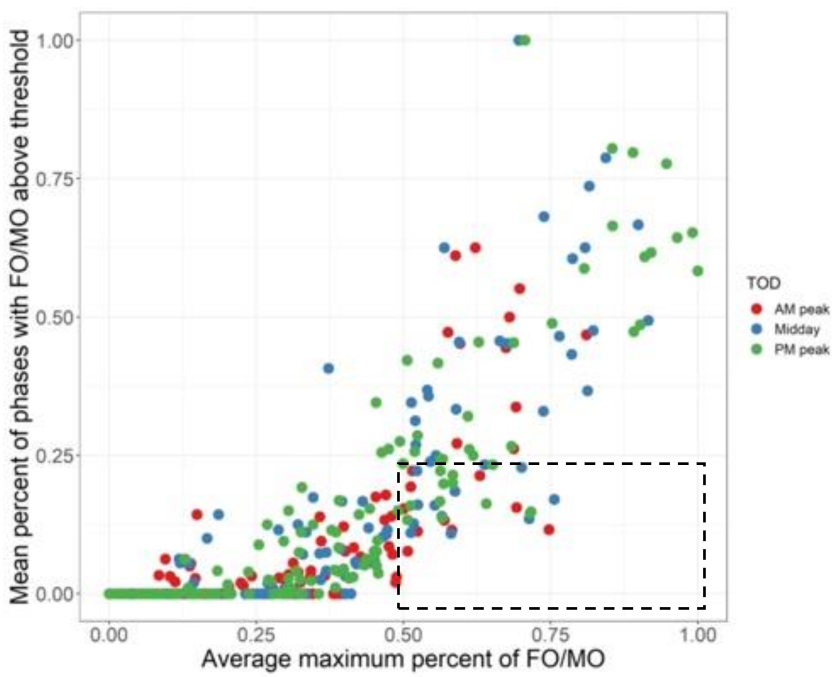


Figure 29. Scatterplot comparing the two criteria

Each dot represents the result for one intersection, for a particular time of day. The horizontal axis represents the worst-performing phase at the intersection, while the vertical axis represents the overall intersection utilization. Points that are closer to the lower right portion of the curve represent situations where there are phases with high utilization, yet the intersection has low utilization. To find candidate locations, the research team filtered the intersections by average maximum percent FOMO greater than 50% and mean percent of phases above threshold less than 25%, as indicated by the dashed line in Figure 29.

Next, the data were examined for each of these to find that each intersection had both one phase with high utilization along with another phase with low utilization that could potentially give up green time to the high-utilization phase. For this study, these exchanges were limited to phase pairs within phase concurrency groups (i.e., the phase pairs {1,2}, {5,6}, {3,4}, and {7,8}). Table 6 presents the results of this analysis.

Table 6. Phase pairs warranting additional investigation for potential split rebalancing

Intersection	Phase pairs
10th Street SE & 3rd Avenue SW	1-2, 3-4
Edgewood Road NW & F Avenue NW	1-2, 5-6
Kirkwood Boulevard SW & US 30	1-2, 5-6
6th Street SW & 33rd Avenue SW	3-4, 7-8
32nd Street NE & Oakland Road NE	1-2, 5-6
Blairs Ferry Road NE & 10th Avenue NE	1-2, 5-6
C Avenue NE & Boyson Road NE	1-2, 5-6
Blairs Ferry Road NE & Council Street NE	1-2, 5-6
Collins Road NE & Northland Avenue NE	5-6
Collins Road NE & Council Street NE	5-6

The list in Table 6 shows 10 intersections where the analysis indicates that there are pairs of phases for which operation would be worthwhile to examine in greater detail to identify opportunities for improvement by rebalancing splits.

4.3 Comparison of ATSPM and INRIX Travel Times

To gain additional insight about the use of INRIX data for signalized arterials, the researchers compared performance measures with the two data sets to identify whether there was any correlation between the average minute speeds in the INRIX data and the measures of the quality of progression in the high-resolution data. Correlation between the two data sets can indicate the suitability of using INRIX data for corridor-level analysis in the absence of high-resolution data. The research team estimated a series of econometric models to investigate correlation between these variables. Data from January and February of 2020 along US 20 in Dubuque, Iowa were employed.

US 20 in Dubuque is one of the busiest roadways in that region of the state, connecting one of two Mississippi River crossings with points west of Dubuque. The roadway passes through eight signalized intersections, as shown in Figure 30.

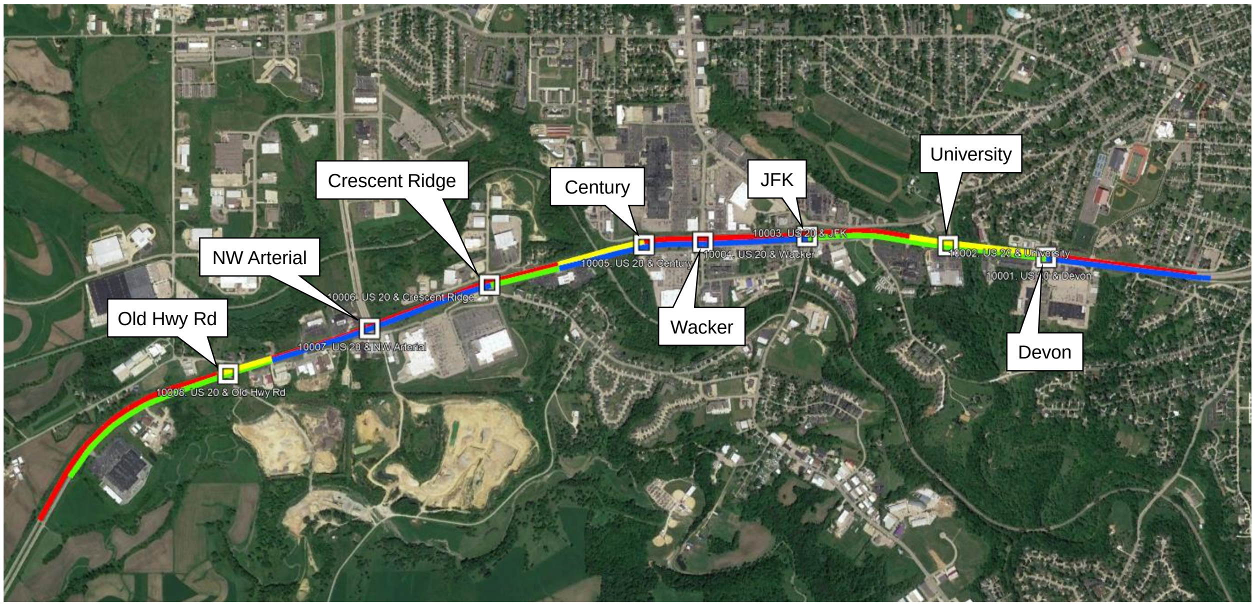


Figure 30.Intersections with high-resolution data and INRIX XD segments on US 20 in Dubuque

Figure 30 also shows the locations of the INRIX XD segments that pass through those eight intersections. After receiving high-resolution data and detector mapping information from the city, the research team first investigated the quality of performance measures that could be extracted from these data. For this study, the focus was on corridor operation, so the team examined performance measures aimed at evaluating coordination, specifically the POG. To ascertain whether the detector data existed to permit this analysis, the team began by creating coordination diagrams (Day et al. 2010b) for each of the approaches among the eight intersections.

Figure 31 presents one of these diagrams for eastbound US 20 and University Avenue; this intersection exhibits very good coordination, with most of the vehicle arrivals (dots) occurring during the green intervals (green shaded regions).

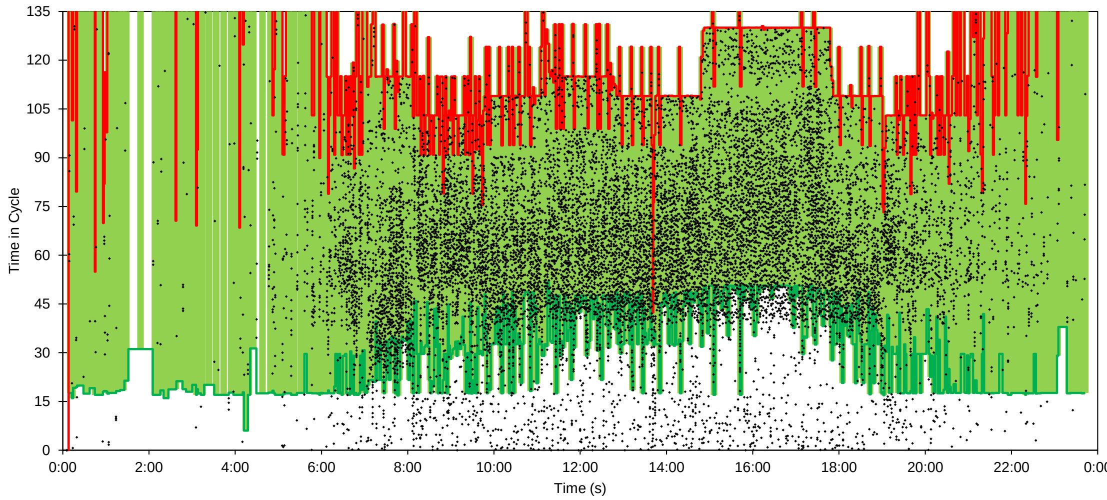


Figure 31. Coordination diagram from westbound US 20 and University Avenue, January 28, 2021

The dots in this figure are dense, with counts that are of the same order of magnitude as the roadway AADT, and this location can be used for further analysis.

Figure 32, in contrast, shows a very irregular pattern.

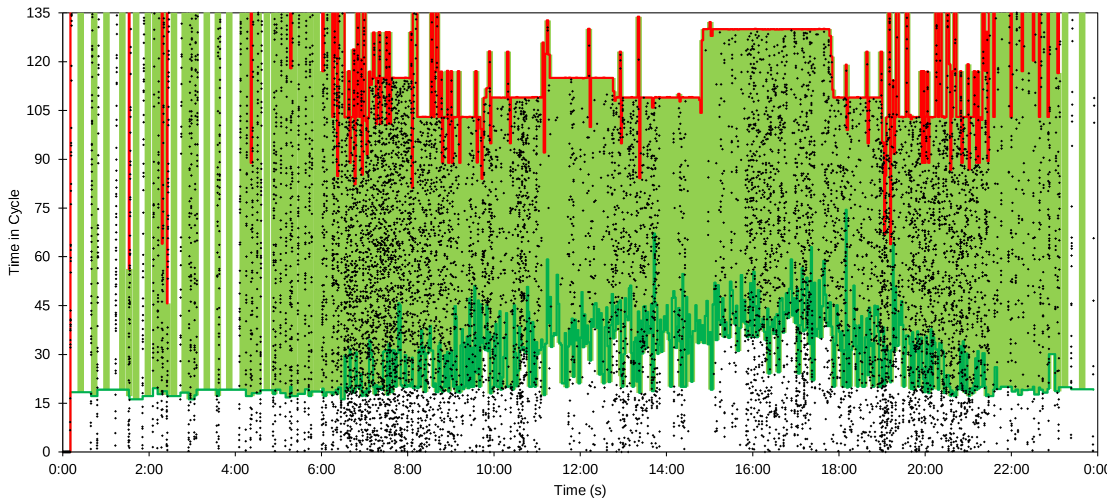


Figure 32. Coordination diagram from eastbound US 20 and Century Drive, January 28, 2021

The figure shows that there are periods with no detection, which is not likely to be due to a lack of traffic during those time periods. The arrivals appear to be random, but this is unlikely because there is a neighboring coordinated intersection within 1 mile. Finally, there are high amounts of detections during the early morning period. In summary, these detector data are not believable.

After examining the setback detector data for all of the intersections, the research team found that the two easternmost intersections in the system (US 20 and University Avenue and US 20 and Devon Drive) appeared to have good detection in both directions. These two intersections are associated with three INRIX segments, as shown in Figure 30. One segment in the eastbound direction passes through both of the intersections, and in the westbound direction, there is one segment passing through each intersection. US 20 and Devon Drive is the first intersection of the corridor in the westbound direction, and it has random arrivals, while US 20 and University Avenue is the next intersection and has platoons (as shown in Figure 31). In the eastbound direction, both of the approaches have platoons, because they are at the end of the corridor in that direction. Thus, there are three different comparison pairs.

Data from all days of January and February were used for the comparison. The day was divided into 5- and 15-minute intervals for which all of the dependent and independent variables were calculated. This eliminated issues with some individual minutes of speed data missing observations throughout the day, as well as variations in the cycle length by time of day. For each interval, three different performance measures were calculated from the high-resolution data as follows:

- The POG was calculated by counting the number of arrivals (detections) occurring during green and dividing by the total number of arrivals during the interval.
- The v/c ratio was calculated by dividing the total number of arrivals during the interval by the capacity, determined by the total amount of green time occurring in the interval multiplied by the saturation flow rate. For this study, a saturation flow rate of 1,800 vehicles/hour (0.5 vehicles/second) was used.
- The percent of time that the signal was green was calculated by dividing the total amount of green time in the interval by the total interval duration. This is equivalent to the green-to-cycle (g/C) ratio used in the HCM analysis.

In addition to these, indicator variables were established for the day of week and time of day. The time-of-day variables were determined using changes in the traffic pattern, which were determined from the coordination events that are written in the high-resolution data when a new time-of-day plan is implemented by the controller. The dependent variable for this comparison was the average speed. To calculate speed, only those records of the INRIX data having a score of 30 (indicating actual observations) were included. These data sets were then compiled into tables containing all of the variables for all the 5- and 15-minute periods in January and February.

The independent variables are summarized in Table 7.

Table 7. Independent variables for INRIX and high-resolution data comparison

No.	Independent variable	Type	Code/Range
1	POG	Quantitative	0–1
2	Day of week	Categorical	Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday
3	Time of day	Categorical	5, 9, 19, 37, 45, 55, 64, 254
4	v/c ratio	Quantitative	0–1.3
5	Green duration (%)	Quantitative	0–1

Figure 33 shows a scatterplot and correlation matrix for all the quantitative variables in the data set for the westbound segment passing through US 20 and University Avenue.

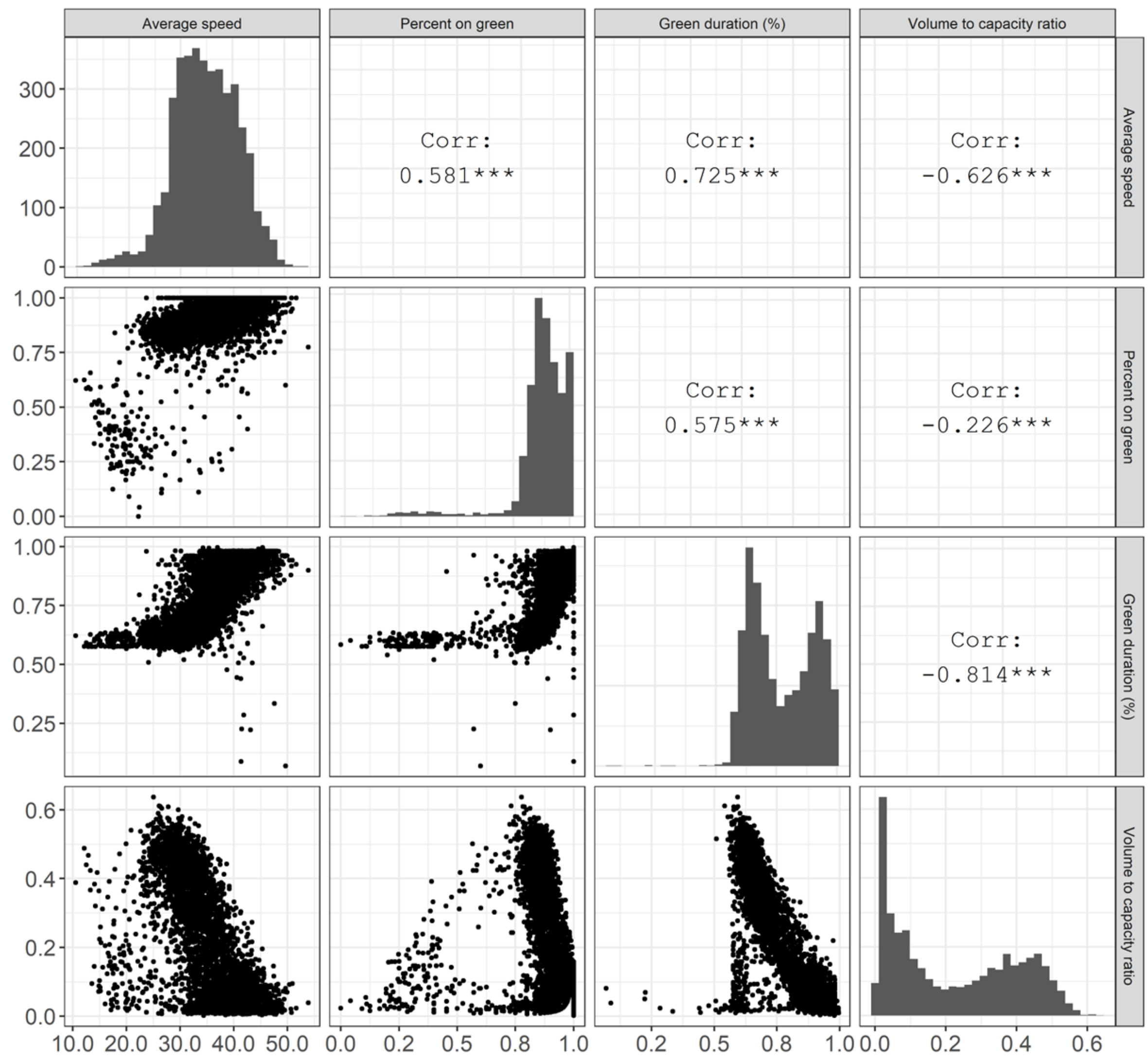


Figure 33. Scatterplots and distributions of quantitative variables

Figure 33 is arranged with each square showing a comparison between variables as indicated by the labels at the top and right sides of the diagram. The diagonals show histograms that represent the spread of the data. The lower left squares show scatterplots that illustrate the potential correlation between variables. The upper right squares show the correlation coefficients. The magnitude of this number represents the strength of the linear relationship between the two variables, while the sign of the number indicates whether the variables are positively or inversely correlated.

These results show that the duration of green is highly correlated to the POG and to the v/c ratio. Similar results were obtained for other segments. Because these variables are likely to contribute similar information to the model, the research team included only one of them in the regression models. In addition, it is apparent that the average speed correlates with the independent variables.

Next, the data were organized to begin creating regression models. Because there were three INRIX segments, three models were created. However, because one of the INRIX segments contained two intersections, there were various options for combining the independent variable data for those two intersections (e.g., whether to include them as separate variables or combine them into an average value). Because the eastbound INRIX segment passes through two intersections, the research team developed two models with different sets of variables. In one case, the team considered separate quantitative variables for the two intersections and in another case average values of the variables were considered. Thus, the team ended up with five models altogether.

The scatterplot of INRIX average speed versus POG shows that there is a positive linear relationship between the two variables. Therefore, a starting point for the analysis was to consider a very simple linear model for average speed as a function of POG (i.e., with one independent variable). However, the R^2 values obtained for these simple models were not very high. This indicated that a very low percentage of variation in the observed values of average speed could be explained by the linear regression model with POG. To improve the models, the day-of-week indicator variable was included, which yielded minor improvement of the R^2 values. Next, the time-of-day indicator variable was added, which greatly increased the R^2 values of the models in all cases. Two additional models were considered by adding v/c ratio and green duration to the previous model. The five regression models that were considered are shown in Table 8**Error! Reference source not found.**.

Table 8. Regression models

No.	Model
1	Average speed ~ POG
2	Average speed ~ POG + Day of week
3	Average speed ~ POG + Day of week + Time of day
4	Average speed ~ POG + Day of week + Time of day + v/c ratio
5	Average speed ~ POG + Day of week + Time of day + Green duration

Models with interval lengths of 15 minutes performed better in terms of R^2 values compared to those with 5-minute interval lengths. Model outputs for interval lengths of 15 minutes are shown in the following tables.

Table 9 shows the regression model outputs for the westbound segment passing through the intersection of US 20 and Devon Drive.

Table 9. Regression model outputs (US 20 and Devon westbound segment)

	Dependent variable:				
	Average speed				
Independent variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
POG	10.58***	10.70***	6.53***	5.13***	
Day: Monday		0.99***	0.28	0.24	0.33
Day: Saturday		1.20***	-0.13	0.03	-0.07
Day: Sunday		2.84***	1.47***	1.19***	1.54***
Day: Thursday		-0.12	-0.02	0.09	0.06
Day: Tuesday		1.22***	0.51**	0.56***	0.53**
Day: Wednesday		1.55***	0.78***	0.81***	0.84***
Timing plan: 5			-4.00***	-1.96***	-3.92***
Timing plan: 9			-4.48***	-2.52***	-4.50***
Timing plan: 19			-2.93***	-0.40*	-2.81***
Timing plan: 37			-3.40***	-0.66*	-3.27***
Timing plan: 45			-1.79***	0.33	-1.65***
Timing plan: 55			-5.39***	-2.58***	-5.28***
Timing plan: 64			-8.24***	-8.53***	-8.18***
Timing plan: 254			0.93***	-0.49***	0.72***
v/c ratio				-6.76***	
Green duration (%)					8.15***
Intercept	31.99***	30.86***	35.36***	38.10***	34.57***
Observations	4,297	4,297	4,297	4,297	4,297
R^2	0.09	0.13	0.37	0.41	0.37
Adjusted R^2	0.09	0.12	0.36	0.40	0.37
Residual Std. Error	4.46	4.37	3.72	3.60	3.70
F Statistic	410.47***	87.65***	165.48***	183.02***	169.83***

*p<0.1; **p< 0.05; ***p<0.01

The R^2 value for Model 1 is only 0.09, which means that the model with POG cannot explain the variations in average speed very well. The addition of other categorical and quantitative variables to this model improved the value of R^2. R^2 was the highest for Model 4 with two categorical and two quantitative variables. All the variables considered were statistically significant at 0.05 level of significance. For all the models, p-values from F-tests were statistically significant, which indicates there is strong evidence that the regression models are significant in explaining the variation in average speed.

The model outputs for the westbound segment passing through the intersection of US 20 and University Avenue are shown in Table 10.

Table 10. Regression model outputs (US 20 and University westbound segment)

	Dependent variable:				
	Average speed				
Independent variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
POG	28.92^{***}	28.46^{***}	14.51^{***}	6.28^{***}	
Day: Monday		0.18	0.26	-0.02	0.28
Day: Saturday		0.26	-0.83^{***}	-0.37^{*}	-0.61^{***}
Day: Sunday		1.71^{***}	0.88^{***}	0.40^{*}	0.70^{***}
Day: Thursday		0.55^{*}	0.23	0.48^{**}	0.37^{*}
Day: Tuesday		0.55^{*}	0.64^{***}	0.81^{***}	0.70^{***}
Day: Wednesday		0.71^{**}	0.86^{***}	1.05^{***}	0.94^{***}
Timing plan: 5			-5.39^{***}	-1.11^{***}	-2.94^{***}
Timing plan: 9			-6.87^{***}	-1.35^{***}	-4.22^{***}
Timing plan: 19			-4.77^{***}	-0.96^{***}	-2.32^{***}
Timing plan: 37			-4.89^{***}	-0.69^{*}	-2.61^{***}
Timing plan: 45			-2.34^{***}	0.07	-1.88^{***}
Timing plan: 55			-7.00^{***}	-2.00^{***}	-4.41^{***}
Timing plan: 64			-7.68^{***}	-12.74^{***}	-11.29^{***}
Timing plan: 254			2.99^{***}	0.61^{***}	1.29^{***}
v/c ratio				-19.12^{***}	
Green duration (%)					20.27^{***}
Intercept	9.37^{***}	9.21^{***}	23.49^{***}	33.97^{***}	20.08^{***}
Observations	4,042	4,042	4,042	4,042	4,042
R^2	0.34	0.34	0.58	0.64	0.63
Adjusted R^2	0.34	0.34	0.58	0.64	0.62
Residual Std. Error	5.05	5.03	4.02	3.72	3.80
F Statistic	$2,057.30^{***}$	302.41^{***}	374.25^{***}	452.64^{***}	448.96^{***}

*p<0.1; **p< 0.05; ***p<0.01

For this segment, the values of R^2 obtained are comparatively better than those obtained for the previous segment. The results were similar in terms of significance. P-values from F-tests were statistically significant for all the models as before. Model outputs also showed that each of the variables is statistically significant in predicting average speed.

Table 11 and Table 12 show the regression model outputs with different sets of quantitative variables for the eastbound segment that contains both the intersections.

Table 11. Regression model outputs with separate quantitative independent variables for each intersection (eastbound segment with two intersections)

Independent variable	*Dependent variable:* Average speed				
	Model 1	Model 2	Model 3	Model 4	Model 5
POG (Devon)	0.76^{*}	0.75^{*}	2.11^{***}	1.95^{***}	
POG (University)	34.13^{***}	33.74^{***}	27.77^{***}	18.67^{***}	
Day: Monday		1.37^{***}	0.63^{***}	0.74^{***}	0.94^{***}
Day: Saturday		1.16^{***}	-0.23	0.23	-0.20
Day: Sunday		1.46^{***}	0.45^{**}	0.31	1.24^{***}
Day: Thursday		0.36	0.58^{***}	0.79^{***}	0.53^{**}
Day: Tuesday		1.41^{***}	0.61^{***}	0.78^{***}	0.69^{***}
Day: Wednesday		1.08^{***}	0.33	0.68^{***}	0.63^{***}
Timing plan: 5			-0.21	1.40^{***}	-2.94^{***}
Timing plan: 9			-4.37^{***}	-1.39^{***}	-5.29^{***}
Timing plan: 19			-1.32^{***}	2.95^{***}	3.00^{***}
Timing plan: 37			-0.22	3.77^{***}	1.03^{**}
Timing plan: 45			0.97^{**}	2.37^{***}	0.86
Timing plan: 55			-3.58^{***}	0.62	-2.75^{***}
Timing plan: 64			-8.75^{***}	-10.54^{***}	-9.21^{***}
Timing plan: 254			3.20^{***}	1.08^{***}	2.81^{***}
v/c ratio (Devon)				-7.72^{***}	
v/c ratio (University)				-8.19^{***}	
Green duration (Devon)					6.80^{***}
Green duration (University)					17.42^{***}
Intercept	5.33^{***}	4.70^{***}	9.53^{***}	20.71^{***}	16.91^{***}
Observations	3,996	3,996	3,996	3,996	3,996
R^2	0.64	0.65	0.74	0.77	0.69
Adjusted R^2	0.64	0.64	0.74	0.77	0.69
Residual Std. Error	4.29	4.26	3.65	3.45	3.96
F Statistic	$3{,}548.77^{***}$	907.19^{***}	708.97^{***}	731.73^{***}	563.46^{***}

*p<0.1; **p< 0.05; ***p<0.01

Table 12. Regression model outputs using average values of quantitative independent variables (eastbound segment with two intersections)

	Dependent variable:				
	Average speed				
Independent variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
POG	25.26***	25.02***	14.26***	8.83***	
Day: Monday		1.94***	0.86***	0.75***	0.75***
Day: Saturday		3.19***	-0.19	0.49**	-0.19
Day: Sunday		4.60***	1.38***	0.73***	1.34***
Day: Thursday		0.85**	0.37	0.66***	0.25
Day: Tuesday		1.95***	0.84***	0.99***	0.74***
Day: Wednesday		1.90***	0.72***	1.04***	0.76***
Timing plan: 5			-7.74***	-2.01***	-6.15***
Timing plan: 9			-12.03***	-4.10***	-9.19***
Timing plan: 19			-3.79***	3.38***	-0.17
Timing plan: 37			-8.03***	1.44***	-3.92***
Timing plan: 45			-6.46***	-0.77	-2.73***
Timing plan: 55			-11.59***	-2.06***	-7.36***
Timing plan: 64			-11.45***	-12.60***	-10.74***
Timing plan: 254			4.35***	0.81***	3.63***
v/c ratio				-22.65***	
Green duration (%)					16.90***
Intercept	14.54***	12.71***	24.77***	33.14***	23.71***
Observations	4,342	4,342	4,342	4,342	4,342
R^2	0.18	0.21	0.66	0.74	0.67
Adjusted R^2	0.18	0.21	0.66	0.74	0.67
Residual Std. Error	6.63	6.49	4.24	3.76	4.21
F Statistic	933.34***	166.80***	569.40***	754.42***	583.80***

*p<0.1; **p< 0.05; ***p<0.01

The models in Table 11 were developed by considering separate independent quantitative variables for the two intersections. Here, slope estimates of POG for the intersection of US 20 and University Avenue are much higher than those for the intersection of US 20 and Devon Drive. This indicates that a 1% increase in POG for the intersection of US 20 and University Avenue results in a higher increase in the average speed for this segment. One possible reason could be the extent of the segment passing through these intersections. The values of R^2 were higher than 0.6 in all the models, which indicated that there is strong linear association between average speed and the independent variables in consideration. P-values show that all the variables and models were statistically significant.

Similar results were obtained when average values of quantitative independent variables were taken as shown in Table 12.

For the simple model, the value of R^2 was very low, but as new variables were added, R^2 values increased and attained a maximum value of 0.74 for Model 4. For all models, the researchers obtained statistically significant p-values for all variables and F-tests.

Slope estimates in all the models developed in this chapter indicated that the average speed decreases with the increase of v/c ratio and increases with the increase of POG and green duration. In all cases, Model 4 with four independent variables namely, POG, time of day, day of week, and v/c ratio performed better in predicting average speeds in terms of the value of R^2. P-values from all F-tests were statistically significant at 0.01 level of significance, which indicated that the independent variables developed using high-resolution data did a better job in explaining the variability of INRIX average speed.

5. METHODOLOGY FOR OPERATIONAL IMPROVEMENT CANDIDATE LIST

5.1. Introduction

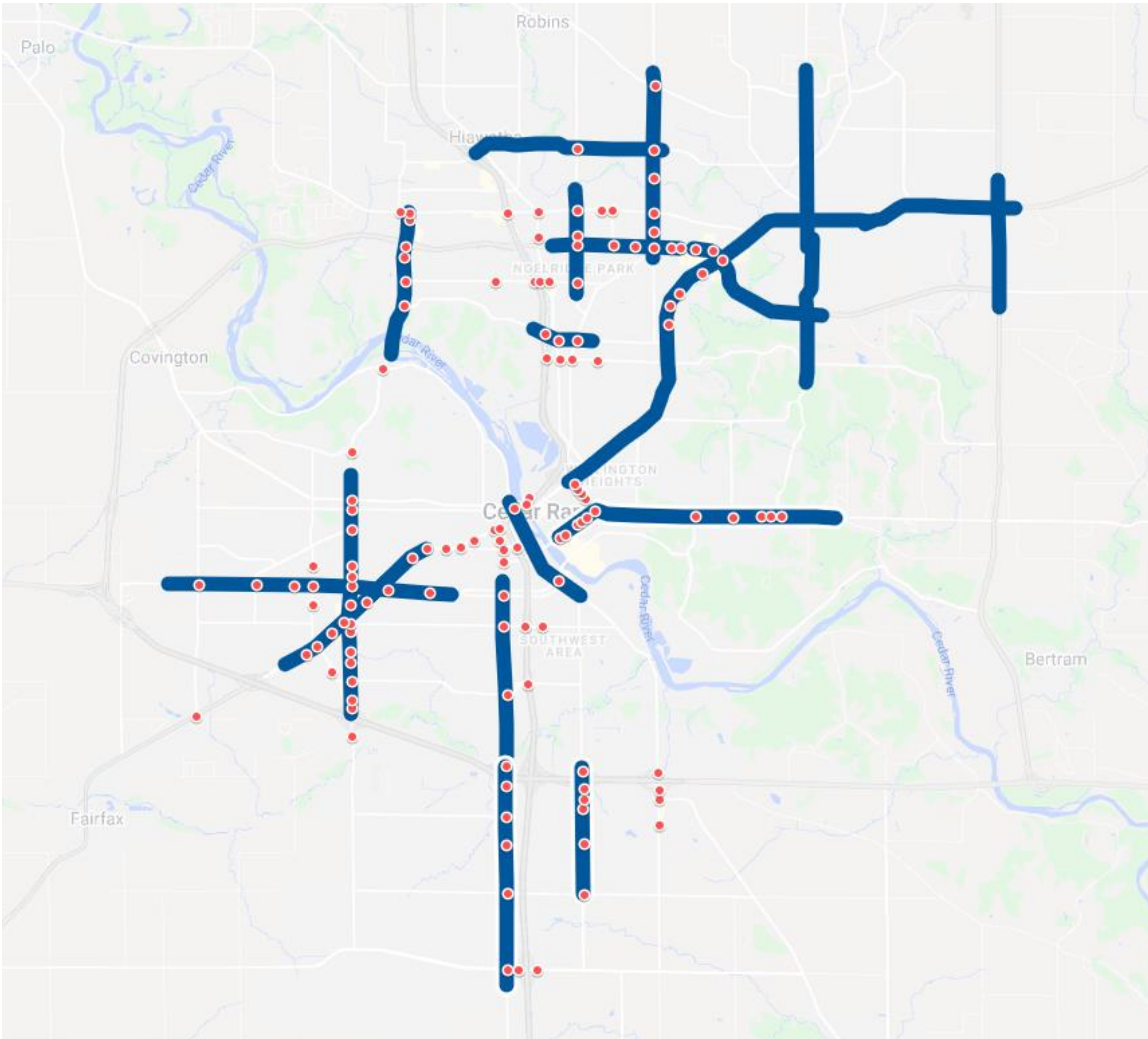
The previous chapters discussed the use of segment speed data (in this report, data provided by INRIX, which the Iowa DOT uses for mobility assessment) for performing a ranking of corridor performance, and the use of high-resolution data for performing a ranking of intersection performance. In addition, a comparison of the two data sets was undertaken, which showed that the segment speed data had a good degree of correlation with the POG performance measure obtained from high-resolution data.

From these findings, a reasonable approach to undertaking a system-wide assessment of performance would use both data sets, which are a good complement to each other. The segment speed data are used for evaluating corridor performance, given advance detection is not always available, or may not be high enough quality to calculate an accurate POG, whereas the high-resolution data can be used to assess the performance of crossing street movements, which the segment speed data are not generally capable of assessing directly.

This chapter presents a case study application of data fusion with an application to the city of Cedar Rapids, Iowa.

5.2. Case Study of Cedar Rapids

In 2019, the city of Cedar Rapids, Iowa implemented the collection of high-resolution data at about 150 of its intersections. These data were used for the analysis in Chapter 4. For this project, an analysis of corridors included 21 corridors in the Cedar Rapids area. Figure 34 presents a map illustrating the overlap between the two data sets that were used in this research.



Base map image © 2021 Google Maps

Figure 34. Corridors selected for ranking and signalized intersections with high-resolution data in the Cedar Rapids area

The blue lines show the segments selected for corridor analysis, while the red dots show the locations of signals having high-resolution data. Not every signal in the city is equipped with high-resolution data collection, and the corridor definitions are not fully comprehensive. However, the two data sets provide a coverage of key corridors in the area, including major state highways with signalized intersections.

One thing that this map reveals is that corridors tend to cross each other. While state highway networks tend to exclude the denser urban core areas with grid patterns, signalized arterials in even moderately built up areas still tend to form such a pattern, and Cedar Rapids has several such areas, in particular on its north and west sides. Corridors tend to cross each other at signalized intersections, so it is possible for intersections to belong to two corridors (or more if the intersection is used as a corridor boundary).

To perform a ranking, it is necessary to match corridors to intersections. This was done for the Cedar Rapids area, with part of the overall matching matrix shown in Table 13.

Table 13. Portion of the matrix of intersections to corridors

	Corridor ID																				
Signal ID	**230**	**231**	**232**	**234**	**235**	**236**	**237**	**238**	**239**	**240**	**241**	**243**	**245**	**246**	**248**	**249**	**250**	**251**	**252**	**253**	**2115**
112118							X														
112119							X														
112120							X														
112121							X														
131122							X										X				X
221203			X															X			
330342												X									
330343												X									
330344											X										
330345													X								
331305												X				X					
331306																X					
331307																X					
331308													X			X					
331309													X								
331310																X					
331311																X					
331312																	X				
331313																	X				
332318												X									
332321													X								
332322													X								
332323											X		X								
333324	X																				
333325	X																				
333326	X																				
333327	X																				
333328	X																				
333329	X																				

The full matrix is 147 × 21, so only a portion is included here. Of the 21 corridors defined for this study, 17 had intersections with high-resolution data while 4 did not. Of the 147 intersections included, 12 belonged to more than one corridor, 61 belonged to one corridor, 59 were not included in any of the defined corridors, and 15 did not have useable data (these are not shown on the map). The median number of intersections per corridor (excluding those without any intersections) was 6, with individual corridors having between 1 and 11 intersections.

5.3. Data Fusion for Comparing Corridors

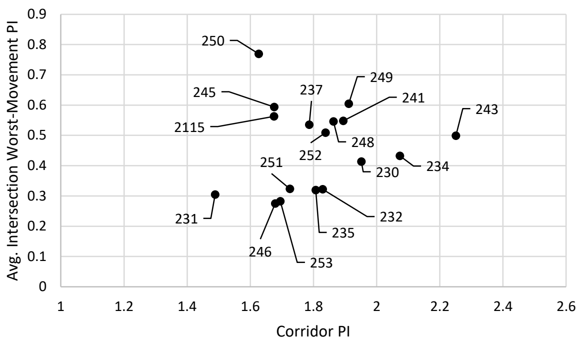
Chapter 3 described the development of a corridor PI that considered the travel time and travel-time reliability of both directions of a corridor across various times of day, while Chapter 4 described the development of two intersection PIs that either revealed the state of the worst-performing movement at the intersection, or the overall level of intersection utilization. In this section, these developments will be called the worst-movement and utilization intersection PIs.

Using the intersection-corridor mapping in Table 13, it is possible to obtain all of the intersection performance measures for each corridor. This yields an array of as many as 11 intersection-level metrics for each corridor, of which two intersection-level metrics are available. The values for the various intersections can then be averaged to yield an average intersection performance score, or the maximum value taken.

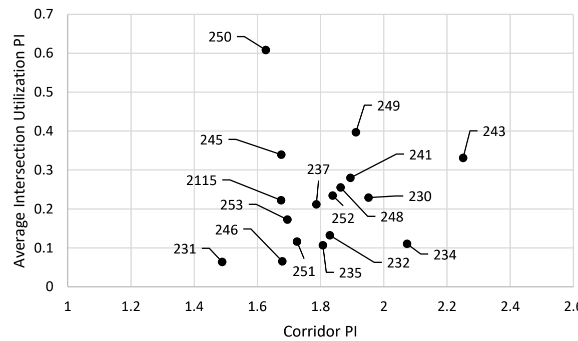
Table 14 shows a summary of these possibilities and what these aggregated performance measures would imply when conducting a ranking of the corridors, while Figure 35 shows charts of each possible ranking with values plotted in comparison with the corridor ranking.

Table 14. Options for aggregating intersection performance measures

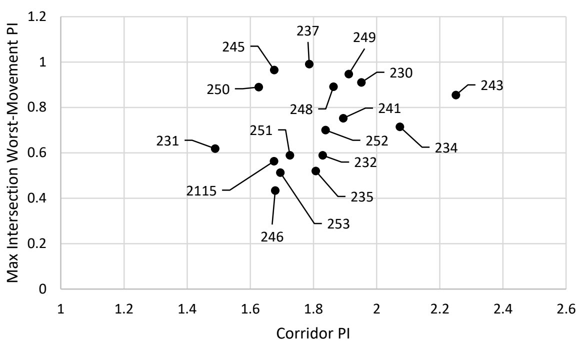
Aggregation method	Intersection worst-movement PI	Intersection utilization PI
Average	(#1) The average is taken of the worst-performing movements at all the intersections. The corridor containing more intersections having poorly performing intersections rises to the top of the list.	(#2) The average utilization of all intersections is taken as the corridor score. Corridors having a higher number of busy intersections rise to the top of the list.
Maximum value	(#3) The worst-performing movement across all the intersections becomes the aggregate score. The corridor containing the worst-performing movement rises to the top of the list.	(#4) The corridor score is equal to the level of utilization of the busiest intersection. Corridors having one very busy intersection rise to the top of the list.



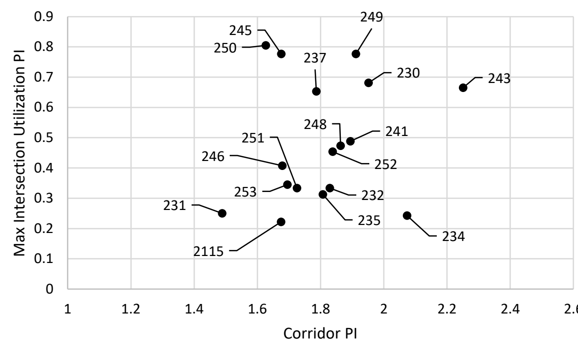
(a) Average worst-movement PI



(b) Average utilization PI



(c) Maximum worst-movement PI



(d) Maximum utilization PI

Figure 35. Options for selecting an aggregated intersection PI

Note that for Figure 35 the four options are each plotted against the corridor PI. As the charts reveal, the choice of the performance measure and the choice of the aggregation method tends to cause different corridors to ascend to the top of the chart, thus having the worst performance (although there are a few corridors that tend to stay near the top in any case).

Ultimately, the choice of aggregation depends on the objective of the analyst. If the goal is to identify the worst locations, then the maximum-value method makes sense. However, considering that intersections may belong to multiple corridors, this value may be less useful for ranking corridors. The utilization PI is able to assess the overall usage of capacity, whereas the worst-movement method identifies the locations with the most problems. Ultimately, it was decided to proceed with method #1 (average worst-movement PI), given it would indicate corridors having problem movements at multiple intersections instead of simply containing one problem movement.

The final step was to develop a composite metric that would consider both the corridor and intersection PIs. The interpretation of the diagrams in Figure 35 is rather simple: the intersections more to the right and further to the top of the chart have worse performance in terms of progression or capacity utilization, respectively. An issue with combining the two measures is that the corridor PI has a base value of 1, whereas the intersection PI has a range of values between 0 and 1. Thus, the tendency would be for the corridor PI to dominate the combined score. To avoid this, the corridor PI was normalized by dividing all values by the maximum value obtained in the study.

Next, the Euclidean distance to the point was calculated by the following equation:

$$PI = \sqrt{CPI_n^2 + IPI^2} \tag{3}$$

where *PI* is the combined PI, CPI_n is the normalized corridor PI, and *IPI* is the intersection PI. Figure 36 shows the results for the 21 corridors, ranked according to the combined PI, with values of the corridor and intersection PIs shown for comparison.

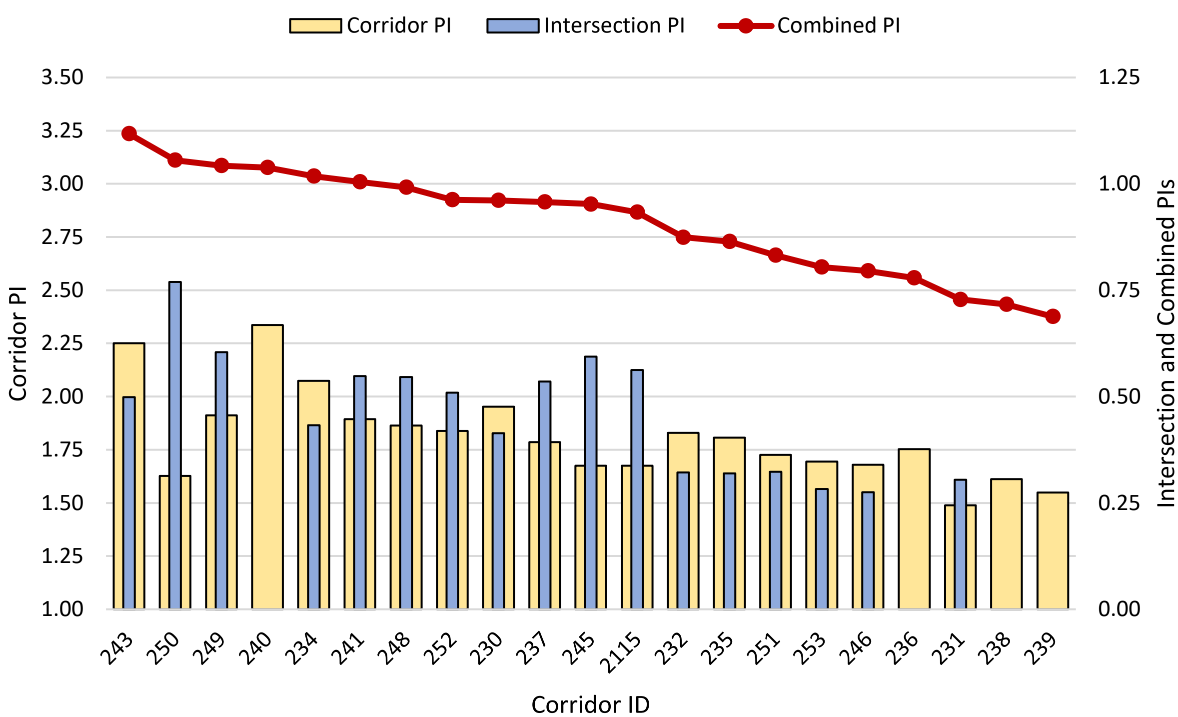


Figure 36. Comparison of corridor, intersection, and combined PI values

Each of those rankings would independently yield a different ranking, but the combined PI takes both into consideration. Note that four of the corridors do not have an intersection value, because no high-resolution data were available from any of their intersections. The results of the ranking are presented in Table 15.

Table 15. OICL for the Cedar Rapids area

Corridor ID	Location name	Corridor PI	Intersection PI	Combined PI
243	Council Street Northeast	2.25	0.50	1.12
250	IA 100	1.63	0.77	1.06
249	IA 100	1.91	0.60	1.04
240	US 151/IA 13	2.34	N/A	1.04
234	Kirkwood Boulevard Southwest	2.07	0.43	1.02
241	Boyson Road	1.89	0.55	1.00
248	32nd Street Northeast	1.86	0.55	0.99
252	6th Street Southwest	1.84	0.51	0.96
230	Edgewood Road Northeast	1.95	0.41	0.96
237	US 151 Business	1.79	0.54	0.96
245	County Road W56	1.68	0.59	0.95
2115	US 151 Business	1.68	0.56	0.93
232	Edgewood Road Southwest	1.83	0.32	0.87
235	6th Street Southwest	1.81	0.32	0.86
251	US 151 Business	1.73	0.32	0.83
253	1st Street Southwest	1.69	0.28	0.80
246	16th Avenue Southwest	1.68	0.28	0.80
236	US 151 Business	1.75	N/A	0.78
231	Mount Vernon Road Southeast	1.49	0.30	0.73
238	10th Street	1.61	N/A	0.72
239	East Post Road	1.55	N/A	0.69

Note: Repeat names indicate different spans of the same roadway. N/A indicates that high-resolution data were not available for this corridor.

The results of the ranking seem to mostly align with the experience of traffic engineers in the Cedar Rapids area. The appearance of Council Street at the top of the list was a bit surprising at first, but its performance is likely a result of its situation as a lower-priority corridor that crosses two other higher-priority corridors that are coordinated. Thus, travelers on Council Street have a few locations where they are likely to incur some delay as they cross these other higher-priority corridors, which can explain the higher corridor PI value. This highlights the need to connect the ranking results with agency objectives. Results for the US 151/Iowa 13 corridor are likely driven by poor performance of one particularly high-utilization intersection that is known to have some issues with detection.

5.4. Outlook for Statewide Integration

This case study focused on the Cedar Rapids area, which was one location in the state of Iowa where high-resolution data were available from a large number of intersections. While probe vehicle data can be used to compare performance for a large number of corridors, as shown in Chapter 3, that data source mostly focuses on the major through movements along the corridor and does not necessarily reveal performance of minor movements. To include a greater portion of corridors into this analysis beyond the use of probe vehicle data, there would need to be additional data from intersections. However, as discussed in Chapter 2, there are several requirements for obtaining such data. It may be a challenge to deploy such technology at a statewide level rapidly, although certain locations may have a lower barrier to entry than others.

In the future, it seems likely that new data sets will be able to provide movement-based performance data for intersections, similar to how probe vehicle data currently offers average speeds for segments per minute. At present, while there are several data vendors starting to market data that could be used for this purpose, or data products that include such information, such data are still in their infancy. It seems likely that within the next year, some early research on potential applications of such data will be feasible. While the quality of such data sources for locations in Iowa is not yet known, it is possible that movement-level data could be obtained in the near future without the need for infrastructure, which may have a potential for conducting a state-level operational analysis. Although such data would not include signal state data, density data at a high enough rate of penetration would make for an effective screening tool to identify locations where collection of signal state data would be useful for performing a more detailed analysis.

6. CONCLUSIONS

6.1 Report Summary

This study investigated the potential uses of currently available data sets for the purpose of developing an OICL. A review of previous data sets was conducted, as presented in Chapter 2, including a survey of legacy data, currently available data, and emerging data sets that are likely to become common in the next few years. The review focused on segment speed data and high-resolution data given their current availability and applicability to OICL development.

Next, the use of segment speed data to rank facilities at the corridor level was conducted. In this study, data from INRIX was used, because the Iowa DOT purchases these data already. Chapter 3 examined the use of these data for ranking arterial corridors in a study that compared the performance of 250 corridors across the state and compared their performance between 2019 and 2020. In Chapter 4, the use of high-resolution data from traffic signal controllers was explored. A case study for Cedar Rapids, which has deployed data collection across most of its intersections, was carried out wherein signals were ranked according to criteria developed from the number of FOMOs per phase. These performance measures can be implemented readily given they do not require knowledge of the detector mapping. A series of criteria were developed to perform an intersection-level ranking. In addition, a comparison was made between segment speed data and POG values (among other variables) obtained from high-resolution data to ascertain the degree to which they correlate with one another.

Finally, in Chapter 5, the corridor-level and intersection-level metrics were combined to yield a composite metric that allowed for the creation of an OICL. This was applied to Cedar Rapids, from which both corridor and intersection data were available.

6.2 Recommendations for Future Research

The present study uses data that are currently available, although coverage of high-resolution controller event data is limited to locations that possess the necessary infrastructure to support it. These data are helpful for evaluating the operation of actuated and actuated-coordinated signal control, for which actual operation can vary substantially from the programmed settings. This would make it challenging to duplicate the present methodology across the entirety of the state. However, at the time of this study, some new data sets have been emerging that may make it possible to obtain movement-based metrics using probe vehicles or vehicle telematics, which would improve the scalability of the method. Future extensions of this study could extend the methodology to incorporate this type of data, and provided these data sets yield enough data to support it, expand the OICL to include the entire inventory of all 2,300 signalized intersections in the state.

Another area in which the present methodology could be expanded would be to include non-signalized intersections. This also would likely be assisted through the introduction of movement-based probe vehicle data.

REFERENCES

Codjoe, J., R. Thapa, and A. S. Yeboah. 2020. *Exploring Non-Traditional Methods of Obtaining Vehicle Volumes*. Louisiana Transportation Research Center, Baton Rouge, LA.

Day, C. M., J. R. Sturdevant, and D. M. Bullock. 2010a. Outcome-Oriented Performance Measures for Management of Signalized Arterial Capacity. *Transportation Research Record: Journal of the Transportation Research Board*, No.. 2192, pp. 24–36.

Day, C. M., R. Haseman, H. Premachandra, T. M. Brennan, J. S. Wasson, J. R. Sturdevant, and D. M. Bullock. 2010b. Evaluation of Arterial Signal Coordination: Methodologies for Visualizing High-Resolution Event Data and Measuring Travel Time. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2192, pp. 37–49.

Day, C. M., D. M. Bullock, H. Li, S. M. Remias, A. M. Hainen, R. S. Freije, A. L. Stevens, J. R. Sturdevant, and T. M. Brennan. 2014. *Performance Measures for Traffic Signal Systems: An Outcome-Oriented Approach*. Purdue University, West Lafayette, IN.

Day, C. M., S. M. Remias, H. Li, M. M. Mekker, M. L. McNamara, E. D. Cox, and D. M. Bullock. 2015a. Performance Ranking of Arterial Corridors Using Travel Time and Travel Time Reliability Metrics. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2487, pp. 44–54.

Day, C. M., D. M. Bullock, H. Li, S. M. Lavrenz, W. B. Smith, and J. R. Sturdevant. 2015b. *Integrating Traffic Signal Performance Measures into Agency Business Processes*. Purdue University, West Lafayette, IN.

Day, C. M., M. L. McNamara, H. Li, R. S. Sakhare, J. Desai, E. D. Cox, D. K. Horton, and D. M. Bullock. 2016. *2015 Indiana Mobility Report and Performance Measure Dashboards*. Purdue University, West Lafayette, IN.

Day, C. M., H. Li, J. R. Sturdevant, and D. M. Bullock. 2018. Data-Driven Ranking of Coordinated Traffic Signal Systems for Maintenance and Retiming. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2672, No. 18, pp. 167–178.

Day, C. M., P. O'Brien, A. Stevanovic, D. Hale, and N. Matout. 2020. *A Methodology and Case Study: Evaluating the Benefits and Costs of Implementing Automated Traffic Signal Performance*. FHWA-HOP-20-003. Federal Highway Administration, Office of Operations, Washington, DC.

FHWA. 2014. National Performance Management Measures; Highway Safety Improvement Program.. *Federal Register* 79, No. 47, pp. 13846–13871.

———. 2016. *Traffic Monitoring Guide*. FHWA-PL-17-003. Federal Highway Administration, Office of Highway Policy Information, Washington, DC.

———. 2021. Transportation Performance Management (TPM) Toolbox. https://www.tpmtools.org/.

Freije, R. S., A. M. Hainen, A. L. Stevens, H. Li, W. B. Smith, H. Summers, C. M. Day, J. R. Sturdevant, and D. M. Bullock. 2014. Graphical Performance Measures for Practitioners to Triage Split Failure Trouble Calls. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2439, pp. 27–40.

Hainen, A. and M. Dunn. 2015. *Alabama Statewide Mobility Report 2014*. University Transportation Center for Alabama, University of Alabama, Tuscaloosa, AL.

Hallmark, S. L., R. Basavaraju, and M. Pawlovich. 2002. *Evaluation of the Iowa DOT's Safety Improvement Candidate List Process*. Center for Transportation Research and Education, Institute for Transportation, Iowa State University, Ames, Iowa. https://intrans.iastate.edu/app/uploads/2018/03/SafetyCandidate.pdf.

Hofleitner, A., R. Herring, and A. Bayen. 2012. Arterial Travel Time Forecast with Streaming Data: A Hybrid Approach of Flow Modeling and Machine Learning. *Transportation Research Part B: Methodological*, Vol. 46, No. 9, pp. 1097–1122.

Hu, H. and H. X. Liu. 2013. Arterial Offset Optimization Using Archived High-Resolution Traffic Signal Data. *Transportation Research Part C: Emerging Technologies*, Vol. 37, pp. 131–144.

Hu, J., M. D. Fontaine, and J. Ma. 2016a. Quality of Private Sector Travel-Time Data on Arterials. *Journal of Transportation Engineering*, Vol. 142, No. 4, pp. 04016010-1–04016010-11.

Hu, J., M. D. Fontaine, B. B. Park, and J. Ma. 2016b. Field Evaluations of an Adaptive Traffic Signal—Using Private-Sector Probe Data. *Journal of Transportation Engineering*, Vol. 142, No. 1, pp. 04015033-1–04015033-9.

INRIX. n.d. INRIX U.S. Signals Scorecard. https://inrix.com/learn/inrix-signals-scorecard-report/.

ITE. n.d. COVID-19 Traffic Volume Trends. https://www.ite.org/about-ite/covid-19-resources/covid-19-traffic-volume-trends/.

Iowa DOT. 2016. *2016 Interstate Congestion Report*. Iowa Department of Transportation and Insitute for Transportation Center for Transportation Research and Education at Iowa State University, Ames, Iowa. https://iastate.box.com/shared/static/gue2p1mc5axzad1r2uh5sb7kfh1chsse.pdf.

Li, H., A. M. Hainen, C. M. Day, G. Grimmer, J. R. Sturdevant, and D. M. Bullock. 2013. Longitudinal Performance Measures for Assessing Agencywide Signal Management Objectives. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2355, No. 1, pp. 20–30.

Mathew, J., D. Krohn, H. Li, C. Day, and D. Bullock. 2017. *Implementation of Probe Data Performance Measures*. Pennsylvania Deptartment of Transportation, Bureau of Maintenance and Operations, Harrisburg, PA.

Mathew, J. K., H. Li, M. Blackwell, and D. M. Bullock. 2020. *Hard-Braking Event Dataset for Interstate Routes in Indiana*. Purdue University, West Lafayette, IN. https://purr.purdue.edu/publications/3656/1.

Mathew, J. K., J. C. Desai, R. S. Sakhare, W. Kim, H. Li, and D. M. Bullock. 2021. Big Data Applications for Managing Roadways. *ITE Journal*, Vol. 91, No. 2, pp. 28–35.

Michigan DOT. 2020. *2019 Freeway Congestion and Reliability Report*. Michigan Department of Transportation, Lansing, MI. https://www.michigan.gov/mdot/0,4616,7-151-9622_11045_25024_75677---,00.html.

Missouri DOT. 2020. *St. Louis District Mobility: Quarterly Regional Mobility Report June/July/August 2020*. Missouri Department of Transportation, Jefferson City, MO. https://www.modot.org/sites/default/files/documents/Regional_Mobility_June_July_August_2020_Final.pdf.

Pennsylvania DOT. 2020. *Green Light-Go: Pennsylvania's Municipal Signal Partnership Program*. Pennsylvania Deptartment of Transportation, Harrisburg, PA. http://www.dot.state.pa.us/portal information/traffic signal portal/fundglg.html.

Pujara, K., S. Mahapatra, C. Delion, M. Hill, K. F. Sadabadi, L. Shemer, and M. Wolniak. 2019. *2019 Maryland State Highway Mobility Report: Eighth Edition*. Maryland Department of Transportation, State Highway Administration, Baltimore, MD.

Quiroga, C. A. and D. Bullock. 1998. Determination of Sample Sizes for Travel Time Studies. *ITE Journal*, Vol. 68, No. 8, pp. 92–98.

Real-Time Analytics of Transportation Data Laboratory. 2021. Historical Interstate Congestion Report. Institute for Transportation, Iowa State University. https://reactor.ctre.iastate.edu/historical-interstate-congestion-report/.

Roll, J. 2019. *Evaluating StreetLight Estimates of Annual Average Daily Traffic in Oregon*. Oregon Department of Transportation, Salem, OR.

Smaglik, E. J., A. Sharma, D. M. Bullock, J. R. Sturdevant, and G. Duncan. 2007. Event-Based Data Collection for Generating Actuated Controller Performance Measures. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2035, pp. 97–106.

Tsapakis, I., L. Cornejo, and A. Sanchez. 2020. *Accuracy of Probe-Based Annual Average Daily Traffic (AADT) Estimates in Border Regions*. Center for International Intelligent Transportation Research, Texas A&M Transportation Institute, El Paso, TX.

Turner, S., J. Carson, L. J. Wilkinson, K. Travis, and C. Zimmerman. 2010. *Traffic Monitoring: A Guidebook*. FHWA-WFL/TD-10-002. Federal Highway Administration, Western Federal Lands Highway Division, Washington, DC.

Turner, S. M., W. L. Eisele, R. J. Benz, and D. J. Holdener. 1998. *Travel Time Data Collection Handbook*. FHWA-PL-98-035. Federal Highway Administration, Office of Highway Information Management, Washington, DC.

Turner, S., I. Tsapakis, and P. Koeneman. 2020. *Evaluation of StreetLight Data's Traffic Count Estimates from Mobile Device Data*. Minnesota Department of Transportation, Office of Research & Innovation, St. Paul, MN.

U.S. DOT. n.d. *Connected Vehicle Basics: Privacy*. U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (OST-R), Intelligent Transportation Systems, Joint Program Office, Washington, DC. https://www.its.dot.gov/cv_basics/cv_basics_privacy.htm.

Wang, Y., B. N. Araghi, Y. Malinovskiy, J. Corey, and T. Cheng. 2014a. *Error Assessment for Emerging Traffic Data Collection Devices*. Pacific Northwest Transportation Consortium (PacTrans), University of Washington, Seattle, WA.

Wang, Y., Y. Zheng, and Y. Xue. 2014b. Travel Time Estimation of a Path Using Sparse Trajectories. *KDD '14: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery (ACM), New York, NY, pp. 25–34.

Wasson, J. S., J. R. Sturdevant, and D. M. Bullock. 2008. Real-Time Travel Time Estimates Using Media Access Control Address Matching. *ITE Journal*, Vol. 78, No. 6, pp. 20–23.

Wolf, J. C., J. Ma, B. Cisco, J. Neill, B. Moen, and C. Jarecki. 2019. Deriving Signal Performance Metrics from Large-Scale Connected Vehicle System Deployment. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2673, No. 4, pp. 36–46.

Wünsch, G., F. Bölling, A. von Dobschütz, and P. Mieth. 2015. Bavarian Road Administration's Use of Probe Data for Large-Scale Traffic Signal Evaluation Support. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2487, pp. 88–95.

Yang, H., M. Cetin, and Q. Ma. 2020. *Guidelines for Using StreetLight Data for Planning Tasks*. Virginia Transportation Research Council, Charlottesville, VA.

Young, S. 2014. Use of Probe and Bluetooth Data for Arterial Performance Measures in the I-95 Corridor Coalition. *NATMEC 2014: Improving Data Collection, Analysis, and Use; Workshop on Arterial Data and Performance Monitoring*, North American Travel Monitoring Exposition and Conference (NATMEC), June 29–July 2, Chicago, IL.

THE INSTITUTE FOR TRANSPORTATION IS THE FOCAL POINT FOR TRANSPORTATION AT IOWA STATE UNIVERSITY.

InTrans centers and programs perform transportation research and provide technology transfer services for government agencies and private companies;

InTrans contributes to Iowa State University and the College of Engineering's educational programs for transportation students and provides K–12 outreach; and

InTrans conducts local, regional, and national transportation services and continuing education programs.

INSTITUTE FOR TRANSPORTATION

IOWA STATE UNIVERSITY

Visit **InTrans.iastate.edu** for color pdfs of this and other research reports.