

INTEGRATING TRAFFIC SIGNAL PERFORMANCE MEASURES INTO AGENCY BUSINESS PROCESSES



*Christopher M. Day, Darcy M. Bullock, Howell Li, Steven M. Lavrenz,
W. Benjamin Smith, James R. Sturdevant*



Recommended Citation

Day, C. M., D. M. Bullock, H. Li, S. M. Lavrenz, W. B. Smith, and J. R. Sturdevant. *Integrating Traffic Signal Performance Measures into Agency Business Processes*. Purdue University, West Lafayette, Indiana, 2015. <http://dx.doi.org/10.5703/1288284316063>

Acknowledgments

This work was supported in part by Pooled Fund Study TPF-5(258), led by the Indiana Department of Transportation (INDOT) and supported by the state transportation agencies of California, Georgia, Minnesota, Mississippi, New Hampshire, Pennsylvania, Texas, Utah, and Wisconsin, the Indiana Local Technical Assistance Program (LTAP), the Chicago Department of Transportation, and the Federal Highway Administration Office of Operations and Resource Center. Some material in this monograph was compiled from previous studies that were made possible under National Cooperative Highway Research Program project 3-79a, INDOT State Planning and Research (SPR) projects, Indiana LTAP projects, and USDOT through Small Business Innovation Research (SBIR) projects with Traffax, Inc., and through a joint research project with Marshall University. We would like to thank Rick Schuman and colleagues at Inrix, Inc., for provision of sample vehicle trajectory data in Chapter 7. We are grateful to these sponsors and research partners for their support over the years.

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein, and do not necessarily reflect the official views or policies of the sponsoring organizations. These contents do not constitute a standard, specification, or regulation.

Copyright

Copyright 2015 by Purdue University. All rights reserved.

Print ISBN: 978-1-62260-376-3

ePUB ISBN: 978-1-62260-377-0

ABSTRACT

INTEGRATING TRAFFIC SIGNAL PERFORMANCE MEASURES INTO AGENCY BUSINESS PROCESSES

This report discusses uses of and requirements for performance measures in traffic signal systems facilitated by high-resolution controller event data. Uses of external travel time measurements are also discussed. The discussion is led by a high-level synthesis of the systems engineering concepts for traffic signal control, considering technical and nontechnical aspects of the problem. This is followed by a presentation of the requirements for implementing data collection and processing of the data into signal performance measures. The remaining portion of the report uses an example-oriented approach to show a variety of uses of performance measures for communication and detector system health, quality of local control (including capacity allocation, safety, pedestrian performance, preemption, and advanced control analysis), and quality of progression (including evaluation and optimization).

CONTENTS

1. Introduction	1
1.1 Objective	1
1.2 Opportunities for Improvement	1
1.3 A Language for Objectives	3
1.4 Interdependency of Components	6
1.5 Business Processes	8
1.6 Organization of the Report	11
1.7 Conclusion	12
2. Requirements for Signal Performance Measures	12
2.1 Introduction	12
2.2 Cabinet Types	12
2.3 Controllers and High-Resolution Data Logging	12
2.4 Detection Technology	13
2.5 Detector Positioning	15
2.6 Detection for Actuated Signal Control	16
2.7 Detector Mapping for Performance Measures	18
2.8 Supported Performance Measures	19
2.9 Travel Time Data	22
2.10 Communication	23
2.11 Data System	25
2.12 Data Scalability	26
2.13 Displaying Performance Measures	27
2.14 Conclusion	30
3. Communication	30
3.1 Basic Signal Communications Architecture	30
3.2 Accessing High-Resolution Signal Controller Data	31
3.3 Identification of Failed Communication	31
3.4 Long-Term Use of Signal Communications Failure Dashboard	33
3.5 Conclusion	34
4. Detection	34
4.1 Background for Traffic Signal Loop Detector Repair	35
4.2 Case Study: Corridor-Level Detection Failure	35
4.3 Case Study: US 421	36
4.4 Detector Repair Results	38
4.5 Benefit/Cost Analysis of Detector Repairs	42
4.6 Conclusion	43
5. Evaluation of Local Signal Control	43
5.1 Overview	43
5.2 Basic Phase Service Concepts	43
5.3 Flow Rates	46
5.4 Evaluating Local Timing in Detail	46
5.5 Case Study of a Single Intersection during a Freeway Diversion	70
5.6 Example Network and Corridor Analysis	78
5.7 Chapter Conclusion	85
6. Evaluation of System Control	86
6.1 Overview	86
6.2 Basic Traffic Progression Concepts	86
6.3 Evaluation of the Quality of Progression	88
6.4 Operational Diagnostics with PCDs	92
6.5 Optimizing Signal Timing with High-Resolution Data	101
6.6 System and Corridor Analysis	114
6.7 Chapter Summary	129
7. Conclusion	130
7.1 Report Summary	130
7.2 Path Forward for Implementation	130
7.3 Future Fusion with Connected Vehicle and Similar Data Types	131
7.4 Report Conclusion	132
References	134

LIST OF TABLES

Table	Page
Table 1.1 Signal system objectives by mode, based on discussion in the <i>Signal Timing Manual</i>	4
Table 1.2 The RATER model of service quality, applied to traffic signal systems	5
Table 1.3 Uses of performance measures at different stages of the signal systems engineering life cycle	11
Table 2.1 High-resolution event data logging-supported controllers	13
Table 2.2 Parameters for measuring performance per detection zone, Indiana Department of Transportation	14
Table 2.3 Acceptance criteria for detection missed and false calls, Indiana Department of Transportation	14
Table 2.4 Example detector mapping and descriptive information for an eight-phase intersection	19
Table 2.5 Data process requirements	25
Table 4.1 Changes in median corridor travel times (in seconds) as result of loop detector maintenance	42
Table 5.1 Performance measures for evaluating capacity utilization	47
Table 5.2 Odds ratio test for RLR rates per 1,000 entering vehicles before/after split adjustment	62
Table 6.1 Performance measures for evaluating the quality of progression	88
Table 6.2 Benefits (user cost reductions) due to offset optimization on SR 37 North, 2010–2015. All figures in USD. Negative numbers indicate increases in user costs	116

LIST OF FIGURES

Figure	Page
Figure 1.1 Comparison of the relative degree of development in use cases related to traffic signal systems	2
Figure 1.2 The concept of a Pareto Front applied to two competing objectives in signal operation, after Gartner et al	5
Figure 1.3 The seven-layer Open Systems Interconnection (OSI) model	6
Figure 1.4 Translating traffic signal control components into a hierarchy of needs	7
Figure 1.5 The systems engineering “V diagram”	8
Figure 1.6 Life cycle of a signal control policy	9
Figure 1.7 Life cycle of signal control equipment	10
Figure 1.8 Tracking of user costs for (a) Saturdays and (b) Weekdays across a 5-year period for operation on SR 37. Three “before” and “after” periods are included in the analysis relative to when signal offsets were optimized. The user cost reductions across the 5-year period total to \$3.6 million. Additional details are provided in Chapter 6	11
Figure 2.1 NEMA TS/2 Type 1 cabinet	13
Figure 2.2 Conceptual comparison of outputs from various detector modes	15
Figure 2.3 Example of lane-by-lane configuration for vehicular detection	16
Figure 2.4 Example detection layout with stop bar zones that are “tied together.” Such a configuration will still support phase utilization performance measures, but it will have less detailed information than lane-by-lane detection	17
Figure 2.5 Detector bus interface unit (BIU) with approach, phase, and lane labeled	18
Figure 2.6 Example intersection with assigned phasing and detection zones	20
Figure 2.7 Selective alternative detector configurations without count detection	21
Figure 2.8 Selective alternative detector configurations with count detection	22
Figure 2.9 Detector configuration requirements matrix for performance measures	23
Figure 2.10 Vehicle re-identification using nonpermanent Bluetooth sensing technology for travel time measurement	23
Figure 2.11 Travel time computation using crowd-sourced probe data segmentation	24
Figure 2.12 Example communication infrastructure using fiber optics and broadband wireless communications	24
Figure 2.13 Scheme for high-resolution data collection and clock synchronization using an embedded computer	25
Figure 2.14 Example of data record accumulation by record type over six months	26
Figure 2.15 Example of using MS Excel to visualize performance measure data	27
Figure 2.16 Example of a route definition table	28
Figure 2.17 Example of using a web application-integrated charting API to plot travel time data	28
Figure 2.18 Example of using a web application-integrated charting API to plot performance measures	29
Figure 2.19 Architecture of a REST system for displaying performance measures	29
Figure 2.20 Model–view–controller (MVC) software design pattern	29
Figure 2.21 Architecture using MS Excel for displaying performance measures	30
Figure 3.1 Overview of INDOT high-resolution data communications network.	31
Figure 3.2 Statewide signal controller communications failures, by district and corridor	32
Figure 3.3 Statewide signal controller communications failures, by district, corridor, and communications technology	32
Figure 3.4 Longitudinal plot of signal communications failures for LaPorte district, May 2015	33
Figure 3.5 Four-month plot of signal communications failures for LaPorte maintenance district	34
Figure 4.1 Longitudinal plot of phases from high-resolution equipped intersections with five or more forced terminations between 2AM and 3AM, grouped by corridor	36
Figure 4.2 Overview of US 421 study corridor	37

Figure 4.3 NB and SB PCDs for selected intersections along US 421, 12/3/2014	38
Figure 4.4 PCDs for select US 421 intersections before and after detector repairs	39
Figure 4.5 Comparison of forced phase terminations and green time, two weeks before and after detector repair at US 421 and 99th Street	40
Figure 4.6 Average weekday green time by hour of day before and after detection repairs for US 421 and 99th Street	41
Figure 4.7 US 421 cumulative distribution functions of travel time, by direction and TOD	41
Figure 4.8 Annualized user benefits from travel time improvements due to detector repairs	42
Figure 4.9 BCRs for detector repairs along US 421. The shaded areas correspond to actual costs between 50% and 150% of those estimated	42
Figure 5.1 Phase initiation and termination processes	44
Figure 5.2 Illustration of local control of a phase, including an example of a split failure	44
Figure 5.3 Illustration of a split failure. The highlighted truck is the same vehicle in the three images	45
Figure 5.4 Flow rates for US 231 and River Road intersection	46
Figure 5.5 Cycle length, US 31 and 126th Street	48
Figure 5.6 Artifacts of signal timing that can be observed from a close-up view of green duration. Data shown for SR 37 and SR 32 in Noblesville, Indiana	48
Figure 5.7 Cycle-by-cycle green times, US 31 and 126th Street	49
Figure 5.8 Changes in cumulative green time at US 31 and 126th Street, with changes to splits noted	49
Figure 5.9 Split monitor plot from UDOT Signal Performance Measures website. Data shown for 700 E at 800 S, 12/2/2015, for phases 2 and 4	50
Figure 5.10 Purdue Phase Termination plots.	51
Figure 5.11 US 31 and 126th Street, Carmel, Indiana.	52
Figure 5.12 Illustration of the calculation of ROR and GOR.	53
Figure 5.13 Red Occupancy Ratio (ROR) versus Green Occupancy Ratio (GOR) at US 31 and 126th Street. The numbers indicate the cycles within the red region where ROR and GOR are both greater than 0.8	54
Figure 5.14 Cycle-by-cycle volume-to-capacity ratios for mainline through phases 2 and 6 at US 31 and 126th Street	55
Figure 5.15 Cumulative frequency diagrams of volume-to-capacity ratio for mainline through phases 2 and 6 at US 31 and 126th Street	56
Figure 5.16 Number of split failures before and after split adjustment at US 31 and 126th Street: Thursday, 7/18/2013, versus Thursday, 7/25/2013	56
Figure 5.17 Split failures by phase, by day of week, before and after split adjustment at US 31 and 126th Street: week of 7/15/2013, versus week of 7/22/2013	57
Figure 5.18 Cumulative split failures on phase 3 at US 31 and 126th Street, before and after a split adjustment	57
Figure 5.19 Cumulative split failures on 8 phases at US 31 and 126th Street, before and after a split adjustment	57
Figure 5.20 Input–output delay (advance detector)	58
Figure 5.21 Approach Delay plot from UDOT Signal Performance Measures website. Data shown for 700 E at 800 S, 12/2/2015	58
Figure 5.22 Random-arrival delay (stop bar detector)	58
Figure 5.23 Analysis of total delays for free versus coordinated operation for US 231 during the AM Peak (06:00–09:00) during timing plan design in 2014	59
Figure 5.24 Cumulative frequency diagrams of average delays per cycle for 5 “before” and 5 “after” days at US 31 and 126th Street, before and after a split change	59
Figure 5.25 Conceptual overview of RLR detection using loop detector and phase event data	60
Figure 5.26 Example video validation of RLR data at US 231 and State Street (facing northbound US 231)	60
Figure 5.27 Comparison of RLR data events and video validation, along with safety risks for valid RLRs. Callout <i>i</i> shows the refined search area for RLR events, which captures all valid RLRs, except for two large semi-trucks (callout <i>ii</i>). Callout <i>iii</i> identifies particularly high-risk RLR events, which occurred nearly 3 s after the start of red	61

Figure 5.28 Effect of phase 8 split increase on red light running (RLR) vehicles during the 09:00–15:00 TOD plan at US 31 and 126th Street. (a) Change in daily RLR counts before/after split increase; (b) Change in average daily RLRs per 1,000 entering vehicles before/after split increase	61
Figure 5.29 Intersection phase configuration at Northwestern and Stadium, West Lafayette, Indiana	62
Figure 5.30 Count of the number of vehicles conflicting with pedestrian phase 4 per cycle	63
Figure 5.31 Percentage of cycles with pedestrian phases (during 06:00–22:00) crossing Northwestern Avenue	63
Figure 5.32 Percentage of cycles with pedestrian phases crossing Northwestern Avenue before and after exclusive pedestrian phase implementation (data from November 2008 and November 2009, 06:00–22:00, excluding special events, breaks, and inclement weather)	64
Figure 5.33 Daily, weekly, and annual variation in pedestrian phase actuation (all data from after implementation of exclusive pedestrian phase)	64
Figure 5.34 Variation in pedestrian phase actuation due to weather effects (all data from 06:00–22:00 during academic year after implementation of exclusive pedestrian phase)	65
Figure 5.35 High-resolution phase and preemption data, US 36 at Ronald Reagan Parkway, 10/14/2014	65
Figure 5.36 Phase and preemption data, US 36 at Ronald Reagan Parkway, 10/14/2014	66
Figure 5.37 Railroad preemption events	67
Figure 5.38 Evaluation of track clearance phase timing during signal preemption	67
Figure 5.39 Measurement of right-of-way transfer time (RTT) during signal preemption	68
Figure 5.40 Evaluation of track clearance green performance: time difference between the end of TCG and gate down confirmation	69
Figure 5.41 Phase green and red times with and without phase reservice	70
Figure 5.42 Analysis of signal operations with phase reservice. Data are shown from US 231 and Jischke Boulevard on 2/3/2014	71
Figure 5.43 US 231 and Jischke Boulevard, West Lafayette, Indiana	72
Figure 5.44 Max vehicle delay at Jischke Drive with phase reservice and different split extensions. The dot shows the median value and the bar shows the interquartile range spanning the 25th and 75th percentiles	73
Figure 5.45 US 231 diversion route	74
Figure 5.46 US 231 and River Road, West Lafayette, Indiana	75
Figure 5.47 Northbound traffic volumes at US 231 and River Road	75
Figure 5.48 Volumes on 8 phases at US 231 and River Road	76
Figure 5.49 Cycle lengths at US 231 and River Road on various dates in 2015	76
Figure 5.50 Split failures per phase at US 231 and River Road on different dates	77
Figure 5.51 CFDs of cycle-by-cycle average delay by phase at US 231 and River Road, before I-65 diversion (7/22/2015) and during I-65 diversion running free (8/12/2015)	77
Figure 5.52 CFDs of cycle-by-cycle average delay by phase at US 231 and River Road, during I-65 diversion running free (8/12/2015) and during I-65 diversion running the detour coordination plan (8/19/2015)	77
Figure 5.53 CFDs of cycle-by-cycle average delay by phase at US 231 and River Road, during I-65 diversion running the detour coordination plan (8/19/2015) and after the end of the diversion still running the detour coordination plan (9/9/2015)	78
Figure 5.54 CFDs of cycle-by-cycle average delay by phase at US 231 and River Road, after the end of the diversion still running the detour coordination plan (9/9/2015) and after the implementation of a new coordination plan for normal traffic (9/16/2015)	78
Figure 5.55 INDOT network map of online corridors for split failure analysis	79
Figure 5.56 Example split failure ticker view showing data for 11/20/2015–12/1/2015. The graphic shows the sum of all interstate system miles operating under 45 mph per 15-minute bin over the dates shown. The data are colored by INDOT district	80
Figure 5.57 System-wide count of split failures per corridor by day from 10/1/2015 to 12/1/2015	81
Figure 5.58 System-wide count of split failures by intersection and movement from 10/1/2015 to 12/1/2015	81
Figure 5.59 Map of US 31 in Greenwood (Corridor #7)	82
Figure 5.60 Count of split failures per intersection by day on US 31 Greenwood, 10/1/2015–12/1/2015	82
Figure 5.61 Split failures per intersection by time of day on US 31 Greenwood, 10/1/2015–12/1/2015	83

Figure 5.62 Split failures per movement by day at Intersection #1, US 31 and County Line Road, 10/1/2015–12/1/2015	83
Figure 5.63 Split failures per movement by time of day at Intersection #1, US 31 and County Line Road on 10/30/2015	84
Figure 5.64 Split failures per movement by day at Intersection #3, US 31 and Fry Road, 10/1/2015–12/1/2015	84
Figure 5.65 Split failures per movement by time of day at Intersection #3, US 31 and Fry Road, 11/27/2015 (Black Friday)	85
Figure 5.66 Example of an agency-wide summary report, from the UDOT Performance Measures website	85
Figure 6.1 Example arterial trajectory data from the NGSIM Peachtree Street dataset	86
Figure 6.2 Example of time–space diagram views in design software	87
Figure 6.3. A cyclic flow profile from Transyt-7F	88
Figure 6.4 Three graphical representations of vehicle arrival and phase status data based on conventional (complete) detection: (a) a time–space diagram, and a coordination diagram for one cycle; (b) a Purdue Coordination Diagram (PCD) covering 24 hours; (c) a cyclic flow profile showing arrival and probability of green distributions	90
Figure 6.5 Example PCDs from INDOT signal performance measures website. Data are shown for three intersections along US 231 in West Lafayette, Indiana for 12/2/2015	91
Figure 6.6 PCDs from Utah’s Signal Performance Metrics website for the intersection of 700 East and 800 South, 12/2/2015	92
Figure 6.7 Example executive report summary from the UDOT Signal Performance Measures website. The report shows corridor performance for 12/2/2015 from 6:00 to 9:00	93
Figure 6.8 Example of exceptional progression: Eastbound, Pendleton Pike at I-465 Northbound Ramp, 7/1/2015	93
Figure 6.9 Random arrivals and queuing at Northbound, SR 37 and 126th Street, 7/1/2015	94
Figure 6.10 Shorter and longer cycles due to preemption, and transition back to normal operations: Eastbound, Pendleton Pike and Carroll Road, 7/1/2015	94
Figure 6.11 Cycle length mismatch: Northbound, SR 37 and SR 32, 4/13/2009	95
Figure 6.12 The interchange of S.R. 1 and I-69 in Fort Wayne	96
Figure 6.13 Evaluation of diverging diamond interchange signal timing: exiting movements at S.R. 1 and I-69 in Fort Wayne, from 7/1/2015	96
Figure 6.14 “Three-phase” operation of a DDI. Data shown are from SR 201 and Bangerter Highway in Salt Lake City, Utah	97
Figure 6.15 Visual assessment using video captured from a drone at SR 201 and Bangerter Highway in Salt Lake City, Utah. The video is available at http://dx.doi.org/10.4231/R7C24TC4	98
Figure 6.16 Portion of the simulation test network for testing advanced control	98
Figure 6.17 Advanced control diagnostics using PCDs from simulation data	99
Figure 6.18 Experimental control diagnostics using PCDs from simulation data	99
Figure 6.19 Experimental control diagnostics using PCDs: real-world example. Data were obtained from the UDOT Signal Performance Measures website	100
Figure 6.20 Predicting new arrival conditions after an offset adjustment	101
Figure 6.21 Comparison of predicted and actual vehicle arrivals from the pilot corridor study	102
Figure 6.22 Comparison of flow profiles: before offset adjustment, predicted offset adjustment, and after offset adjustment	103
Figure 6.23 Explanation of the progression potential graph	103
Figure 6.24 Progression potential graph for the 2010 offset optimization study	104
Figure 6.25 Performance comparison of several alternative offset optimization algorithms	104
Figure 6.26 Computational performance of several alternative offset optimization algorithms	105
Figure 6.27 Hill-climbing algorithm	106
Figure 6.28 Link pivot algorithm	107
Figure 6.29 Implementation of link pivot on INDOT Signal Performance Measures website	108
Figure 6.30 Implementation of link pivot on UDOT Signal Performance Measures website	109
Figure 6.31 Map of SR 37 North	110

Figure 6.32 Selected PCDs from SR 37 before optimization, Saturday, 5/29/2010	111
Figure 6.33 Selected PCDs from SR 37 after optimization, Saturday, 7/10/2010	112
Figure 6.34 Travel times before and after 2010 offset optimization	113
Figure 6.35 Comparison of offset optimization outcomes. (a)/(c)/(e) demonstrate northbound and southbound travel time improvements, and (b)/(d)/(f) show northbound improvements at the expense of southbound vehicles	114
Figure 6.36 Comparison of arrivals on green (AOG) and percent arrivals on green (POG) for weekdays and Saturdays from 06:00–19:00, before/after offset optimization	115
Figure 6.37 Comparison of arterial green time, excluding 135th Street, for weekdays and Saturdays from 06:00 to 19:00, before/after offset optimization	115
Figure 6.38 Annualized user costs for (a) Saturdays and (b) weekdays, from 06:00 to 19:00, before/after offset optimization	116
Figure 6.39 Bar chart showing the breakdown of the total user benefit by analysis year	116
Figure 6.40 Plot of AOR, 10/1/2015–11/30/2015	117
Figure 6.41 Sorted list of average POR values by approach. Data represent the average POR between 10/1/2015 and 11/30/2015. Each row shows the system, intersection, and approach	118
Figure 6.42 The top 5 POR approaches in the INDOT system, 10/1/2015–11/30/2015	119
Figure 6.43 Closure of northbound I-65 in August/September 2015, and diversion of traffic onto US 231	120
Figure 6.44 US 231 in West Lafayette, Indiana	120
Figure 6.45 Dashboard showing integration of signal data (flow rates measured using high-resolution data) with probe vehicle speed data	121
Figure 6.46 View of the INDOT performance measure dashboard	121
Figure 6.47 Northbound PCDs from 9/1/2015 along US 231, as visible from the INDOT performance measure dashboard	122
Figure 6.48 Longitudinal POG for US 231 corridor, 7/1/2015–12/1/2015 (overall 24-hour performance shown)	123
Figure 6.49 AOR per intersection by date, 7/1/2015–12/1/2015	124
Figure 6.50 AOR per intersection by time of day, 7/1/2015–12/1/2015	125
Figure 6.51 POG heatmap for Intersection #4 (US 231 and River Road), 7/1/2015–12/1/2015: (a) northbound; (b) southbound	126
Figure 6.52 Selected PCDs at Intersection #4 (US 231 and River Road), for various times during the I-65 diversion	127
Figure 6.53 Selected PCDs for Intersection #4 (US 231 and River Road) showing differences between weekday (11/11/2015) and weekend (11/14/2015) operations	128
Figure 6.54 POG heatmap for Intersection #6 (US 231 and Airport Rd.), 7/1/2015–12/1/2015: (a) westbound (northbound along the corridor); (b) eastbound (southbound along the corridor)	128
Figure 6.55 Selected PCDs for Intersection #6 (US 231 and Airport Road), showing change in signal timing plan related to reduction in POG observed at a higher level	129
Figure 6.56 POG heatmap for Intersection #10 (US 231 and Old US 52), 7/1/2015–12/1/2015: (a) northbound (arterial left and right turns); (b) westbound (crossing arterial)	129
Figure 6.57 Selected PCDs for Intersection #10 (US 231 and Old US 52), showing change from coordination along US 231 to coordination along Old US 52	130
Figure 7.1 Example raw real-world probe vehicle position data from a section of SR 37 North. Data are shown for a portion of Tuesday, 4/14/2015.	132
Figure 7.2 Sensitivity of optimization outcome to market penetration p for (a) $T = 3$ h and (b) $T = 15$ min. Lines marked “CV” show the distribution of results for 100 iterations at each trial value of p	133

1. INTRODUCTION

1.1 Objective

The first Pooled Fund Study [1] report, *Performance Measures for Traffic Signal Systems: An Outcome-Oriented Approach* [2], documented an extensive portfolio of performance measures for evaluating traffic signal systems. The goal of that report was to define performance measures that could be introduced into system specifications for agencies that wanted to begin using such tools to **measure** the quality of service in their systems. The emphasis of the document was largely on performance measures obtained from high-resolution data and from external travel time measurements.

The present report takes the next step toward implementing and using performance measures to **manage** traffic signal systems. The objective is to provide a resource for agencies that are in the process of developing an active traffic management program. To address this need, this report takes the following approach:

1. To begin, the current state of practice is briefly reviewed and critical opportunities for improvement are identified where performance measures can assist.
2. To formulate a solution, we need a **language** with which we can discuss the objectives of signal performance: namely, the goals that should be sought for a specific system or subsystem (network, corridor, or intersection).
3. We need to understand the interdependence among different elements of the system, because the success of many applications at high levels depends on the success of other applications at lower levels. There is a **hierarchy of system needs** that any performance-based solution needs to be cognizant of.
4. We also need to identify the **business processes** in signal system management to understand how those processes can be improved through the implementation of performance measures.
5. Finally, we need to connect specific objectives with specific performance measures in a structured way to make improvements and validate them, and we need to understand the requirements to obtain performance measures.

The first chapter of the report walks through the first four items above, and the subsequent chapters discuss collection and use of performance measures for system management. (This chapter builds on a paper originally published in IEEE Conference Proceedings [3].) The performance measure use cases are presented using the hierarchy of system needs and incorporate a variety of examples from the past 10 years of signal performance measure research.

1.2 Opportunities for Improvement

There is a great deal of opportunity to improve the state of practice in traffic signal system management, as has been documented in numerous studies over the past 20 years. Assessments conducted in the 1990s first showed a widening gap between the state of practice and the state of the art [4]. These studies were followed up by efforts

by the Institute of Transportation Engineers [5] and the National Transportation Operations Coalition [6] to understand the causes and offer solutions to the problem.

Notably, the *Traffic Signal Report Card* self-assessment conducted by the NTOC yielded consistently low scores for the nation over the 7-year period of its existence, with a subscore of “F” in the area of traffic monitoring. The lack of information about system performance has arguably limited the ability of the traffic industry to ask for more resources to do its job better. Unlike the world of pavement management, where engineers are able to make an assessment of the likely performance that can be delivered with a given treatment, there are no such parallels that can be drawn within arterial management, because comparatively few data have been developed for the potential treatments that are available for managing traffic [7].

A 2009 FWHA report [8] summarized the situation with the following assessment: “The industry often promises what it cannot deliver, and then fails to deliver what it could, with better commitment and resources.” This state of affairs was attributed in part to the fact that the main performance measures being used to measure success were (and often still are) the amount of investment rather than outcomes.

Historically, most agencies have had limited usable feedback from signal infrastructure. Citizen complaint calls have been the primary source of information. Improvements such as retiming of signals would be scheduled on arbitrary, multi-year timelines. The only available quantitative data would be the input-oriented measures of effort expended by the agency—with little assessment of the actual performance outcomes. In the case where outcomes were assessed, this would typically be done by labor-intensive means such as a floating car study. Those studies could only cover a brief time period and only assess a subset of overall system objectives. The floating car study in particular emphasizes arterial progression, and does not provide the ability to balance that objective against other objectives such as minimization of excessive side street delay. Furthermore, these efforts are only feasible for individual intersections or at most a single corridor, thus betraying the sense of measuring true “system level” performance.

There is also a lack of cognizance of the interdependent nature of the subsystems involved in keeping traffic signal systems fully functional. The literature on new technology is brimming with advanced control concepts. Far less attention is given to the subsystems needed to support those concepts. Although there is certainly a *prima facie* case for supporting good maintenance practices (control systems will not work if their components do not), in practice this motivation is infrequently realized, for the same reason that signal timing is not kept up to date: there is a lack of system feedback to facilitate the business processes of maintenance and operation.

It is helpful to briefly consider the state of development for technology uses in traffic signal systems. Figure 1.1 presents a diagram that examines a handful of selected use cases for technology in traffic signal systems management

and their relative level of development. For now, this discussion excludes the details of traffic control methods. This is intentional for several reasons. For one, any statement regarding the “level of development” for specific technologies would be contentious, given the variation in agency practices. However, and more importantly, the task of performance management is necessarily independent of the selected traffic control method: free, coordinated, responsive, adaptive, manual, cyclic, non-cyclic, self-organizing, rolling-horizon, or any other combination of attributes that could possibly be imagined. For *all* methods of signal control, there is an assignment of right-of-way, or *capacity*, and a utilization of it by traffic, or *demand*. The effectiveness of the utilization, or the overlap between capacity and demand, can be analyzed to understand whether the system objectives are being attained.

The first two items in Figure 1.1, namely, clock synchronization and controller database upload/download, both have a very mature state of development. The first feature synchronizes multiple signal controllers in a system, and the second enables the engineer to remotely adjust the control parameters. These two features have been fairly common for over 30 years. Indeed, they are the most common functions of the “closed-loop” systems currently in widespread use. Some of these systems can report average green times and aggregated detector occupancies, but these rarely allow one to make detailed inferences about the quality of service.

The third item is real-time status display. This feature is in stark contrast with the limited feedback present in most closed-loop systems. Such displays enable the user to view the current state of the signal heads and the detectors, for all connected intersections, or to view video feeds of current traffic conditions. These are extremely valuable pieces of information for active traffic management and for manual spot-checking of the control plan. However, those processes are manual rather than automated.

Despite the enhanced functionality and data availability that real-time status displays can facilitate, they are limited in their ability to systematically identify intersection performance problems. This is because many of the performance validation abilities that real-time displays enable are done on a manual, rather than automated, basis. To evaluate how the system is doing in general or over a long period of time, it is not feasible to sit in the Traffic Management Center and make continuous manual observations. Although real-time status displays convey detailed information about the signal state, if that information is not written down and processed into performance measures that can express how the system state changed over time, then the greater potential for system monitoring is lost.

To partially address some of the shortcomings of manual data processing and observation-based conclusions about intersection performance, outcome-oriented performance measures are considered as the fourth item in Figure 1.1. Such performance measures have been the subject of considerable research in the past decade or so [2,9–23]. These are performance measures that focus on directly measuring the outcomes of system investments, in contrast to the more readily available investment- or input-oriented performance measures that quantify money or resource spent, rather than the outcome.

In the last decade, two automatically generated datasets have emerged that have allowed the implementation of data collection on a broad scale.

The first of these is the development of high-resolution controller event data. These data consist of timestamped logs of “events” occurring in a signal controller, with a time resolution of a tenth of a second or smaller. Typical events include changes of signal outputs (phases, overlaps, etc.) and of detector states, but could record any information internal to the control system. The “high-resolution” term is a relative moniker, intended to contrast this type of data with legacy volume and occupancy metrics that are

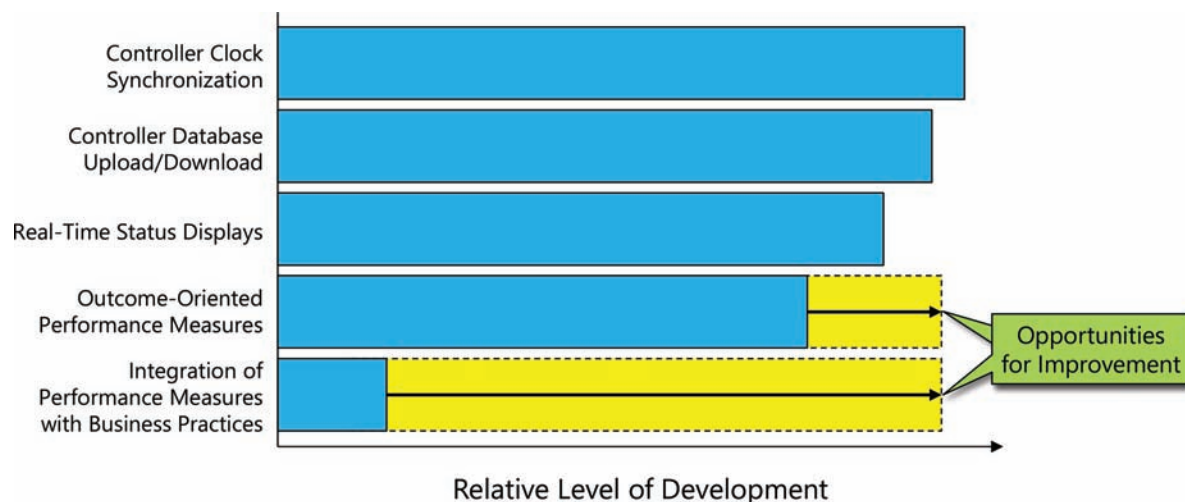


Figure 1.1 Comparison of the relative degree of development in use cases related to traffic signal systems.

aggregated on a minute-by-minute basis. Although these data existed in one form or another since the very beginning of traffic control, the ability to record them has only recently been applied to performance measure uses external to the control system.

Since the introduction of the high-resolution data concept [10], high-resolution data have evolved from a research tool into a standard offering now available from at least five North American controller manufacturers (at the time of writing) with a uniform specification [24]. In addition, there are other means of logging similar data for other controller types, such as external data collection units [25,26]. The ability of such data to monitor operations in great detail, and to improve operations, has been very extensively documented [2,9–23].

The advantage of high-resolution data logging is that it can be implemented modularly with routine controller upgrades, and does not necessarily require capital infrastructure projects to build out systems. What sets this strategy apart from existing monitoring solutions, such as real-time status displays, is the concept of *logging by the controller*, which means that low-latency, always-on communication is not strictly required, since the events are written down locally and then harvested for analysis when the next opportunity occurs.

Another enabling technology is related to the increasing pervasiveness of mobile cellular devices. Because of the nearly ubiquitous distribution of these among the vehicle fleet, it is now possible to obtain highly detailed records of vehicle travel times and average speeds along roadways, with little or no investment in new infrastructure, through vehicle re-identification-based travel times and average probe vehicle segment speeds. These new datasets have proven extremely useful for evaluation of freeway facilities and quantifying congestion on large scales [27–31]. That utility has also been demonstrated for arterials [32,33].

The modular nature of these developments, scaled toward incremental growth, marks a potential sea change in implementation strategies compared with traditional methods of implementation. Whereas most proposed technological solutions start from the *top down*—for example, with a high-performance central system and/or a novel traffic-control algorithm—other opportunities offer the possibility of taking on the problem from the *bottom up*, through incremental, modular improvements that are achievable with investments that are better aligned with many agency budgets.

With these enabling technologies in mind, we return at last to the fifth item in Figure 1.1: **Integration of performance measures with agency business processes.** The greatest prospects for advancing the state of the practice occur when the performance measures are used to inform and to drive agency decision making. In addition to better scoping the resources that have been dedicated to signal systems management, the ability to articulate the system state and to define the returns on past investment would also afford system operators

greater confidence in asking for more resources. The present report focuses on this item.

The traditional solution to signal system deficiencies is to seek a capital project for equipment upgrades en masse, or advanced control systems intended to solve all of the problems at once. However, the effectiveness of these solutions is relatively unknown, because most of them still do not provide any additional feedback from which to draw conclusions. Whereas the capital project approach to procurement can help make gains for critical parts of the arterial network, one corridor at a time, the problem of better programming resources at the agency level really needs data collection to be expanded more rapidly and more broadly. It is difficult to envision how this can be done via large capital projects while still enabling a competitive, multi-vendor environment. This suggests that a *modular* approach could be more feasible. This option is not often considered for system implementation. Often, the culture of public works generally orients itself toward ribbon-cutting events rather than incremental improvement. This overlooks opportunities for improvement, which despite perhaps having subtler immediate effects, could have enormous cumulative potential.

1.3 A Language for Objectives

To begin integrating performance measures into agency business processes, even before considering which business processes are affected, it is necessary to first identify the objectives to be achieved by making the investment. Performance measures are not implemented for their own sake, but rather to achieve improvements. Therefore, we need a language with which to express what those achievements should be.

The goal is essentially to answer the question: **“What should a traffic signal system do?”** A starting point for building an answer is to take a very high-level view. The provision of safe and efficient separation of right-of-way at intersections encompasses the most fundamental objectives. The following statement, taken from a recent FHWA report [8], succinctly states a basic general goal for the system:

“We will do our best to avoid making drivers stop, but when we must make them stop, we will delay them as little as possible, within the context of safe operation.”

Each individual agency would have its own more specific objectives that relate to how this overall goal would be obtained. The Basic Service Model related to the above quotation, for example, breaks out five specific objectives in more detail:

- *Field Infrastructure Reliability.* The physical components of the system must be kept in working order for the overall system to work properly.
- *Minimizing and Balancing Congestion.* Green times should be efficient and equitable, balancing the needs of the heaviest movements of traffic without causing undue delay to lighter movements.

- *Smooth Traffic Flow*. Progress traffic (i.e., enable vehicles that are in motion to remain in motion) when it is reasonable to do so, as through groups of intersections relatively close together.
- *Consistent and Predictable Response*. Provide signal control that does not excessively defy the expectations of users.
- *Signal Timing Versatility*. Provide signal control that is capable of serving a variety of likely field conditions in an effective manner.

To consider another perspective, the *Signal Timing Manual* [34] lists 15 specific objectives for different modes. Table 1.1 sorts these into a table, which forms a potential “menu” of services to be provided by the system, if any of the listed modes are present. Although this list does not enumerate every possible objective, it does contain those that are likely those commonly regarded as important for many agencies.

After organizing these objectives into a table view, it is possible to generalize them across all modes, as follows:

- *Safety*. Minimize the likelihood that system users of the system (and nonusers) will be harmed.
- *Mobility*. Avoid incurring delay for travelers as much as possible—i.e., attempt to make the travel time as brief as possible.
- *Accessibility*. Ensure that all users have access to the service/facility.
- *Environmental Impact*. Minimize the amount of pollution generated by the transportation activity and/or make it environmentally sustainable.
- *Queue Length Management*. Prevent queues from interfering with other traffic movements, as much as possible.
- *Operating Cost*. Minimize the cost of the transportation activity and/or make it economically sustainable.

These objectives are primarily oriented toward system *users*. It is also important to consider services provided to system *operators*. Examples include the following:

- *Equipment maintenance*. Ensure that the deployed components (communication system components, vehicle detectors, signal controllers, etc.) are in working order.
- *Diagnostics*. Validate and resolve reported problems (as from citizen complaint calls), and proactively detect such problems.
- *Asset management*. Inventory deployed components, and digital “assets” such as software or control parameters.
- *Development of datasets*. Compile information for administration, planning, incident management, public relations, and other agency needs.
- *Demonstration of accountability*. Report system performance and longitudinal trends to correlate with system investments, for executive reports, budget justification, and other purposes.

Connecting these ideas back to the Basic Service Model concept, we can begin considering notions such as reliability or predictability. A model that breaks out several dimensions of service quality in business operations is the “RATER” model [35]. The acronym stands for Reliability, Assurance, Tangibles, Empathy, and Responsiveness. These are defined in Table 1.2, and translated into terms that are applicable to traffic signal systems.

Most of these elements of service quality are easily interpreted into a traffic signal systems context:

- “Reliability” refers to whether the service quality can generally be expected to have similar properties over time—that is, whether it is predictable.

TABLE 1.1
Signal system objectives by mode, based on discussion in the *Signal Timing Manual* [34].

Objective	Mode of Transportation			
	Vehicle	Pedestrian	Bicycle	Transit
Safety	Minimize vehicle collisions and conflicts. Provide vehicles with sufficient time to execute movements.	Minimize pedestrian collisions and conflicts. Provide pedestrians with sufficient time to execute movements.	Minimize bike collisions and conflicts. Provide bikes with sufficient time to execute movements.	Minimize transit collisions and conflicts. Provide transit with sufficient time to execute movements.
Mobility	Serve vehicle movements efficiently and equitably. Progress vehicle routes according to priority.	Serve pedestrian movements as efficiently as possible.	Serve bicycle movements as efficiently as possible.	Serve transit movements as efficiently as possible.
Accessibility		Ensure special needs groups can execute pedestrian movements.		Ensure special needs groups can access transit.
Environmental Impact	Minimize pollution produced by traffic.	Facilitate use of alternative modes.	Facilitate use of alternative modes.	Facilitate use of alternative modes.
Queue Length Management	Prevent spillback of queues at critical locations.			Prevent blockage of transit routes.
Operating Cost	Minimize stops and delays.	Minimize delays.	Minimize delays.	Minimize delays.

TABLE 1.2
The RATER model of service quality [35], applied to traffic signal systems.

General Service Dimension	Description	Application to Traffic Signal Systems	Description in a Signal Systems Context	Example Drill Down
Reliability	The service is dependable and accurate.	“Reliability”	The quality of service is likely to be similar over time.	It takes about the same amount of time to get through the corridor every morning.
Assurance	The service quality builds trust in the system.	“Robustness”	The system effectively handles a wide variety of operating conditions.	The “off-peak” timing plan works well from 9:00 to 14:00 and on weekends.
Tangibles	The physical qualities of the system are acceptable.	“Consistent and Predictable”	The system fulfills basic user expectations.	The timing plan is configured to clear long queues to minimize split failures.
Empathy	The system is attentive to the needs of all stakeholders.	“Equitability”	The system is attentive to the needs of all stakeholders.	Excessively long cycle lengths are avoided so that arterials do not restrict pedestrian mobility.
Responsiveness	The system provides help and answers complaints promptly.	“Responsiveness”	The system can adjust to accommodate expected or unexpected changes on different time scales.	The maximum time automatically increases to handle a spike in demand.

- “Assurance” really means that the system can be trusted to work well under a variety of conditions, which has been expressed with the concept of “robustness” [36,37].
- The notion of “tangibles” can be related to the degree of comfort provided to system users. This includes the notion of “consistent and predictable” expressed in the Basic Service Model.
- “Empathy” is perhaps more commonly expressed in traffic engineering as “equitability”, meaning whether the operation can be said to provide reasonable level of service to all those who it affects.
- “Responsiveness” means the ability of the system to react to changes in the operating environment, which may occur suddenly or gradually.

Having identified some language to express potential system objectives, it is not difficult to imagine some constituents saying: “Yes. Do *all* of these things, and make it reliable, robust, comfortable, equitable, and responsive.” However, such a solution is not realistically attainable. Rather, the objectives have to be prioritized.

The reason for this is that some objectives are in tension with others. There is, for example, a fixed amount of green time at an intersection, and it is possible to draw an analogy with the concept of a “production potential” in economics [38]. Based on a limited amount of available resources, it is only possible to produce a certain amount of each competing product. Generally, because of the principle diminishing returns, there is a tendency for efficiency to diminish as one particular product is increasingly emphasized.

Consider the competition between two streets for green time at a signalized intersection. The quality of service for each street can be measured by the amount of delay incurred on each. Potential statements of two competing objectives could be stated as follows: “Minimize delay on

the north–south street” and “Minimize delay on the east–west street.”

Figure 1.2 shows a plot of the production frontier in terms of the delay on both movements [38]. It is desirable to be as far to the bottom and to the right as possible in this plot, but because of the fixed amount of resources, it is infeasible to reach beneath the curve (points well above the curve are possible, and represent suboptimal operation). It is possible to achieve extremely low delay for either the east–west street (point “A”) or the north–south street (point “B”), but only at the cost of extremely high delay for the other. Rather, a reasonable

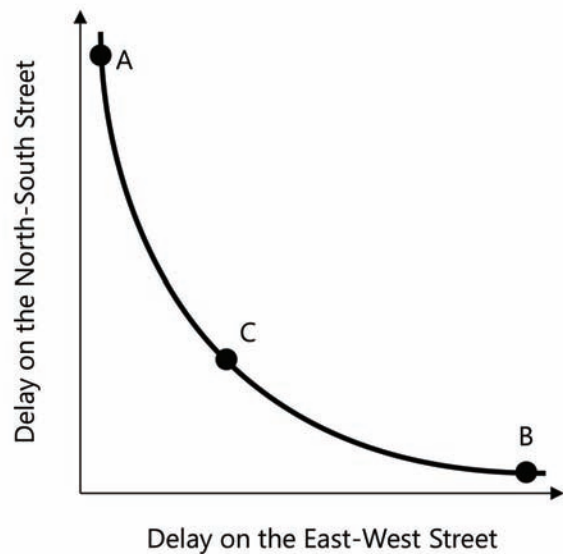


Figure 1.2 The concept of a Pareto Front applied to two competing objectives in signal operation, after Gartner et al. [38]

solution would seek a balanced compromise between the two (point “C”).

The same concept can be applied to many of the different services that we stated earlier. For example, the two vehicle mobility objectives in Table 1.1, equitable and efficient service versus progression along routes, are not always in harmony with one another. Consider a corridor of multiple intersections, some of which are highly utilized while others are not. For coordination, it may be necessary to use a long cycle length to progress traffic while also accommodating the busiest intersections. For some of the other intersections, using that cycle length will likely introduce additional delay that would not be needed if those intersections were dropped from coordination. This is a good example of a tradeoff between two objectives in tension.

Because many of the services compete with each other, the engineer must establish priority among them. To do so, the systems engineering approach would be to consult the stakeholders who are involved in the planning, design, operation, and maintenance of the system. These are the actors who participate in the business processes, who make use of the services provided by the system, who experience the characteristics of those services, and who are the potential audiences for performance measures that measure those characteristics. The main stakeholders are as follows:

- *Engineers.* Depending on the agency, the engineering staff manage signal-related projects, which might be done in-house or contracted out to consultants or other agencies (as in “home rule” states). These individuals need to independently verify whether the system is performing as expected.
- *Technicians.* Equipment is installed and maintenance is performed by technicians. This also may be done in-house or contracted out to consultants or other agencies. Technicians need to know whether the equipment is functional.
- *Vendors and consultants.* Traffic signal projects involve the use and/or purchase of control equipment and may involve consultants or bring in expertise from the vendor to verify and validate the systems.
- *Agency administration and planning staff.* Executives, managers, and planners are usually concerned with demonstrating that investments are yielding a benefit and with knowing the state of overall performance of the overall network or of an individual facility.
- *Users.* The system users become involved as the system is operated and maintained, and are interested in their travel characteristics. Regarding traffic signals, users are unforgiving of highly visible “inefficiencies,” such as extremely long red times while the signal serves another movement with no traffic, or when an expected phase is “skipped,” etc. Users generate complaint calls that sometimes are the only available system feedback, but are often difficult to substantiate, and therefore to resolve. This is a broad category that can be broken down further by mode.

Traffic engineers engage in these processes whenever they determine the control settings—regardless of whether they are mindful of it or not. Consider, for example, the use of a feature such as “Rest in Walk” on a corridor. This

feature relates to the “default” behavior of an actuated traffic signal when there is no demand. The default behavior of most actuated signals is to rest in green on the mainline. The Rest in Walk feature causes it to also remain in the Walk state for the compatible pedestrian movements. When a vehicle or pedestrian arrives on a conflicting movement, that user cannot proceed until after the mainline Walk and Ped Clear intervals have finished timing. Resting in walk benefits pedestrians moving *parallel* to the mainline—and perhaps secondarily benefits any parallel traffic by extending the green. The decision whether to use the feature asks the engineer whether pedestrians using those movements should be prioritized over users making other movements. This could be needed in some systems, whereas for others it would not.

To summarize, agencies determine the objectives to be pursued by their signal systems, and set the relative priority of these. The discussion provided here has developed a provisional menu of objectives and characteristics. Common signal system objectives, by mode, have been presented in Table 1.1. Desirable characteristics of system performance for any of these objectives have been presented in Table 1.2. By expressing the needs of an agency in this language, it becomes possible to connect a service objective to specific performance measures, and thus to begin *measuring* the service quality.

1.4 Interdependency of Components

The previous section discussed objectives as combinations of a service for the system to perform. So far, this discussion has been somewhat external to the system. An examination of the technical details of the system will reveal that many objectives are dependent on the success of others.

Modern traffic signal systems are essentially a networked computing application. The International Organization for Standardization (ISO) proposes a “seven-layer” model called the Open Systems Interconnection (OSI) model [39] that explains the functionality of networked computer systems. Figure 1.3 presents a diagram of the OSI model, with the seven layers numbered from bottom to

Layer	Examples
7. Application	Web browser
6. Presentation	Encoding of text, graphics, etc.
5. Session	Management of data exchange
4. Transport	Exchange of data segments
3. Network	Node identification and routing
2. Data Link	Transfer of bits between nodes
1. Physical	Electrical connection

Figure 1.3 The seven-layer Open Systems Interconnection (OSI) model.

top in order of their complexity and dependency. Each higher level element depends on the functionality of all the lower level layers. The figure includes some examples that illustrate how each layer would play a role in the application of a web browser (a software utility that exists on practically every smartphone, computer, and tablet in existence today).

From the perspective of the OSI model, the traffic engineering community could be regarded as focusing mostly on the uppermost layer, while in practice the lower layers are not always stable enough to facilitate all of the desired applications.

A few nuances are particular to traffic signal systems that set them apart from a more generalized networked computer application:

- Signal systems deal with not one, but at least *three* types of communication systems: (1) connections to the outside world (such as Ethernet or cellular IP); (2) connections within the traffic cabinet (such as through the SDLC bus in a TS/2 Type cabinet); and (3) connections to local field devices such as detectors. The signal control will completely fail if the cabinet connections break, but will degrade rather than fail if the other communications stop working.
- Signal systems are intended to be completely automated, always-on, and continuously running, differentiating them from on-demand types of applications.
- Signal systems are relatively robust at providing functionality within the application layer, even as various components fail. For example, a failed detector will cause most controllers to place the associated phase in recall. If the controller itself malfunctions, other cabinet equipment will take over and place the intersection in flash. What this means is that the system can tolerate many minor failures (which cause operational deficiencies), and these will often go unnoticed as long as the lights are continuing to change through green, yellow, and red.

With these characteristics in mind, Figure 1.3 presents a simplified hierarchy of signal system needs inspired by the OSI model but focusing more on the roles played by engineers and technicians in traffic signal operations. As with the OSI model, however, performance in the upper levels requires that all of the lower levels function. Conceptually, the upper layer functions deal with different aspects of traffic control, and are related to the upper layers of the OSI model dealing with exchange and use of data. The lower levels are more related to the functionality of system components and are related to the lower levels of the OSI model.

From bottom to top, the layers in Figure 1.3 are as follows:

- *Detection*. This refers to the situational awareness of signal controllers to their physical environment, which is currently achieved by using detectors (e.g., vehicle detectors of different types, and pedestrian push-buttons). In the future, this layer will likely be augmented with connected vehicle data.
- *Communication*. This means the ability to remotely reach signal controllers and other equipment. In the past, this was mainly achieved mainly through dial-up connections,

but today it is increasingly managed via internet connections such as by cellular IP drops [40].

- *Local control*. This refers to the phase-switching operations of a signal controller at the local intersection, which are accomplished based on the detection input and the control parameters. Although this functionality will persist even if detection and communication systems fail, it will be considerably degraded.
- *System control*. This refers to signal coordination, or the creation of structured patterns of the local phase-switching operations to establish relationships within groups of intersections. An example is the creation of a green band or green wave along a corridor to facilitate smooth traffic flow.
- *Advanced applications*. The uppermost layer refers to more sophisticated system functions that rely on the functionality of all the lower layers. An example would be the adaptive adjustment of control parameters in response to measured changes in traffic conditions.

The key element expressed in Figure 1.4 is the *dependency* of each set of functions on the lower levels. This implies that some objectives have inherent dependencies on others. These relationships are usually implicit in the proposal of any traffic control system. However, a focus on end applications can lead to the undervaluing of good maintenance practices.

At present, if we consider which system components could be called extremely reliable, with a high probability of being “always on” for most of the time, this is really only true of the signal controller, conflict monitor, and signal heads. Failures of those components are extremely visible and are completely unacceptable to the system operation. The other components, in contrast, are far less visible and therefore unlikely to generate an active response. For example, failed detectors effectively put phases into recall states. Because the signal continues to cycle through green, yellow, and red on every movement, the operation is still considered acceptable even though it may be inefficient.

Communication failures are even more subtle than detector failures. If two neighboring intersections cannot communicate, they will be unable to coordinate effectively. However, it is impossible as an external observer (such as a driver) to tell the difference between poor progression caused by bad signal timing and that caused by unsynchronized controllers. Once again, without active system feedback—that which is acted upon—the problem is likely to go unnoticed.

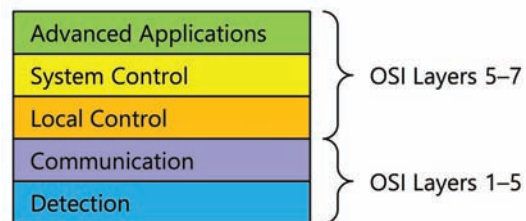


Figure 1.4 Translating traffic signal control components into a hierarchy of needs.

If this is true for present systems, what of more advanced systems that are supposed to be fully automated? Those systems tend to focus mostly on the high-level applications in the hierarchy of needs; “maintenance” is not generally a selling point for them so much as a promise of greater responsiveness and fewer engineer hours devoted to adjusting signal timing. However, if those systems do not provide active feedback to locate component failures (where they are now and their history), then they too will surely suffer from the same types of maintenance problems that plague conventional systems. One might also ponder whether the operator expectation that the system will look after itself would encourage them to pay *less*, not more, attention to potential problems.

To summarize, the old adage “you have to learn to walk before you can run” is eminently applicable to the present situation of signal maintenance. The problem of underinvestment in infrastructure is much larger than the world of signal systems, and it is to some degree rather nontechnical. However, the technical characteristics of signal systems acts to *conceal* problems of maintenance, such that noncritical failures can go unnoticed for very long periods of time. To improve maintenance is to win incremental gains in the higher level performance of the system.

1.5 Business Processes

So far, the discussion has focused mainly on the technical side of traffic signal systems. The other side is the institutional side, namely, the business processes and the personnel involved in the deployment and management of signal systems. An entry point to begin discussing how these can be managed is the “V diagram” from systems engineering that conceptualizes the life cycle of an engineered system. Figure 1.5 presents a view of this model. The left-hand side of the V represents system *definition*, wherein it is determined what the system should be doing, and what the needs are to achieve that. The right-hand side of the V represents *evaluation*, wherein the components are made to work properly. The steps in the process are as follows: The left-hand side of the V diagram represents the system definition

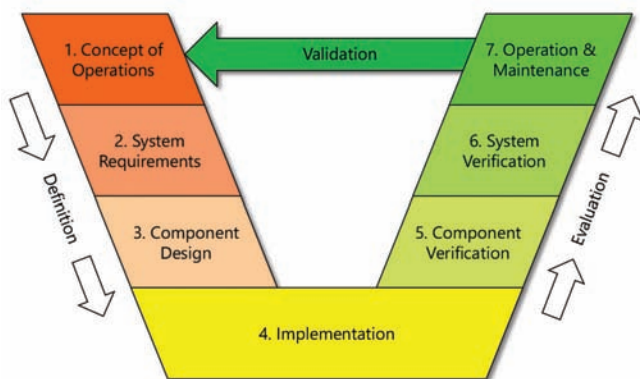


Figure 1.5 The systems engineering “V diagram.”

process. The implementation phase is shown along the bottom. Here, project development pivots from a “defining” phase to the “evaluation” phase represented by the right side. A feedback loop is created by the final validation step.

The articulation of a concept of operations establishes the system requirements that drive the design process. In signal systems, this generally means the adjustment of specifications to achieve specific types of operation. Implementation and evaluation are where these are realized in the field. Outcome-oriented performance measures are vital to the verification processes that ensure the system requirements are met, and the validation process that ensures the system satisfies the user needs [41]. Validation offers the opportunity to evolve the concept of operations.

When introducing this systems engineering concept, it is important to make a distinction between the somewhat idealized view that can be attached to the V diagram and the way in which systems are actually developed. One may argue that most existing signal systems were not developed by a formal systems engineering process. For newly constructed roadways, the installation of traffic signals is a task given to traffic engineers after the pavement has set, so the systems engineering process occurs in a greatly abbreviated fashion. Perhaps an even greater number of “systems” were not designed as such, but are de facto systems that developed organically over many years, with “corridors” forming from isolated intersections that grew together over time.

However, the same systems engineering processes still occur as the system comes together, even if those involved are not cognizant of it. In the example of a corridor that developed from isolated intersections, the initial concept of operations was most likely developed from a template idea of how a fully actuated intersection should work. The system requirements and components needed for that type of operation are relatively straightforward, and the verification stages could be limited to mostly ensuring that the detection works. Operation and maintenance of isolated intersections is also rather simple, as fully actuated control will generally perform as expected so long as the detection is in working order. The *validation* step occurs after the corridor has developed to a point where fully actuated operation is no longer adequate, and a critical reassessment finds that it is necessary to develop a coordinated timing plan, which requires a new concept of operations. Again, this too is likely would follow a template based on how other agency corridors are managed.

One of the most important elements of signal control is the *control policy*, which is the set of parameters that determines how the signals will operate. The life cycle of a signal control policy offers a number of examples to demonstrate this in more detail. Figure 1.6 shows a block diagram showing the activities that are part of this process, with six steps defined as follows:

1. *Define Objectives, Assess and Prioritize Activities.* Equivalent to developing a concept of operations, at an early stage,

the engineering team will assess and prioritize the needs of the system to be achieved in the subsequent steps.

2. *Gather Information Needed to Develop the Control Plan.* The design of a control policy requires some intelligence about the likely level of demand in the corridor. The conventional way of obtaining such information is through turning movement counts. Other information, such as intersection geometry, presence of detection, and so forth, is also needed.
3. *Model System Performance in Software.* The information gathered in Step 2 is fed into a model that yields a control policy. Typically, this consists of entering traffic volumes into timing plan design software and optimizing the timing within that system.
4. *Design and Document Control Parameters.* The initial design of the control is developed with assistance from software. This process is often iterative, with the engineer making tweaks to the design based on intuition or feedback from simulation. The iterative step is shown as Feedback Loop 1 in Figure 1.6. This is equivalent to the systems engineering Definition processes. The result is a set of parameters (such as a coordination pattern, or controller logic to allow more complex operation) that can be taken to the field.
5. *Implement Control Parameters in the Field.* After preparing an initial design, it is documented and programmed in the signal controller in the field. Translation of the control parameters from a design environment to a field environment is sometimes needed [42].
6. *Monitor System Performance.* After initial programming of the field controllers, the engineer will typically spend some time in the field to make adjustments in response to the response of actual traffic. This field fine-tuning step is shown as Feedback Loop 2 in Figure 1.6. This is equivalent to the systems engineering Verification processes.

The third feedback loop linking step 6 back to step 1, in which the overall system objectives are occasionally reassessed, is equivalent to the systems engineering Validation process.

In the current state of practice, Feedback Loop 1 in Figure 1.6 is the place where the existing tools for documentation and assessment are the most developed. There are several mature software tools for adjusting signal timing and simulating how signal systems will likely work in the field, and it is possible to generate an

enormous amount of modeled data from these and to refine a control plan.

In contrast, Feedback Loops 2 and 3 have historically lagged behind in terms of data availability, which has limited practitioners to mostly incremental adjustments based on manual observations for the most part. Feedback Loop 2 has been improved by some technological developments. Tools such as real-time status displays and video feeds have made these manual tweaks more convenient, and other tools such as adaptive control have attempted to automate the adjustment process. However, these technologies do not scale to system-wide coverage, and they do not provide the ability to make the longitudinal performance assessments that would be needed for Feedback Loop 3.

The three feedback loops represent specific business processes relevant to signal timing, all three of which would be improved by use of automated performance measures:

- *Feedback Loop 1*, where signal control policies are designed, relies on data that historically have been obtained manually (e.g., turning movement counts). Automation of that data collection process was proposed over 10 years ago, using existing detection systems [43]. High-resolution data provide a second means of obtaining the equivalent volume data. There is a tremendous potential to build on this concept, not only through vehicle counting but also by calibration of the design models [44].
- *Feedback Loop 2*, where existing signal control policies are adjusted to field conditions, benefits tremendously from the availability of detailed operational information to verify that the control policies and adjustments to them achieve what they are intended to do. This is one of the key areas where performance measures can increase agency capabilities, as has been demonstrated by their extensive use by engineers in Indiana and Utah [45,46].
- *Feedback Loop 3*, where overall system assessments are made, would also benefit from the availability of information to validate the effectiveness of the control plan over time and to determine whether objectives are being met or whether the policies need to be revised. This runs the broad gamut of all high-level arterial decisions, from whether or not to implement coordination [47], to determining where bottlenecks exist in the system [48], to producing executive reports for administration [49].

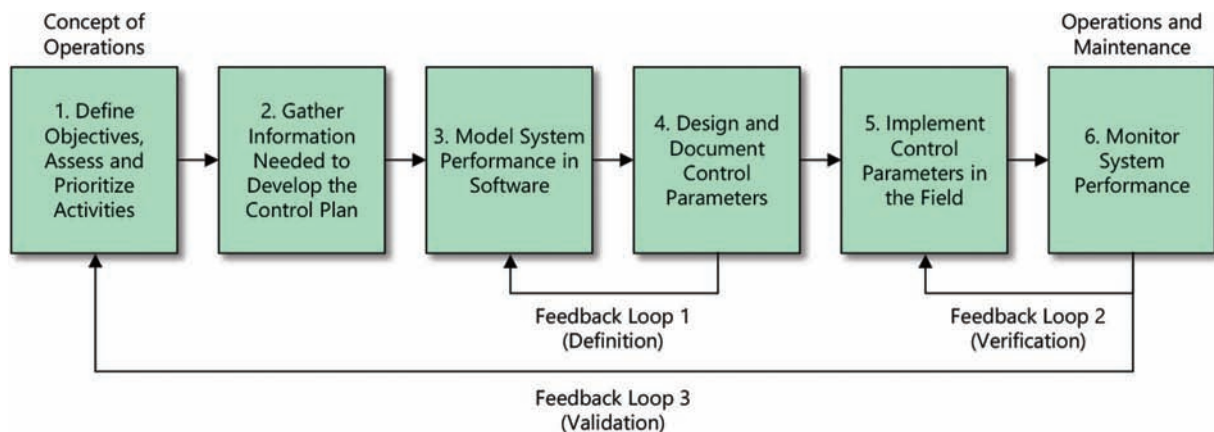


Figure 1.6 Life cycle of a signal control policy.

The three feedback loops represent key business processes that are part of the greater systems engineering process. Local control, system control, and advanced applications are all dependent on the detection and communication components that enable them to work properly, as shown in the hierarchy of needs presented earlier (Figure 1.4). Therefore, another cyclical process that needs to be examined from the systems engineering perspective. Figure 1.7 expresses this process as a block diagram representing specific activities related to the life cycle of any component of a signal system, from a vehicle detector to a central system. The six steps are as follows:

1. *Define Objectives, Write Specifications and/or RFPs.* Initially, system needs must be defined and translated into documentation to facilitate procurement. For example, the range of expected performance for a vehicle detector might be specified [50].
2. *Make Initial Selection.* Based on results of Step 1, the agency must make some initial decisions on procurement, such as by selecting candidate component alternatives that could be used in the field.
3. *Test Component Performance on the Bench.* Initial tests of a component would typically be done in a controlled environment. The terminology of testing “on the bench” refers to the manner in which signal controllers can be programmed and tested before they are put into the field, but for some other components this might consist of a limited deployment at one or more test sites, or running an advanced control system component in a human-supervised mode before allowing it to run on its own.
4. *Adjust Component Configuration on the Bench.* A test environment affords sufficient diagnostic capability to enable component adjustment. During this “bench” testing, the component will be adjusted to achieve the expected performance. For example, the spatial configuration of vehicle detectors might be adjusted to achieve expected performance [51]. This iterative process is shown as Feedback Loop 1 in Figure 1.7.
5. *Deploy Component in the Field.* If the component passes the initial evaluation in a test environment, it will then be deployed in a *non-test* environment.
6. *Monitor Component Performance.* After the initial deployment and turn-on of the component in the field, there will usually be a period of observation that will verify that the component works. Often, some fine-tuning is required to make the component perform as

expected, hopefully drawing on lessons learned in the test environment. An example would be the adjustment of the sensitivity of an inductive loop detector. The adjustment process is represented as Feedback Loop 2 in Figure 1.7.

The third feedback loop, Validation, is once again a higher level assessment of whether the system is achieving its overall objectives.

Similar to the signal control policy life cycle, Feedback Loop 1 in Figure 1.7 is the area where the most data can be collected, with the component being evaluated and its configuration modified in a test environment. The other feedback loops (2 and 3) do not typically have the same measurement capabilities; therefore, the feedback process is limited to what can be observed by an engineer or technician installing the equipment and verifying it. If the equipment does not work, or if it loses functionality over time, this can often go unnoticed. Preventative maintenance on a regular schedule can mitigate this to some degree. However, similar to scheduled signal retiming, some of those resources could be targeted to known problems, if there was a way to detect them at a broad scale.

Another set of business processes that relate to control policy and component life cycles are those of planning and programming. These represent perhaps the most important processes, because they determine the amount of resources that ultimately will be invested in arterial management. Traffic signals must compete with other potential agency investments, such as improving pavement condition or safety performance. There is a long history of documented return on investments in such areas, such as the crash reduction factors that relate to safety improvements, or pavement condition, which can be rather easily predicted based on the expected traffic demand and the amount of investment provided. It is still rather difficult to make similar predictions of traffic performance with the same degree of confidence under alternative investments.

The present report includes one assessment based on five years of performance-based management on one corridor, which yielded a net benefit of approximately \$3.6 million (Figure 1.8) [52]. That result was obtained for one particular corridor; expansion of analysis over more areas that is sustained over time will greatly improve the amount of information that engineers can

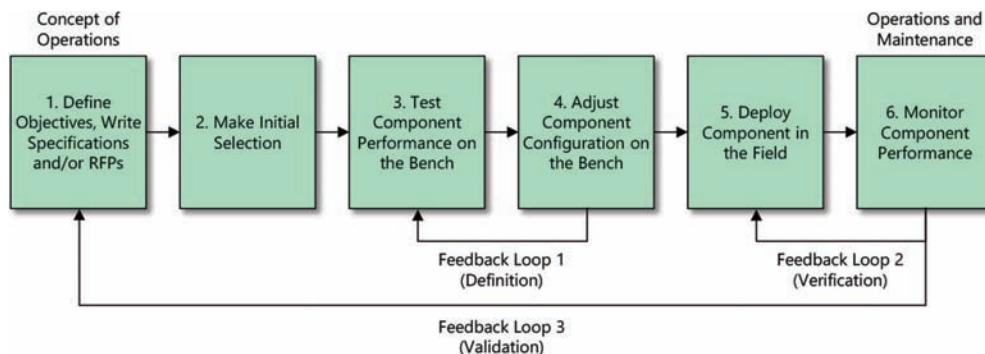


Figure 1.7 Life cycle of signal control equipment.

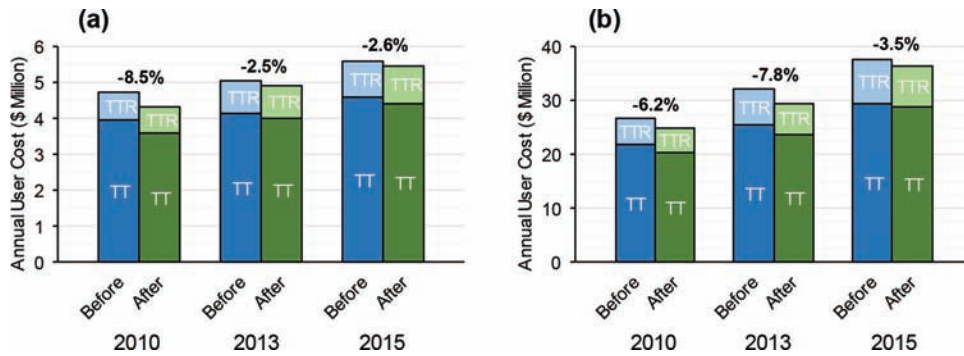


Figure 1.8 Tracking of user costs for (a) Saturdays and (b) Weekdays across a 5-year period for operation on SR 37 [52]. Three “before” and “after” periods are included in the analysis relative to when signal offsets were optimized. The user cost reductions

potentially use to ask decision makers for more investments in signal systems.

1.6 Organization of the Report

The present report demonstrates uses of performance measures to support the business processes represented in the feedback loops of Figure 1.6 and Figure 1.7. The introduction of performance measures into these business processes can greatly improve their effectiveness, as has been demonstrated extensively by those agencies that have pioneered them [2,7,45,46]. This report documents those efforts with the objective of providing an informational guide for those who are intending to implement a similar program for their own agency.

While the objectives of each individual agency will vary according to their stakeholders, this document attempts to capture some use cases that most agencies will have in common. Table 1.3 provides a high-level overview of the performance measures incorporated into this report, tying these into the five objectives of the Basic Service Model [7]. The table also maps the objectives and performance measures into specific chapters of the report. These segment into the following discussion topics:

- *Field Infrastructure Reliability.* Performance measures to identify detection and communication system failures are presented. These are demonstrated for system-wide views for the INDOT signal network.
- *Minimizing and Balancing Congestion.* Performance measures that characterize the level of demand for service,

TABLE 1.3
Uses of performance measures at different stages of the signal systems engineering life cycle.

Objective	Main Layers (Figure 1.4)	Report Chapters	Main Performance Measure Audiences	Example Performance Measures and Visualizations
Field Infrastructure Reliability	Communications, Detection	Chapter 3 Chapter 4	Engineers, Technicians, Vendors, Consultants	Diagnostic analysis Communications/Detector Health
Minimizing and Balancing Congestion/ Intersection Safety	Local Control	Chapter 5	Engineers, Administrators, Users	Pedestrian Utilization Vehicle Flow Rates Volume-to-Capacity Ratio Phase Termination Type Red/Green Occupancy Ratio Estimated Delay Red Light Running Oversaturation Severity Index Time to Service
Smooth Traffic Flow	System Control	Chapter 6	Engineers, Administrators, Users	Time Space Diagram Percent on Green Platoon Ratio Purdue Coordination Diagram Cyclic Flow Profile Measured Travel Time
Signal Timing Versatility	Local and System Control	Chapter 5 Chapter 6	Engineers	Longitudinal analysis
Consistent and Predictable	Advanced Applications	Chapter 5 Chapter 6	Engineers, Vendors, Consultants	Longitudinal analysis Diagnostic analysis of individual events

and the quality of the service provided, are presented for vehicle, pedestrian, and transit modes. The performance measures evaluate mobility and safety.

- *Smooth Traffic Flow.* Performance measures for evaluating the quality of progression are presented. Methods for using the same data to predict and optimize the traffic control are demonstrated.
- *Consistent and Predictable Response.* The longitudinal performances of signal systems across spatial and temporal distributions are evaluated from a variety of perspectives to locate problems and assess the quality of service over time.
- *Signal Timing Versatility.* Detailed operational impacts of control features are examined across a variety of operating conditions.

In addition to the above, Chapter 2 of this report provides a discussion of the requirements needed for the performance measures used in these examples, including a discussion of detection configurations and possible performance measures that can be obtained from these using the high-resolution data, as well as performance measures that could be developed without high-resolution data. The chapter also provides an overview of the data infrastructure needed to support the use of the data, as it requires an investment into a means of storing and retrieving the data.

1.7 Conclusion

This chapter described several issues relevant to the business processes by which traffic signal systems are developed, operated, and maintained. The discussion touched on technical and nontechnical aspects of the overall problem and dissected the main business processes related to the development of the signal timing and procuring components for signal systems. By connecting these specific business processes to performance measures, agencies can greatly improve the quality of operations. The remainder of this report examines the possibilities in much more detail, assisted with numerous examples of recent accomplishments.

2. REQUIREMENTS FOR SIGNAL PERFORMANCE MEASURES

This chapter presents the technical requirements that are necessary for the implementation of the suite of signal performance measures. The requirements are introduced from an objective-based perspective and identify the technologies needed for bringing together the infrastructure, software, and processes for generating performance measures that would satisfy the initial objectives defined.

2.1 Introduction

As high-resolution controller data logging becomes more widely available on newer NEMA and advanced transportation controller (ATC) controllers, there is a need for agencies not only to acquire these new devices, but to also be cognizant of the complete set of requirements for detection, communication, and software and database systems and processes that are necessary for

transforming the high-resolution data into signal performance measures. There is a need to clarify these requirements and develop a matrix matching detector configurations, storage capacities, and latency for satisfying minimum functionality baselines achievable under such configurations. Identifying these components effectively facilitates the collection, transmission, storage, and visualization of data for signal performance measures.

When agency operating objectives are well defined, performance measures that fit the criteria for achieving those objectives can be targeted for implementation. Consequently, the resources necessary for the production of the desired performance measures can be allocated depending on how effectively the agency's business processes were designed to meet those needs. For instance, although most agencies have considerable experience in designing intersection detector layouts for various *operational* use cases, designing for data collection is a relatively new use case, and there is a lack of institutional knowledge on what design variants can be recommended to accomplish agency objectives. If a location is only equipped with advance detection for dilemma zone protection, but stop bar occupancy data are desired, agency business processes should accommodate for such a need with new detector layouts so that resources can be allocated for stop bar detection installation, configuration, and subsequent data visualization.

2.2 Cabinet Types

Various vendors have developed signal controllers capable of logging high-resolution event data compatible with the NEMA TS-2 Type 1 and Type 2, ATC 170 and 2070 cabinet standards. Generally, each of these formats operate some type of communication interface between the detectors, preemption units, various flash and telemetry hardware, power management, and the controller itself. While each standard has its own connection interfaces and form factors, open architectures such as the NTCIP communications protocol are becoming more pervasive across standards. For example, in the TS-2 Type 1 cabinet (Figure 2.1) the Synchronous Data Link Control (SDLC) serial data bus facilitates data communication between the various load switches and detection racks to the controller. On the detector racks, Bus Interface Units (BIU) interface detector data with the SDLC bus. The data are typically assigned to individual communication channels distinguishing each detector coming into the SDLC. Once the information is received by the controller, configuration parameters determine the channel numbers that will drive the actuation for any assigned outputs. The same information can be timestamped and recorded in controllers with data logging capability.

2.3 Controllers and High-Resolution Data Logging

Historically, most signal controllers provided only a limited amount of data, which consisted of volume and

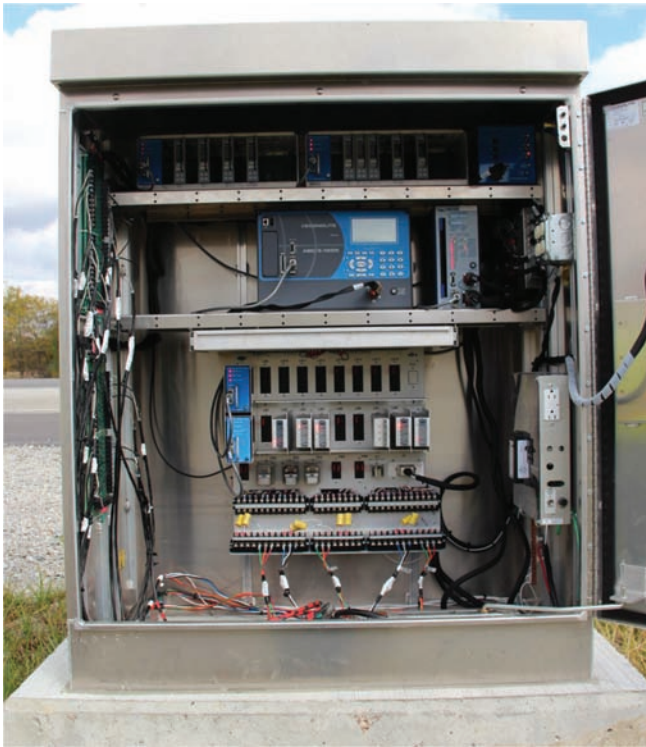


Figure 2.1 NEMA TS/2 Type 1 cabinet.

occupancy measurements aggregated into several-minute-long bins. External, advanced control applications were required to obtain data with more detailed informational content. The emergence of high-resolution signal event data has given a means of developing a more open-platform, information-rich dataset within the controller itself. Table 2.1 shows some controller models capable of high-resolution event data logging as of this writing. Currently, several controller manufacturers have adopted a uniform set of signal event data definitions, and it is anticipated that in the future, many traffic signal controllers will possess, out of the box, an ability to establish a detailed record of intersection operations.

Such event logging functionality should be provided in the newest firmware release for supported hardware, or must be enabled per instructions specific to the manufacturer. For example, NEMA TS-2 type controllers communicate via the NTCIP method, and messages can be delivered using proprietary software to the controller to enable logging capabilities for any supported hardware.

2.4 Detection Technology

The first requirement for signal performance measures is detection. Performance of detection technology can vary among different types, but all types are acceptable. For controllers without any detection but with high-resolution data logging enabled, only phase timings and controller status codes can be monitored for that intersection.

Several types of sensing technologies are used for vehicle detectors. *In-pavement* detectors require sensors

to be embedded in or be underneath the roadway pavement. These include inductive loop detectors, as well as some types of magnetometers. Recently, several compact wireless in-pavement detectors have become available. The other category of detectors is commonly referred to as *above-pavement* or *noninvasive*, which do not require cuts in the pavement. Some types of noninvasive detection include video, radar, and infrared sensors. These sensors generally reside above the lanes of traffic installed on luminaires or signal mast arms.

Detector types have varying conditions under which they are effective and require varying levels of maintenance, and it is beyond the scope of this chapter to compare and contrast their performance. However, it is important to emphasize that the high-resolution data concepts are *independent* of the type of detection available at an intersection, as long as the detection is capable of providing an accurate record of vehicle occupancy and/or count contingent on the application. Moreover, as part of the detector selection and implementation process, agencies can specify acceptable *baselines* for detector sensitivity, responsiveness, and accuracy. For example, the Indiana Department of Transportation has developed criteria for evaluating and accepting the performance of all detectors that are installed for the state (Table 2.2 and Table 2.3).

Although many of the initial studies regarding high-resolution data were carried out at locations where inductive loops were in place, the same data are being collected effectively at many other locations using other types of detection technology. For example, the Utah Department of Transportation (UDOT) makes extensive use of radar detection, while there are several intersections in Elkhart County, Indiana that use video detection. Both of these agencies have made extensive use of performance measures based on high-resolution

TABLE 2.1
High-resolution event data logging-supported controllers.

Vendor	Model	Firmware Version
Econolite	ASC/3	OS version 01.14.03 or higher Application version 12.50 or higher
	Cobalt	Any version
Intelight	All	1.7.0 or higher
Peek	ATC-1000	03.05.0528 or higher
Siemens	M52	3.52 or higher
	M60 ATC	4.52 or higher
Trafficware	970 ATC	76.10 or higher
	980 ATC	
	ATC	
	2070 L	
	2070 LN	
	2070 E	
	2070 EN	
2070 ATC		
McCain	ATC eX	1.7.0.5484 or higher

TABLE 2.2
Parameters for measuring performance per detection zone, Indiana Department of Transportation [50,53].

	Low Performance		Standard Performance	
	During Amber and Red Interval	During Green Interval	During Amber and Red Interval	During Green Interval
Lateral Offset	≤ 3.0'	≤ 3.0'	≤ 3.0'	≤ 3.0'
Activation Position, Upstream Tolerance	≤ 6.0'	≤ 6.0'	≤ 3.0'	≤ 3.0'
Activation Position, Downstream Tolerance	≤ 6.0'	≤ 6.0'	≤ 3.0'	≤ 3.0'
Termination Position, Upstream Tolerance	≤ 6.0'	≤ 6.0'	≤ 3.0'	≤ 3.0'
Termination Position, Downstream Tolerance	≤ 6.0'	≤ 6.0'	≤ 3.0'	≤ 3.0'
Response Time, Typical (85%)	≤ 2 sec	≤ 1 sec	≤ 1 sec	≤ 100 ms
Response Time, Maximum (100%)	≤ 10 sec	≤ 5 sec	≤ 5 sec	≤ 1 sec
False Call Duration	≤ 5 sec	≤ 5 sec	≤ 500 ms	≤ 500 ms

TABLE 2.3
Acceptance criteria for detection missed and false calls, Indiana Department of Transportation [50,53].

	Low Performance		Standard Performance	
	During Amber and Red Interval	During Green Interval	During Amber and Red Interval	During Green Interval
Number of Missed Calls	0 per 24 hours	≤ 10 per 24 hours	0 per 24 hours	≤ 10 per 24 hours
Number of False Calls	≤ 20 per 24 hours		≤ 20 per 24 hours	

data, as these performance measure recommendations are compatible with any form of effective vehicle sensing technology that is able to provide accurate volume and occupancy information.

Besides the sensing technology itself, these are two basic modes of operation that control the types of data that the detector can provide. The current paradigm in which detectors operate requires the provision of what is called a “contact closure,” a simple electrical active or inactive state that is used to inform the signal controller of activity on a detector channel. There are two basic types of information conveyed by this signal, as illustrated in Figure 2.2. The top of the figure shows a time-space diagram in which the trajectories of a series of vehicles moving through a detection zone are represented. Beneath this, a hypothetical graph of the actual sensor response is shown, and the detector output for presence and pulse (count) detection are shown.

- *Presence* data record the detector occupancy whenever there is a sensor response to a vehicle within the detection zone. The presence state is typically active when there is any portion of a vehicle in the detection zone. For example, in Figure 2.2, the first presence output ON state is active from the entry time of the front end of the first vehicle to the exit time of the rear end of the last vehicle. This measurement is used to drive phase actuation, and can be helpful for measuring phase utilization.
- *Pulse* or *count* data consist of a short (~0.1 second or shorter duration) active state that is produced by the detector when the presence state becomes active. For long detection zones, the detector must apply some additional analysis on the actual sensor response to estimate the arrival of new vehicles. This is not required for shorter detection zones when multiple vehicles do not overlap the detection zone.

For most detector types, the detector presence is the “default” output, while count data require special processing of the data for long detection zones. For short detection zones, presence and count data have essentially the same information content.

In general, the detector *presence* is needed on all detector channels to obtain useful information. Detector *count* is desirable for long stop bar detection zones where there is expected to be an overlap of vehicles in the same detection zone. However, it is still possible in the latter case to develop useful measures of phase utilization, even if count data are not available. While it is still *desirable* to have vehicle counts, if possible, it is not *required*.

The choice between whether to provide presence and/or count data is sometimes a matter of cost, because of the need to provide a separate count output for stop bar detection. There may be some marginal increase in cost when including count detection, since the number of detector channels increases, meaning that more detector racks might be needed in the cabinet. These issues may need to factor into intersection detection designs.

It should be noted that some detection types can expand more flexibly than others, particularly some noninvasive types where the detection zones are externally defined. For example, UDOT uses a radar detection system whose cabinet interface can expand to 64 detection channels with no additional cost, because all of the data are passed through the SDLC bus. The detection system enables the user to freely create and change detection zones for both presence and count. When UDOT began deploying performance measures across their system, it was relatively simple to implement lane-by-lane detection.

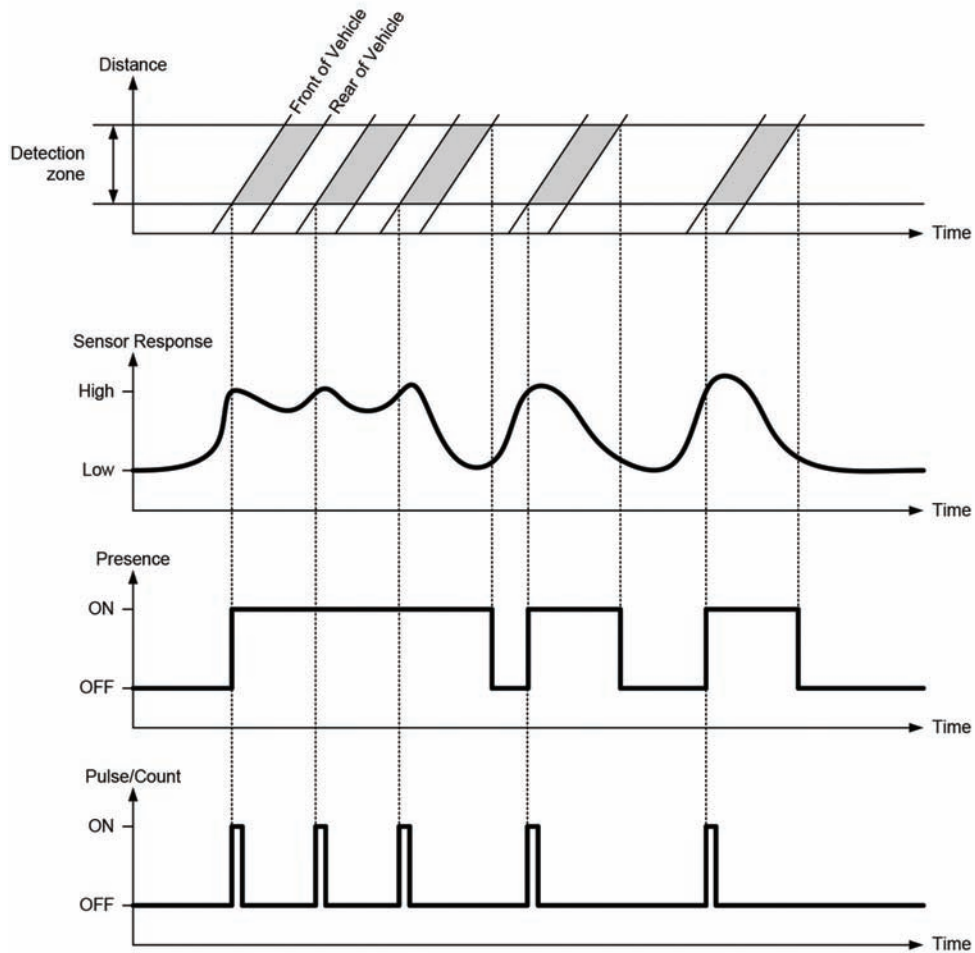


Figure 2.2 Conceptual comparison of outputs from various detector modes.

2.5 Detector Positioning

Detection is typically implemented at an intersection to enable actuated signal control and performance measurement. If existing detector layout designs do not satisfy the requirements of updated agency performance objectives, intersection design manuals could be amended with more detailed layout recommendations. Moreover, design manuals that do provide specifics on detector locations may be based more on conformance to prior practice rather than fulfillment of a performance expectation. Consequently, many existing intersections have an incomplete detection plan because of such a legacy. The authors would encourage engineers to consider developing a standard design based on a higher expectation of performance. This may increase the detection quality arising from future projects, even if some interim projects may have to accept relatively low standards due to cost constraints.

Several options exist with regard to the physical space occupied by detection zones. The detection zone length (within each lane), width (across lanes), and

distance from the stop bar affect the information that can be obtained from the detector, as well as the types of use cases that a detector can satisfy.

Long detection zones are generally used close to the stop bar to accommodate variation in the position where a single vehicle comes to a stop. This pertains to phase actuation. Usually, the same detection zone used for actuation is also used to terminate the phase by detecting gaps between vehicles. *Short detection zones* can also be created, such as for measuring a count of vehicles on the approach to an intersection at an upstream location. There is no need to provide coverage over a long area leading up to the intersection if the objective does not include capturing stopped vehicles.

The *distance from the stop bar* determines the possible use cases for the detector. Several options in signal timing become available with advance detection, including dilemma zone protection, and it also becomes possible to evaluate and optimize coordinated signal timing when the arrival patterns can be measured. On the other hand, *stop bar detection* is useful for measuring phase utilization, in addition to providing a means of measuring gaps for

phase termination. For these reasons, it is recommended that an agency consider installing both stop bar and advance detection on all intersection approaches.

Figure 2.3 shows an example of recommended and desirable detection zone configurations at a four-legged intersection for a crossing mainline and side street. The zones are marked depending on whether they provide critical data for monitoring intersection performance. Advance detection on the mainline movement through lanes is recommended for dilemma zone protection and calculating progression and queue performance measures (A). To capture phase utilization on all lanes, stop bar detection zones are recommended for all mainline and side street movement lanes (B). At intersections with right-turn-on-red (RTOR) allowed, the detection for the right-turn-only lanes can often be omitted, although on heavier movements separate detection may still be desired (C). For crossing-arterial situations, advance detection on the side street is desirable (D).

Some alternative *lateral range configurations* exist in the field. Often, multiple lanes are “tied together,” which ostensibly makes the configuration easier to manage, as only one channel needs to be configured in the controller

rather than multiple ones. In Figure 2.4, a four-legged intersection with stop bar detection zones spanning multiple lanes is presented. Such a setup will have less fidelity at the stop bar than if detection zones are set up in each lane separately. A detection zone that is both long and wide will not provide very detailed information about the approach, other than the overall utilization. It cannot, for example, differentiate vehicle arrivals between lanes. In general, “lane-by-lane” detection is more efficient for driving phases [54], increases robustness for performance measurement, and is recommended when possible. This is especially true in cases where there are substantial differences in vehicle behavior between lanes, such as approaches with a high percentage of right-turning vehicles.

2.6 Detection for Actuated Signal Control

The most common use of detection is *phase actuation*. Phase actuation is the adjustment of signal timing in response to traffic demand as it arrives at an intersection. This is relevant not uniquely to vehicles, but also to bicycles, pedestrians, and transit vehicles, all of which can provide input through detection.

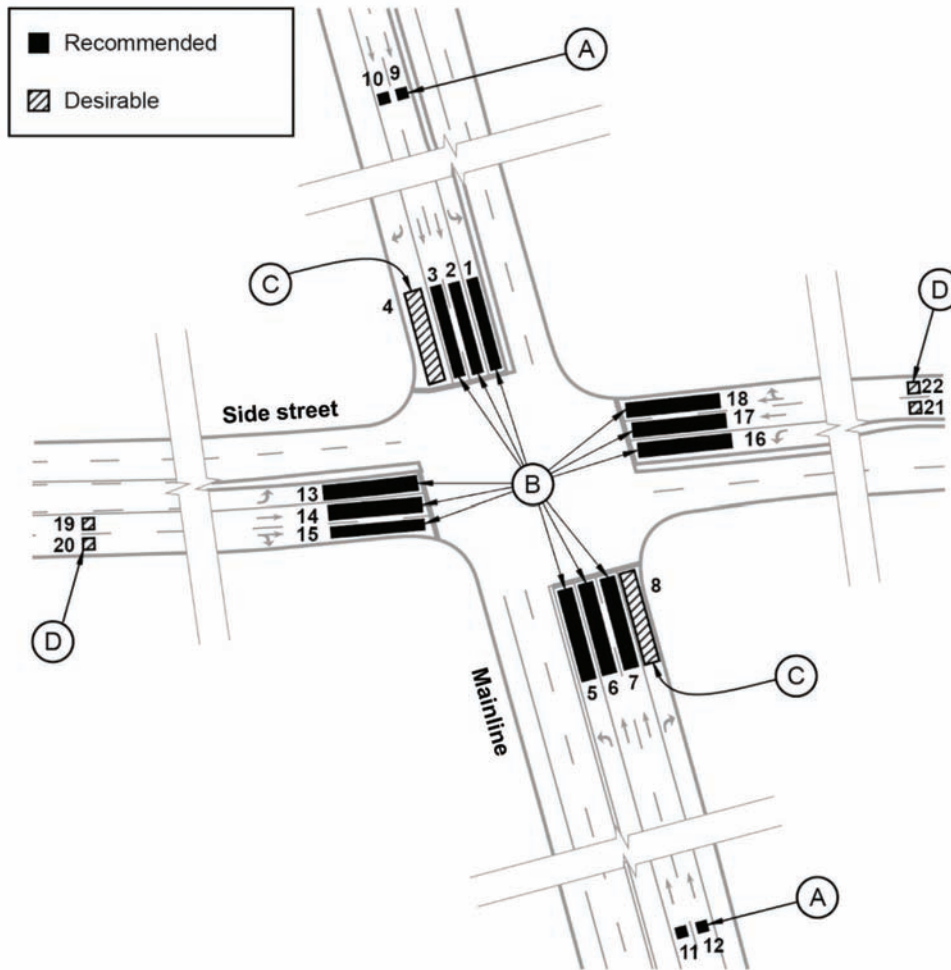


Figure 2.3 Example of lane-by-lane configuration for vehicular detection.

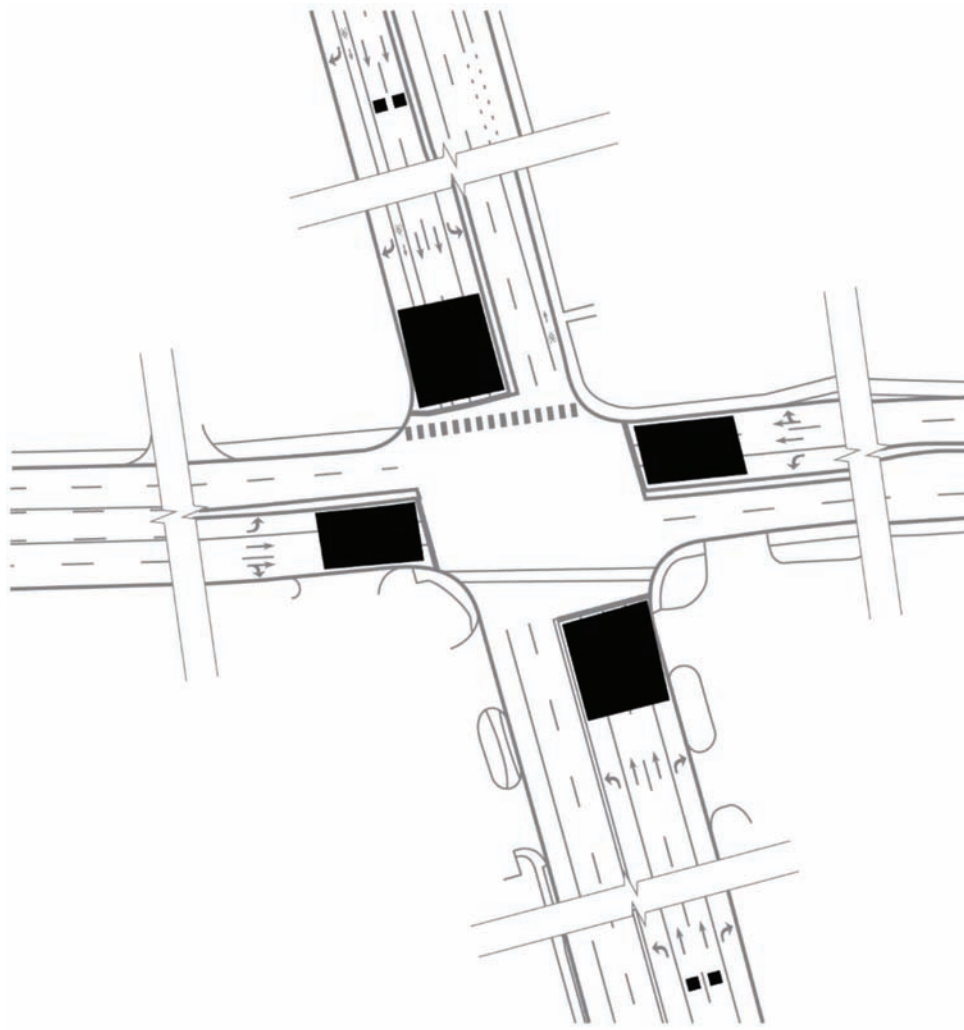


Figure 2.4 Example detection layout with stop bar zones that are “tied together.” Such a configuration will still support phase utilization performance measures, but it will have less detailed information than lane-by-lane detection.

In phase-based signal control, actuation typically involves a call for service on a particular signal phase, in response to the presence of a vehicle.

Actuation typically involves both the initiation and termination of a phase. Detectors are configured in the controller to register calls on associated phases. The same detectors can also be configured to terminate those phases by measuring gaps in the traffic flow over those detectors. Any vehicle sensed by a detector can be set to “extend” a passage timer for the associated phase. When a gap between vehicles renders all of the detectors associated with a phase inactive, the passage timer will count down to zero, thereby terminating the associated phase. Note that this also behaves in tandem with other rules governing the minimum and maximum green time and coordinated split time set in the controller. The way actuated control is implemented varies according to the spatial configuration of the detection zones.

- *Stop bar detection* detects vehicles that are stopped at an intersection, and gaps in traffic flow currently moving through the lanes at the intersection.
- *Advance (setback) detection* detects vehicles en route to the intersection and gaps in traffic flow as it approaches the intersection at a certain distance.

Historically, stop bar detection has been used for “minor” movements at an intersection. The major movements (typically the mainline through movements) are often provided with advance detection, although in some cases, no detection is provided at all; in this case, the green simply reverts to the mainline once the side street has served all vehicles. A movement that lacks any detection at all must be timed without reference to traffic demand.

Another application related to phase actuation is *dilemma zone protection*, which is related to phase termination using advance detection. A dilemma zone exists when the driver is forced to make a decision whether to stop or to proceed following a signal

indication change from green to yellow. Although it is possible to avoid dilemma zones by selecting appropriate clearance times, it is also possible to reduce the number of times that drivers are forced to make stop-or-go decisions by extending phases to accommodate approaching traffic.

For this purpose, the advance detector is located upstream of the stop bar at the stopping sight distance for the approach speed, which is usually the posted speed limit. For example, at 55 mph, a safe stopping sight distance on level, dry pavement is 408 ft, so the detector is located at about that distance, which corresponds to a 5-second travel time to the stop bar.

For detecting arrivals for performance measure purposes, it is desirable to locate the advance detector beyond the extent of queuing. The detector distance should be upstream of where vehicles are likely to stop because of the initial queues during red, but should not be located so far upstream that the measurements become inaccurate. About a 400 ft setback distance is reasonable for intersections with light to medium traffic; the distance may need to be increased for more congested intersections. Such changes are often easier to make with noninvasive detection types. In situations where intersections are closely spaced, it may make more sense to use detection at an upstream intersection to measure the arrivals at the downstream intersection.

2.7 Detector Mapping for Performance Measures

The most critical requirement for interpreting the high-resolution signal event data logged in a controller is the detector mapping information. Since any detector event must be understood as a vehicle arriving at *some* right-of-way and distance at the intersection, and any phase event must be understood as an interval of time elapsed for *some* movement, the detector mapping process is the critical step in guiding how signal output states and vehicle arrivals are to be interpreted relative to each other. Without the provision of such information, it would be impossible to generate most of the signal performance measures from the detector and phase data independently.

Associating output phasing and overlap enumerations with movement indications typically requires intersection plan documents, load switch wiring diagrams, or a field visit. Detector channel-to-output assignments may also be documented in a similar fashion, or labeled on equipment within the cabinet. Figure 2.5 depicts a typical detector rack configuration with four two-channel loop detector amplifier cards installed. The labels under each card signify the approach direction, lane, and phase associated with the incoming channels. At the controller, the same detector channels are assigned to a phase under configured timing plans. Controller parameters at the very least tell which detectors are assigned to which outputs. However, details such as whether a detector has a distance set back or lane-by-lane attribution may be more obscured or often impossible to distinguish. For instance, if a lane contains both stop bar and advance



Figure 2.5 Detector bus interface unit (BIU) with approach, phase, and lane labeled.

detection, more specific configuration details such as each detector's extension method need to be explored. Alternatively, for multiple lanes in the same movement, this method of sleuthing will not likely distinguish each individual lane.

For larger agencies that perform much of the systems engineering and asset management in-house, the detector mapping information is typically recorded in the plans or in technical documentation either electronically or on hardcopy format during construction and system installation. Often, engineering plans providing such information may be kept in the field in signal cabinets. Agencies with more sophisticated detection systems or central systems can have such information stored in the detection equipment, software, or database, depending on vendor-specific product features. This may also be the case for smaller agencies that hire consultants for any signal- and detection-related work. Detector mapping information may only be recorded temporarily for system installation purposes and may not be formally documented. Again, in this case, much of the information resides in the physical equipment or software and is not made readily available to agencies, simply owing to its lack of utility to the end-user.

Table 2.4 shows an example detector mapping table for a hypothetical eight-phase intersection with protected-permitted left turns on the mainline and side street. Figure 2.6 is a map of the intersection, with the movement outputs overlaid in black and the detection zones and channel numbers overlaid in red, with count channels labeled with an asterisk (*). Each lane on both the mainline and side street has a 36 ft detection zone at the stop bar, with a separate 6 ft zone set back at 45 ft for counting. Because the mainline is a coordinated arterial that runs in the north-south direction with a speed limit of 50 mph, advance detection zones are configured at 365 ft back from the stop bar for the northbound and southbound through movements for individual lanes. This configuration is acceptable for dilemma zone protection and progression performance measurement.

This example is meant to be purely illustrative and is a common INDOT configuration. Other agencies will likely have very different configurations, yet will still very likely support performance measures. For example, a typical UDOT configuration would include 65 ft stop

TABLE 2.4
Example detector mapping and descriptive information for an eight-phase intersection.

Approach	Movement	Lane	Channel Number	Channel Type	Phase/Overlap	Length(ft)	Stop Bar Setback (ft)
Northbound	Left turn	L	35	Presence	$\phi 5$	36	0
			14	Presence	$\phi 5$	6	45
			16	Count	$\phi 5$	6	45
	Through	A	17	Presence	$\phi 2$	36	0
			19	Count	$\phi 2$	6	45
			37	Presence	$\phi 2$	6	365
		B	18	Presence	$\phi 2$	36	0
			20	Count	$\phi 2$	6	45
			38	Presence	$\phi 2$	6	365
	Right turn	R	36	Presence	$\phi 2$ /Overlap A	36	0
			21	Presence	$\phi 2$ /Overlap A	6	45
			23	Count	$\phi 2$ /Overlap A	6	45
Southbound	Left turn	L	29	Presence	$\phi 1$	36	0
			1	Presence	$\phi 1$	6	45
			3	Count	$\phi 1$	6	45
	Through	A	2	Presence	$\phi 6$	36	0
			4	Count	$\phi 6$	6	45
			30	Presence	$\phi 6$	6	365
		B	5	Presence	$\phi 6$	36	0
			7	Count	$\phi 6$	6	45
			32	Presence	$\phi 6$	6	365
	Right turn	R	6	Presence	$\phi 6$ /Overlap B	36	0
			8	Count	$\phi 6$ /Overlap B	6	45
	Eastbound	Left turn	L	40	Presence	$\phi 7$	36
22				Presence	$\phi 7$	6	45
24				Count	$\phi 7$	6	45
Through		A	25	Presence	$\phi 4$	36	0
			27	Count	$\phi 4$	6	45
Right turn		R	42	Presence	$\phi 4$ /Overlap C	36	0
			26	Count	$\phi 4$ /Overlap C	6	45
			28	Presence	$\phi 4$ /Overlap C	6	45
Westbound		Left turn	L	33	Presence	$\phi 3$	36
	9			Presence	$\phi 3$	6	45
	11			Count	$\phi 3$	6	45
	Through	A	10	Presence	$\phi 8$	36	0
			12	Count	$\phi 8$	6	45
	Right turn	R	35	Presence	$\phi 8$ /Overlap D	36	0
			13	Presence	$\phi 8$ /Overlap D	6	45
			15	Count	$\phi 8$ /Overlap D	6	45

bar lane-by-lane presence zones with separate count detectors and separate advance detector channels. Also, INDOT mainly uses inductive loops, whereas UDOT mainly uses radar detection. Both types of detection and both detector zone configurations will support performance measures. However, complete documentation of the detector configuration is the key step in translating the raw data into performance measures. The geometry of the detection zones and their lane and channel assignments are the key pieces of information.

2.8 Supported Performance Measures

The suite of performance measures that can be implemented at an intersection ultimately depends on the physical space occupied by the detection zones, how far back the zones are configured from the stop bar, and the manner in which the detector zones transmit information to the controller. At intersections operating with at least some amount of actuated control, minor (noncoordinated) phases are typically instrumented with detection zones at

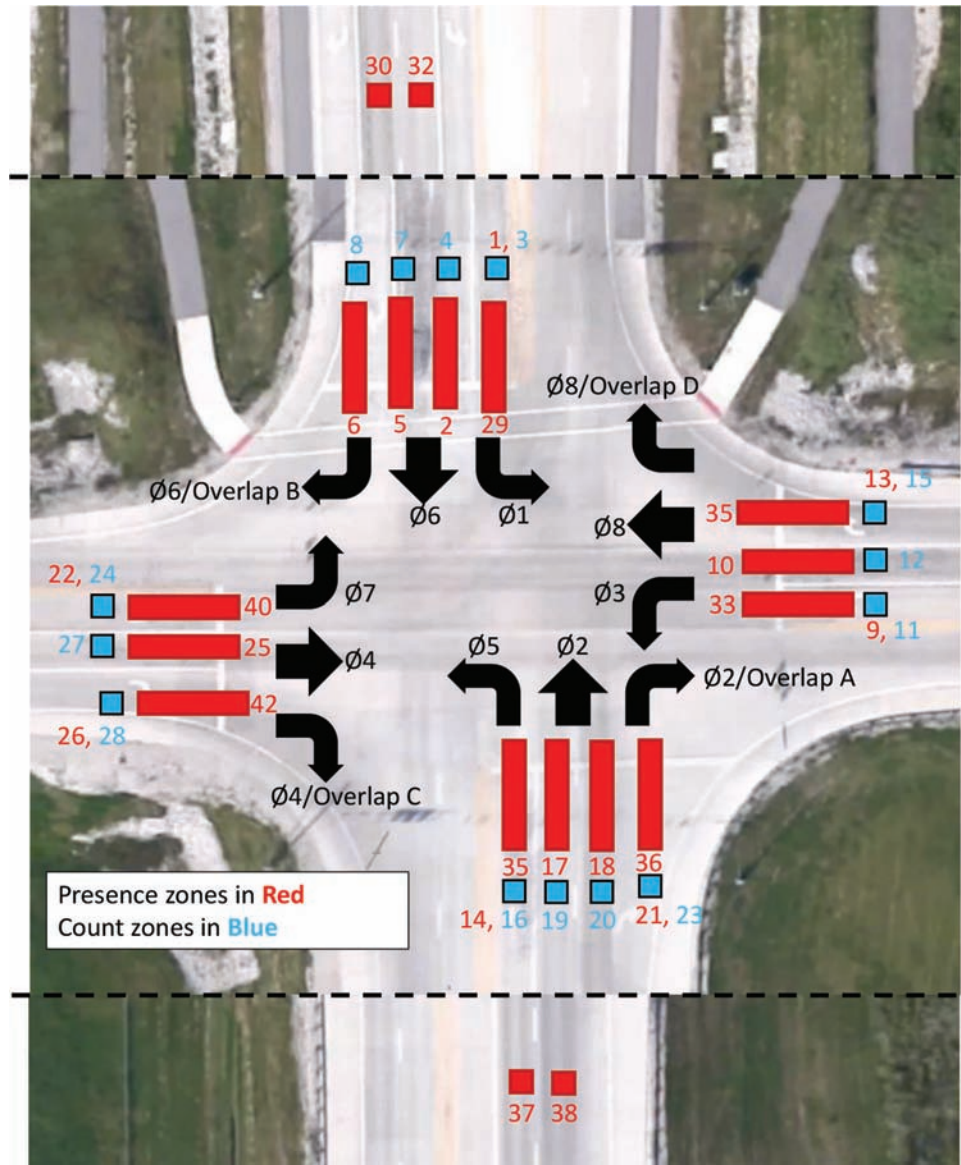


Figure 2.6 Example intersection with assigned phasing and detection zones.

the stop bar so as to allow the phase to be omitted when no vehicles are present, and also allow for the extension of a phase when vehicles are present. Advance detection zones allows for dilemma zone protection for high-speed approaches typically found on suburban arterials. In addition, some implementations include one additional count channel per lane to record volumes.

Figure 2.7 illustrates four common detector zone configurations at an eight-phase signalized intersection for a crossing mainline and side street:

- *Configuration (a)* has stop bar detection zones in the presence mode configured only for the minor movements with no mainline through detection. This configuration assumes the controller will default back to serving the mainline phases when all other phases have been served.
- *Configuration (b)* has stop bar presence detection zones for all major and minor movements, which would allow for

major phases to be omitted when no vehicles are present.

- *Configuration (c)* has stop bar presence zones on the minor movements only and advance presence detection zones for the major movements for dilemma zone protection.
- *Configuration (d)* has stop bar presence zones all movements and mainline advance presence zones.

Figure 2.8 shows four similar detector zone layouts, but with the addition of count detection at each zone. Figure 2.9 is a matrix of attainable performance measures for each of the configurations presented in Figure 2.7. Generally, the more robust the detector layout is configured, the greater the suite of performance measures can be attained.

- **Green/yellow/red duration** is a measure of the amount of green, yellow, and red time served for a particular phase or overlap. No detection setup or mapping information is required.

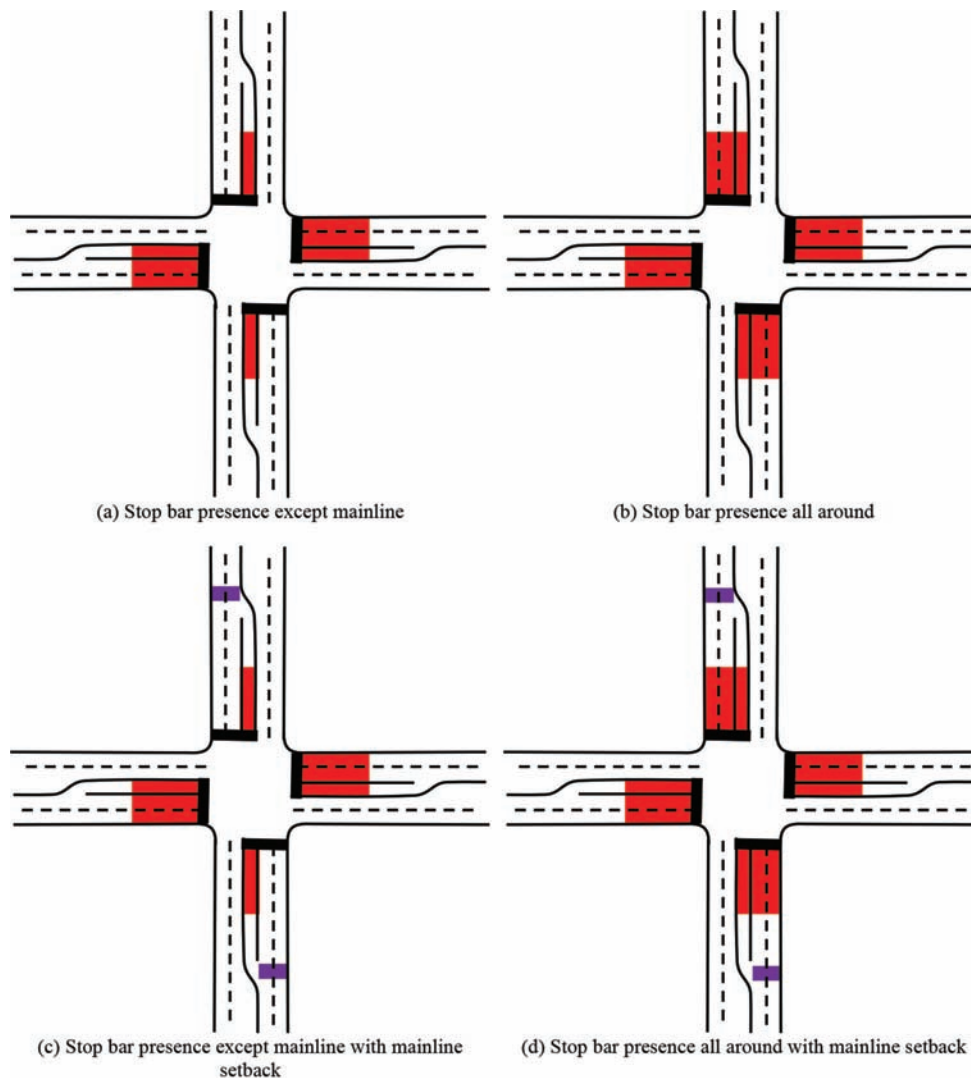


Figure 2.7 Selective alternative detector configurations without count detection.

- **Phase termination type** signifies whether a phase or overlap has gapped out, maxed out, or forced off. This is typically driven by arrivals in an actuated control system. Although detector actuation determines how a phase is terminated, the phase termination type acts as a surrogate for capacity utilization as no detector mapping is necessary.
- **Volume per time period** is a count of vehicles interpreted from the number of activations from a setback short length detector or a dedicated count channel.
- **Vehicles-per-hour (vph)** is the rate of vehicle arrivals for a one-hour period. This value is typically converted from a volume per time period.
- **Volume/capacity ratio (v/c)** is the number of vehicle arrivals during a time period divided by the theoretical capacity of vehicles approximated from the green time (less any startup loss) served in the same period.
- **Maximum vehicle delay** is the greatest number of seconds between a detector activation time at the stop bar during red, to either the beginning of green or detector deactivation time for RTOR vehicles [55].
- **Green/red occupancy ratio (GORROR)** is a measure of capacity utilization that uses the amounts of detector occupancy during green and red portions of a phase, respectively.
- **Split failure** is an indication of overcapacity for a lane, phase, or movement. A combination of a green phase exceeding a certain occupancy ratio followed by a red phase exceeding a certain occupancy ratio triggers the indication. Typically, a threshold occupancy of 80% for the entire duration of the green phase and 80% for the first 5 seconds of the red phase is used [21].
- **Queue estimation** is a measure of the queue length using two possible methods: vehicle arrivals and estimated discharge rates based on green time served (input-output method), or using detector gap durations and counts (shockwave estimation method) [56].
- **Oversaturation severity index** is the ratio of the unusable green time to the total available green time in a cycle [57].
- **Time to service** is a measure of the first detection to the beginning of green.
- **Percent on green (POG)** is the percent of total vehicles arriving during a given cycle in the green.
- **Purdue Coordination Diagram (PCD)** is a visualization of the quality of progression of a movement over a period of time.

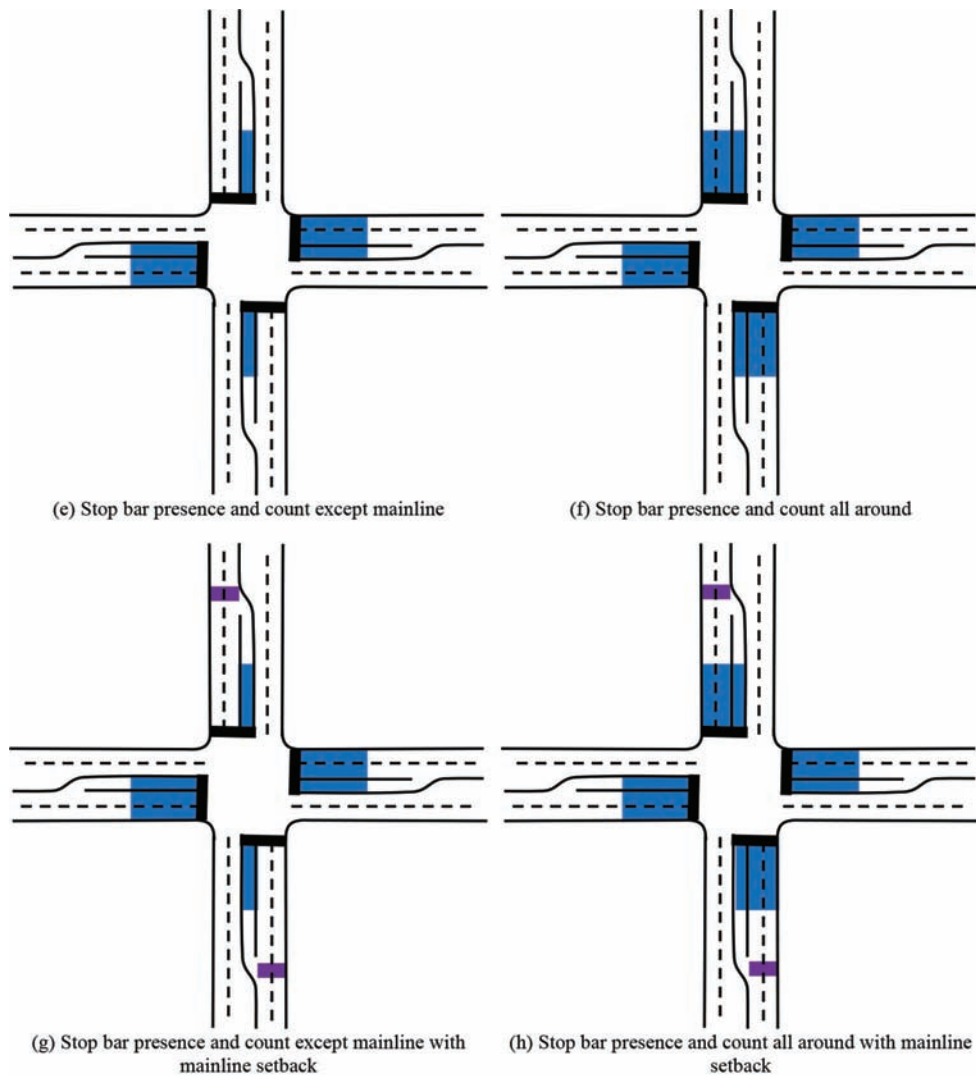


Figure 2.8 Selective alternative detector configurations with count detection.

- **Platoon ratio** is the percent of vehicles arriving on green, adjusted by the green time proportion per cycle (g/C).
- **Cyclic flow profile** is a view that combines the distribution vehicle arrivals and the probability of green, aggregated over a set of fixed-length cycles.
- **Time-space diagram** is a visualization of estimated vehicle trajectories traversing a distance over time. Locations and phase status of signalized intersections are identified in the diagram.
- **Red light running** measures vehicles entering and exiting a stop bar detection zone after the beginning of red.

2.9 Travel Time Data

In addition to performance measures derived from high-resolution signal event data logged by controllers, travel time data provide an additional set of outcome assessment metrics for corridor or system-level performance. Typically, these datasets are acquired independently of controller-logged data through the deployment of agency or consultant-driven probe vehicles, field instrumentation of

vehicle re-identification devices such as Bluetooth and Wi-Fi readers, or downloading of third-party commercial traffic providers in real-time or as a delayed archive.

Figure 2.10 is a conceptual diagram illustrating how a signalized corridor can be instrumented with nonpermanent Bluetooth monitoring stations (BMS) for measuring travel times from Point A to Point B. At each end of the corridor, the BMS acquires and timestamps Media Access Control (MAC) signatures as Bluetooth-enabled vehicles approach each station [58]. As the vehicles enter and exit the corridor, travel times can be computed from the difference between the entry and exit timestamp of distinct MAC address signatures.

Alternatively, Figure 2.11 illustrates how speed data from a third-party commercial vendor can also be used to compute travel times. For each of the three segments in the figure, an average speed is given for a moment in time from which a travel time can be computed. By adding the travel times of the three segments together, an instantaneous approximation of corridor travel time can be computed.

Task Group	Performance Measure(s)	Detector Configuration							
		a	b	c	d	e	f	g	h
Monitor phase events	Green/yellow/red duration	✓	✓	✓	✓	✓	✓	✓	✓
	Phase termination type	✓	✓	✓	✓	✓	✓	✓	✓
Counting	Volume per time period			Δ	Δ	Δ	✓	Δ	✓
	Vehicles-per-hour (vph)			Δ	Δ	Δ	✓	Δ	✓
	Volume/capacity ratio (v/c)			Δ	Δ	Δ	✓	Δ	✓
Monitor capacity	Time to service	✓	Δ	✓	✓	✓	✓	✓	✓
	Maximum vehicle delay	Δ	✓	Δ	✓	Δ	✓	Δ	✓
	Green/red occupancy ratio (GORROR)	Δ	✓	Δ	✓	Δ	✓	Δ	✓
	Split failure	Δ	✓	Δ	✓	Δ	✓	Δ	✓
	Queue estimation			Δ	Δ	Δ	✓	✓	✓
	Oversaturation severity index			Δ	Δ			Δ	Δ
Monitor progression	Percent on green (POG)			Δ	Δ			Δ	Δ
	Purdue Coordination Diagram (PCD)			Δ	Δ			Δ	Δ
	Platoon ratio			Δ	Δ			Δ	Δ
	Cyclic flow profile			Δ	Δ			Δ	Δ
	Time-space diagram			Δ	Δ			Δ	Δ
Monitor safety	Red light running	Δ	✓	Δ	✓	Δ	✓	Δ	✓

NOTE: ✓ all movements; Δ select movements; ■ unsupported.

Figure 2.9 Detector configuration requirements matrix for performance measures.

The aforementioned datasets for assessing system-level travel times will incur additional costs to an agency that has neither existing infrastructure in place for collecting the data nor existing data contracts with commercial traffic data providers for downloading the data. These datasets must be procured and managed as separate items under the umbrella of the greater signal system business process. However, in some cases, infrastructure may already be in place for managing controller event data, as these resources may be leveraged for the transmission, storage, and processing of the travel time data to reduce costs. Some of these modules include power sources, wiring and networking hardware, data connections, data processing software, databases, reporting software, and interactive user interfaces.

2.10 Communication

The primary use of a communication network in traffic signal performance measurement is to transmit data securely and within a tolerable latency threshold between the signal controller and the traffic signal management center or offices. Depending on the geographic expanse, density, and stage of infrastructure investment pertaining to each agency, there can be numerous challenges for enabling data transmission between each cabinet to the

back-office. As such, there is not a “one size fits all” design for communications infrastructure, but, rather, each scenario must be evaluated based on budget, relationships with telecommunication providers, and feasibility of leveraging existing facilities.

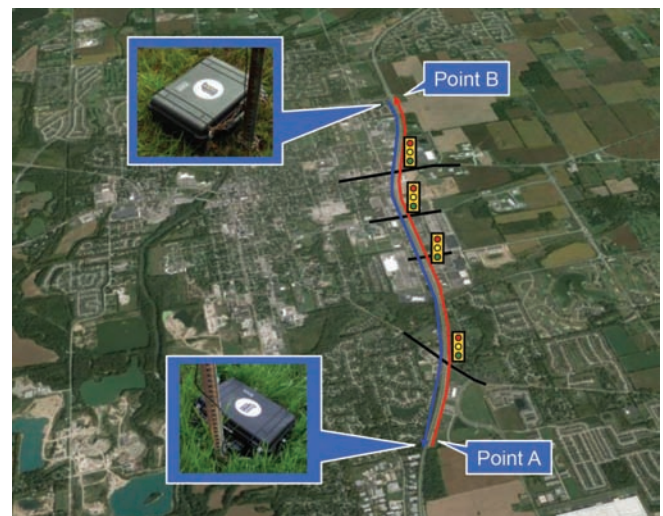


Figure 2.10 Vehicle re-identification using nonpermanent Bluetooth sensing technology for travel time measurement.

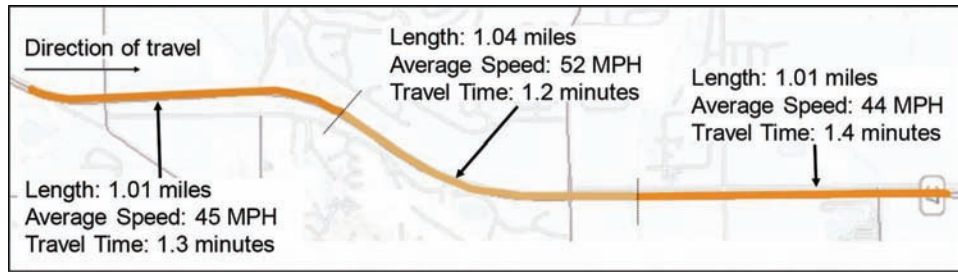


Figure 2.11 Travel time computation using crowd-sourced probe data segmentation.

Communications infrastructure comprises dial-up, fiber optics, Digital Subscriber Lines (DSL), cellular, and/or wireless transmitters connecting a signal cabinet to a remote central system or, on older systems, to other cabinets that contain a main or “master” controller. Historically, communications infrastructure to signal cabinets was necessary for real-time monitoring, remote configuration, and clock synchronization of controllers in a system. Current performance measure communication requirements in many cases can still leverage the existing infrastructure for high-resolution event data retrieval. Using the communication network, the event data can be transported to the regional traffic management center back-office for storage and analysis. Figure 2.12 provides an example

of a hybrid communications network that uses fiber optics along a local arterial with broadband gigabit radio connecting the local network to a fiber backbone. The high-resolution data flow from each controller through an internal secured network to servers at the traffic management center, where the data are processed, stored on a database, analyzed, and visualized. For systems that do not communicate over a dedicated network, Virtual Private Networking (VPN) should be implemented for the secure transmission of data.

For locations without adequate communications infrastructure, localized single-board embedded computers can cache high-resolution signal data. Recent technology has improved the affordability and computing power of

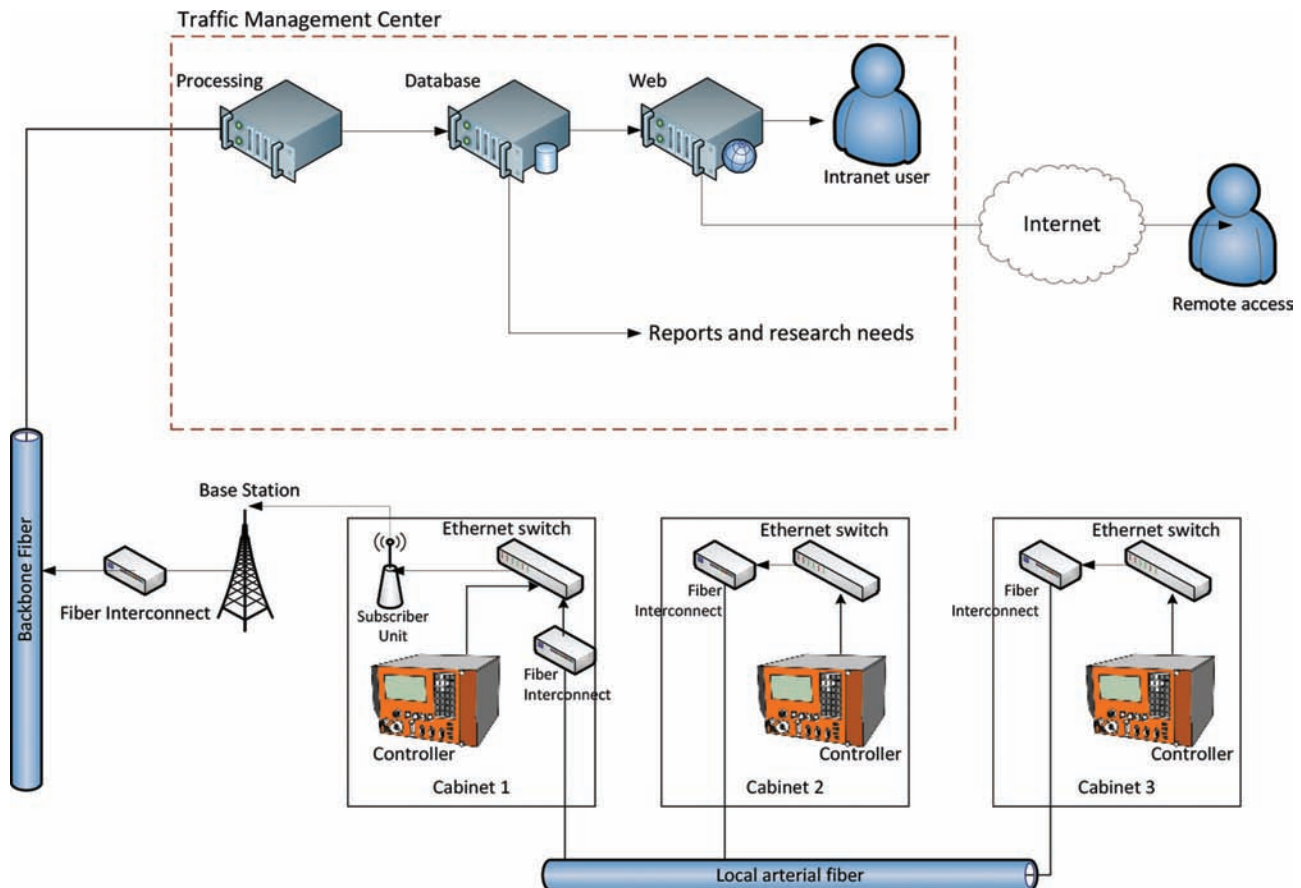


Figure 2.12 Example communication infrastructure using fiber optics and broadband wireless communications.

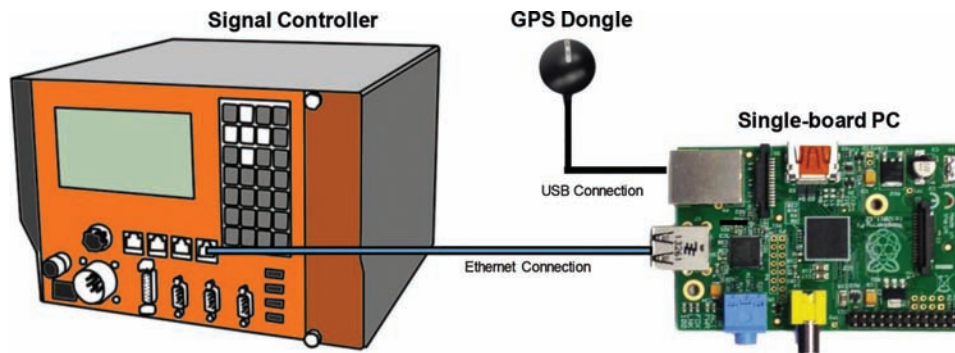


Figure 2.13 Scheme for high-resolution data collection and clock synchronization using an embedded computer.

small-sized, single-board personal computers. These computers typically have integrated onboard graphics, sound, and networking components and store all data on interchangeable, nonvolatile flash memory cards. With current memory card capacities reaching upward of 128 gigabytes, multiple decades of data from individual controllers can be stored. Through field visits, the stored data can be retrieved manually and brought back to the data warehouse. Moreover, these computers can add on additional devices through the Universal Serial Bus (USB) adaptor interface to extend functionality, much like a traditional PC or laptop. Such computing devices used in conjunction with Global Positioning System (GPS) adaptors have enabled time synchronization services that typically require communication infrastructure to operate without any extensive cabinet interconnect. It is important to note that such setups require a nontrivial amount of software development for enabling these functionalities. The components typically required for a single-board PC setup are illustrated in Figure 2.13.

2.11 Data System

The data system required for signal performance measurement is composed of software and hardware components that download, normalize, archive, interpret, and display the signal data logged by the controller. Each step has unique modules that must be identified, procured, and maintained by an agency business process. Moreover, the requirements of each module will vary depending on the number of intersections in the system. An overview of the requirements is outlined in Table 2.5:

- **Downloading** is the transfer of raw, unprocessed signal event data to the data center. Typically, this process runs at set time intervals depending on agency operating objectives.
- **Normalization** is the conversion of logged controller event data into a uniform and consistent format. This is necessary owing to the number of different controller logging formats that are available from various manufacturers. Software should be provided by the manufacturer to translate proprietary formats. Additionally,

TABLE 2.5
Data process requirements.

Data Process	Software Required		Hardware Required	
	Small system (about 1–20 signals, 1 month of data)	Large system (>20 signals, >1 month of data)	Small system (about 1–20 signals, 1 month of data)	Large system (>20 signals, >1 month of data)
Download	Single open connection	Simultaneous open connections	Desktop workstation with additional disk space	Server (network and disk critical)
Normalization	Single-threaded program	Multithreaded program with parallel processing	Desktop workstation	Server (CPU and RAM critical)
Archival	Standard database CSV MS Access Indexing	Enterprise database Indexing Table partitioning Compression	Desktop workstation with additional disk space	Server (disk critical)
Interpretation	SQL queries MS Excel	Stored procedures SQL functions Data structures Dedicated business logic modules	Desktop workstation	Server (CPU and RAM critical)
Display	MS Excel Web applications Other OS or mobile platforms	Web applications Other OS or mobile platforms	Desktop workstation	Server (CPU and RAM critical)

further steps may be taken during normalization to reduce and aggregate the data for computational efficiency.

- **Archival** is the storing of normalized data in a database. Raw, unprocessed data may also be archived for backup or auditing purposes.
- **Interpretation** is the transformation of the normalized event data into actionable information using detector mappings, an event code dictionary, traffic algorithms and heuristics, and business logic.
- **Display** is the visualization of normalized data in a user interface, typically involving charts, graphs, tables, maps, and reports. Displayed data may be customized interactively through user input features.

For small systems with about 20 signals or fewer, or if less than one month of data are archived, less robust software and hardware requirements are required. For the data download and normalization process, typically a single connection and program thread for downloading and normalizing event data should suffice. This will allow for data transfer between the field to the data center archives at a near real-time frequency (e.g., twice an hour). For data interpretation and display, off-the-shelf programs such as MS Excel would suffice to visualize the data. More customizable data visualization software is desirable, but is not necessary. Because not a lot of data will be processed, a dedicated high-end workstation with extra disk capacity would accommodate the computing needs.

For large systems with more than about 20 signals, or if many months or years of data are archived, enterprise software and hardware will be necessary. For the data download process, data retrieval software capable of downloading from multiple controllers simultaneously will be required. At the normalization and archival stages, parallel computations with multithreaded software are necessary to convert and store the data within a reasonable time. In addition, enterprise-level database software is required for handling large volumes of data efficiently. Features such as table partitioning and compression will greatly improve query response times for large datasets. Moreover, substantial development of stored procedures, SQL functions, and data structures will be necessary for applying the performance measures business logic to the normalized data in the interpretation stage. For displaying the data, a number of platforms can be used, ranging from websites to operating system and mobile applications.

The hardware required for large systems will typically be dedicated, enterprise-level servers composed of one or more physical chassis. In some cases, one chassis with high-performing CPUs, disks, and RAM will be sufficient to accommodate all of the data processes, depending on the combination of computing power and system size. If the server performance is limited or if the system is substantially large, there will be significant benefits to using multiple chassis to expand CPU cores, disk arrays, and RAM. For the downloading stage, network and disk-bound operations will be dominant in order to transfer raw event data from multiple controllers to the server simultaneously. During the normalization stage, many processes will kick off at once,

requiring intensive CPU cycles and RAM to convert the data in parallel. For the archival stage, disk performance will be most critical. For example, data input-output (I/O) speeds will be increased by using disks with high revolutions-per-minute, solid-state drives, high-bandwidth host bus adapters (HBA) for connecting the drives to the bus, and an efficiently partitioned RAID array. Performance will also be improved by separating database logging files and data storage files on separate disk arrays to make use of parallel I/O operations. Finally, for displaying the performance measures, fast, multicore CPUs and large amounts of RAM are required for generating charts and graphics to multiple users simultaneously.

Cloud computing is an alternative solution to the in-house deployment of data system software and hardware [40]. By using a distributed system located and operated off-site by a third-party commercial cloud provider, substantial equipment and software acquisition and maintenance costs, as well as risks, can be reduced. However, for agencies with systems that are either very small or very large, the data system's cost per intersection may be high. In addition, the agency will have less control over the system in terms of privacy and network data security.

2.12 Data Scalability

To increase the system scalability for the display of performance measures, data reduction and aggregation techniques can be employed at the normalization and archival stages of the data process. Typically, normalized data consist of stateless events such as a “detector on” indicated by a timestamp and a detector channel identifier of that event. In a noncompressed database, this event record requires a minimum of 10 bytes per event, with potentially up to 100% storage overhead to index the data for performance. To measure detector occupancy with this event alone would not be possible, as another event indicating a “detector off” is needed from which the duration is taken. In such a scenario, there is a nontrivial amount of redundancy for the two events. Interval aggregation would combine these pairs into phase green, phase yellow, phase red, and detection interval records by storing both “on” and “off” times as a single record. This extra process, typically performed during normalization and archival, eliminates the need for specifying the event type and phase or detector identifier in separate records and will reduce the storage requirements and computation time of the dataset.

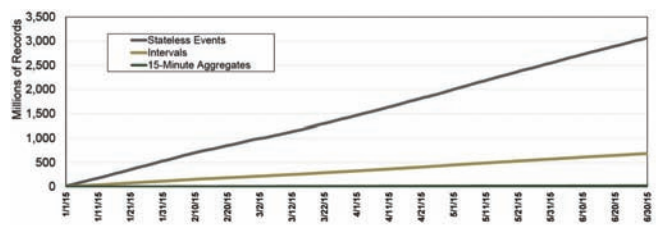


Figure 2.14 Example of data record accumulation by record type over six months.

The data can be further aggregated by 15-minute binning performance measure results. Producing performance measures typically involves some amount of computational time during the interpretation and display stages. By employing a 15-minute binning process, this helps move the computational effort earlier to the normalization and archival stages. In database storage, each 15-minute record can be referenced by the starting bin time and a unique phase, overlap, or detector identifier. An example of a performance measure that can be binned is the number of split failures in a 15-minute period for a particular phase. Computing the individual split failures by itself may involve hundreds of stateless events and heuristic calculations, but a binning process precomputes and stores the data as a single record ahead of time. During the display process, the binned data can be quickly utilized without any additional computation.

Figure 2.14 illustrates an example of data accumulation over a 6-month period for a system with 70 intersections. The raw stateless events add up to over 3 billion records over the period. The interval aggregation process reduces the number of records by 77% from the stateless data, totaling 687 million records. Aggregating split failures into 15-minute bins reduces the number of records even further—a reduction of 99.3% from the number of stateless events, totaling only 9.6 million records.

Scaling high-resolution data for performance measurement also requires the implementation of a comprehensive set of *route definitions*. These definitions are data records that connect lane groups of different intersections together within a road network. These types of route definitions are essential for generating system-level performance measures such as the spatial oversaturation index (SOSI). Figure 2.16 is an example of a route definition table that describes a six-intersection route traversing five links a corridor. In each record, the upstream “from” lane group is mapped to a downstream “to” lane group at different intersections, and so forth. By joining records that have identical “from” and “to” lane groups, a direction of travel can be implied and built using the chaining of the records. The arrows in the table indicate the flow of the route mapping logic using conjoining lane groups, and going down the rows for each link matched.

2.13 Displaying Performance Measures

The simplest way to display performance measure information is in text format. In many cases, because of the nature and scale of the performance measure, text may be sufficient to convey information. For example, hourly volumes of one approach for a 24-hour period may be adequately represented by a comma-separated

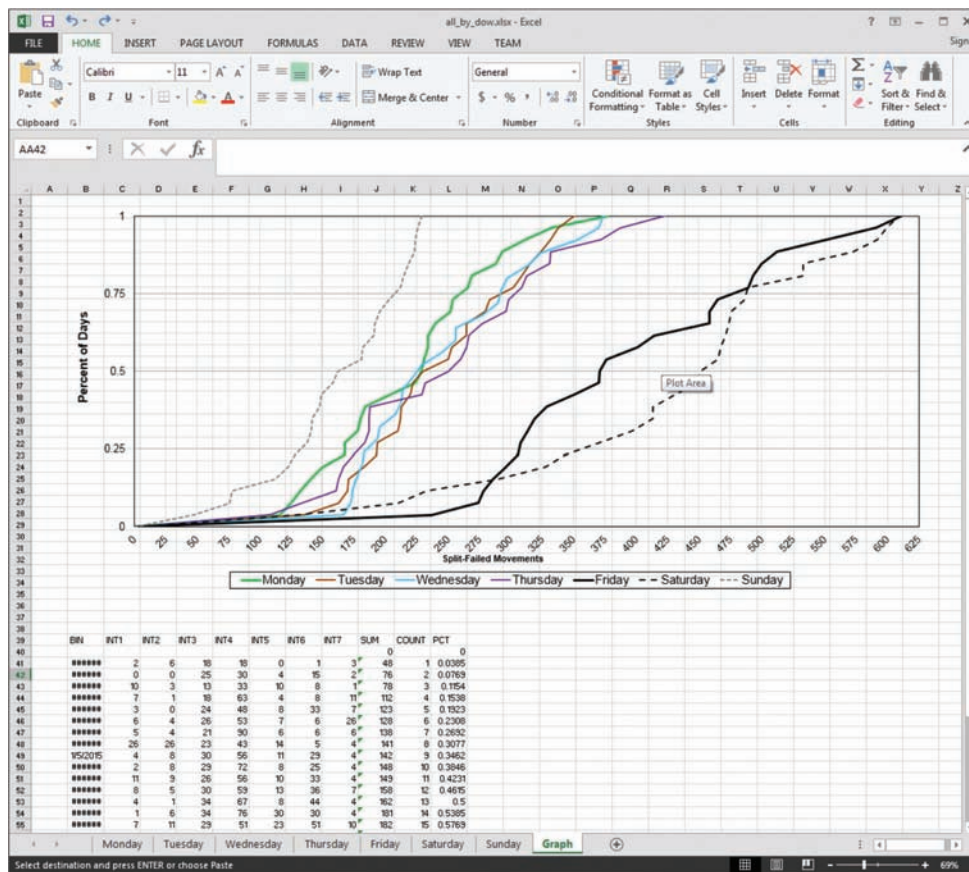


Figure 2.15 Example of using MS Excel to visualize performance measure data.

Link Number	From			To		
	Intersection	Approach	Lane Group	Intersection	Approach	Lane Group
1	Declaration Dr.	Northbound	Through	750 North	Northbound	Through
2	750 North	Northbound	Through	Stop 18 Rd.	Northbound	Through
3	Stop 18 Rd.	Northbound	Through	Apryl Dr.	Northbound	Through
4	Apryl Dr.	Northbound	Through	Madison Av.	Northbound	Through
5	Madison Av.	Northbound	Through	Smith Valley Rd	Northbound	Through

Figure 2.16 Example of a route definition table.

variable (CSV) text file with 24 lines. Obviously, there are many limitations with using only text, such as clarity, conciseness and effectiveness.

Graphically, performance measures can be displayed by off-the-shelf spreadsheet or charting programs, data visualization application user interfaces (API), cloud solutions such as Google Charts, other custom software, or a combination of software modules. Each set of tools comes with different costs, learning curves, and ease of implementation into the existing data system. However, all options ultimately achieve the goal of converting interpreted performance measure data into graphical displays.

Figure 2.15 shows an example of using MS Excel to plot a series of cumulative frequency distribution curves for the number of split failed movements per day of the week. The normalized and interpreted data used to plot the chart are computed from a separate software module that outputs the number of split failed movements per intersection per day. The same interpretation process can usually be performed in MS Excel itself,

albeit with varying degrees of complexity and efficiency. The pivot feature in MS Excel is used to partition and transform the dataset into individual days of the week, to sum across all intersections for a system view, and to provide charting features to plot the graph.

Alternatively, a web-based application leveraging charting APIs can be used. Figure 2.17 shows an example of using the Highcharts API in the HTML5 web environment for displaying travel time data along a signalized corridor, also using cumulative frequency distribution curves [30]. Figure 2.18 shows an example of using the same web libraries to plot split failures by 15-minute bins and PCD graphs for a 24-hour period using interpreted signal event data. In both cases, the display system is built on a distributed client-server architecture (sometimes called representational state transfer, or REST) where the normalized and interpreted data are packaged server-side and sent to the client, where the web browser's graphical functions are used to plot the charts.

Figure 2.19 provides an overview of the REST architecture utilizing web applications to generate

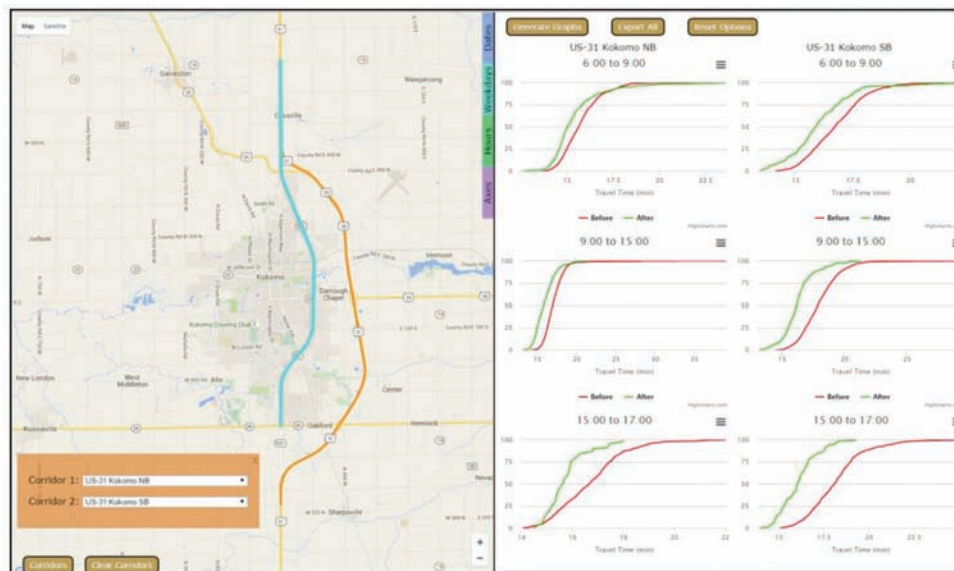


Figure 2.17 Example of using a web application-integrated charting API to plot travel time data.

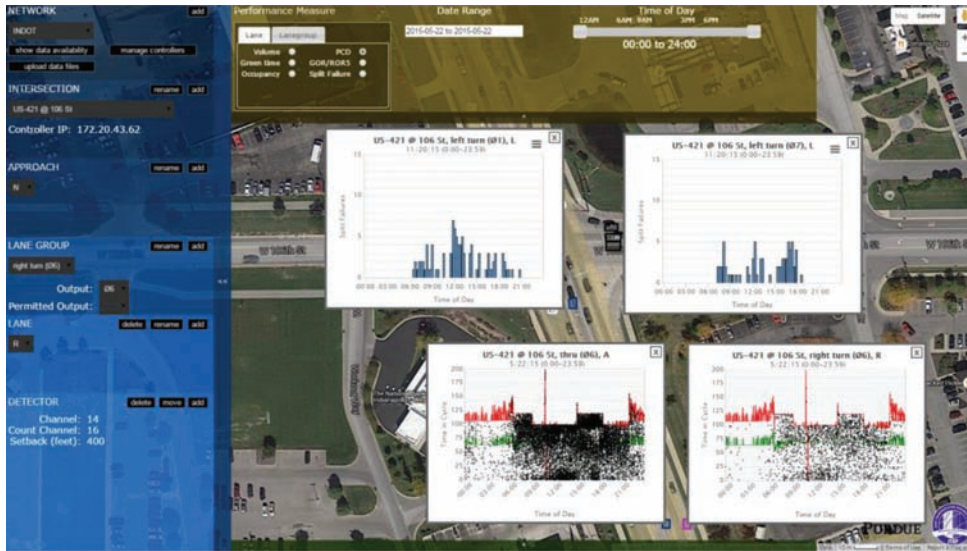


Figure 2.18 Example of using a web application-integrated charting API to plot performance measures.

performance measure graphics for the end-user. This architecture requires a considerable amount of software development, but it allows for greater accessibility and repeatability because multiple users can access the information from different locations (with consideration to encryption and security privileges) using a standard interface. The architecture also allows for a more

simplified process for displaying the performance measures. It can also be interpreted as a model-view-controller (MVC) software design pattern [40] because of how the database (the model), the web browser (the view), and the server-side business logic (the controller) interact together to generate a chart. This workflow of MVC is illustrated in Figure 2.20. Alternatively, Figure 2.21 provides an overview of the workflow using MS Excel to generate performance measures.

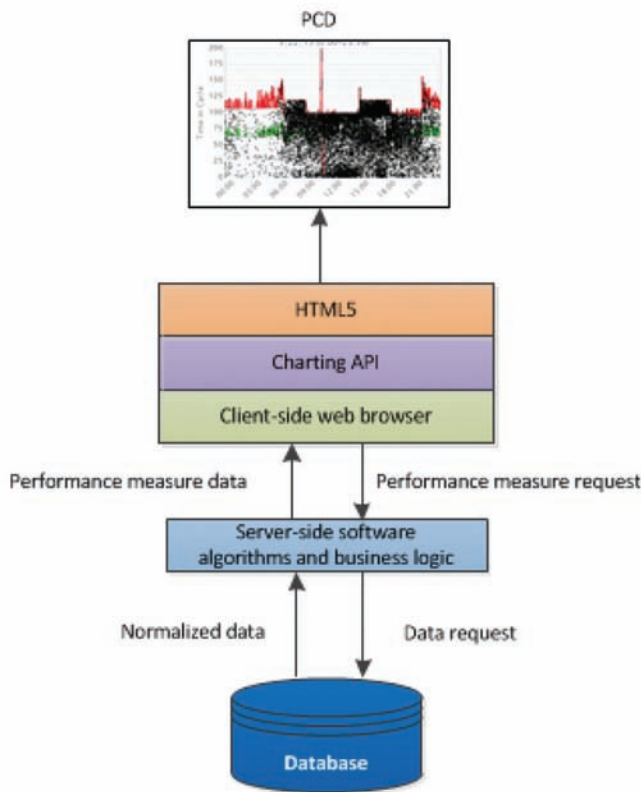


Figure 2.19 Architecture of a REST system for displaying performance measures.

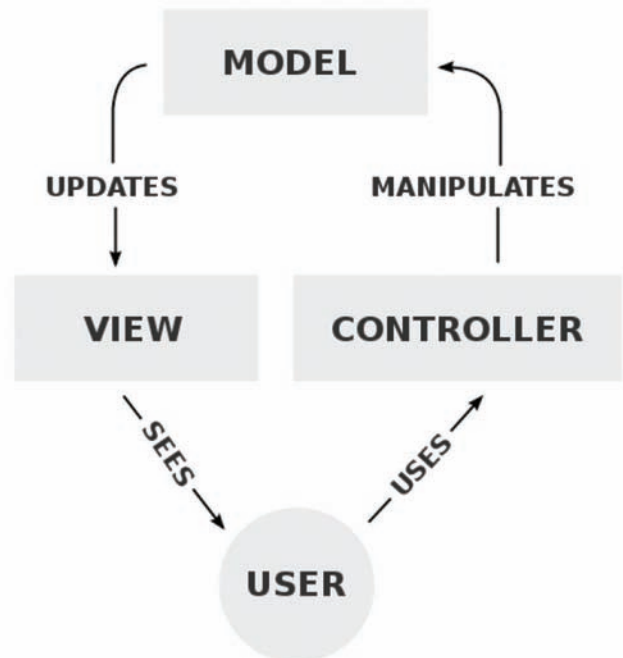


Figure 2.20 Model-view-controller (MVC) software design pattern.

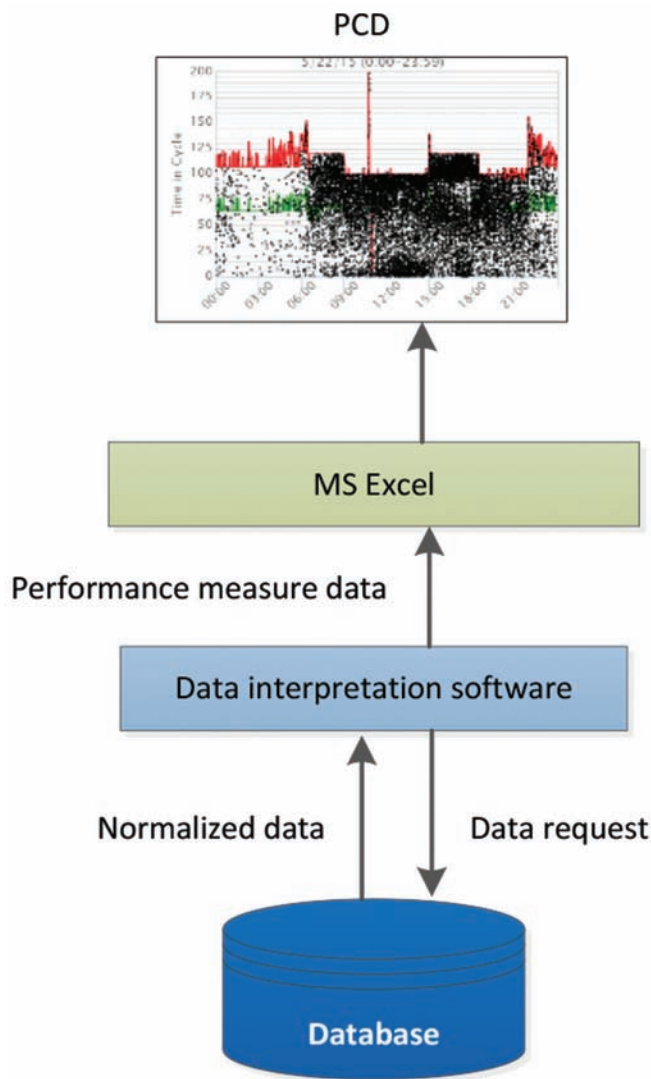


Figure 2.21 Architecture using MS Excel for displaying performance measures.

2.14 Conclusion

This chapter discussed the technical requirements pertaining to the generation of signal performance measures to fulfill agency objectives. First, a cabinet that accommodates the newest generation of NEMA, 170 ATC or 2070 controllers is needed. This will ensure that the data link between the detectors and controller is functioning using the most up-to-date protocols required for event data logging. Next, detector layout designs should be selected carefully for generating the suite of performance measures that will fulfill agency objectives. The detector layout information should be well documented after the layouts have been selected and implemented. Very critically, the detection mapping information is necessary for relating controller logged data to the spatial orientation of the right-of-ways at the intersection.

Once the data are logged in the controller, a communication network is necessary to transfer the data back to the traffic management or data center. Alternatively,

the data can be cached locally in the cabinet using flash memory cards on an embedded computer. In addition to controller data, travel time data provide a different dataset for outcome assessment, and can be measured using additional agency infrastructure or purchased through third-party commercial data vendors.

Finally, the data are downloaded to the traffic management center, where a series of processes normalize, archive, and interpret the event data for display to the end-user. Depending on the system size and archival period, different requirements for software and hardware are needed. Additionally, the data can be reduced and aggregated to increase performance. Once the data have been interpreted into the desired performance measure format, various off-the-shelf or custom software graphical applications could provide information for achieving agency objectives.

3. COMMUNICATION

In pursuit of strategies to enhance TSM&O, Chapter 1 discussed how new technologies like high-resolution controller data can be integrated with performance measurement techniques to improve agency business processes and decision-making. However, a challenge with utilizing these data is to ensure that they are available to a central agency or TMC in the first place. This chapter will present a methodology for utilizing server-side data aggregation tools for the purpose of monitoring the reliability of signal communications, and will detail the development of a visual dashboard intended to help agencies prioritize their communications maintenance resources. (This chapter adapts some material from S. Lavrenz's dissertation [62].)

3.1 Basic Signal Communications Architecture

Modern traffic signals make extensive use of various communications protocols at three primary levels:

1. Communications between electronic components internal to the signal cabinet (for example, the data streams that run between a rack of loop detector amplifiers and the signal controller)
2. Communications between signals along a corridor. These communications are frequently used to ensure proper time synchronization between adjacent intersections, for the purpose of maintaining efficient signal coordination and vehicle progression.
3. Communications to an offsite monitoring location, such as a central agency TMC. Such offsite data communications allow engineers to quickly view performance statuses for intersections over a wide geographic area, and help to facilitate long-term data archival for additional analysis.

The role of data communications by INDOT in traffic signal systems has previously been documented in the literature [20]. INDOT currently uses two methods of connecting to signal controllers for the purpose of clock synchronization and high-resolution data access: modem connectivity to individual controllers, and modem connectivity to a string of controllers

linked through a fiber optic network. These modems relay high-resolution controller data to servers at the INDOT TMC using file transfer protocol (FTP) channels, where they are archived and used to generate various performance measures. Figure 3.1 shows a graphical overview of the system architecture.

In addition to relaying high-resolution signal controller data to the TMC, wireless communications to signal controllers frequently serve a second critical function: clock synchronization. Proper clock synchronization is necessary to ensure that traffic signals start and end specific timing plans at the correct times, and for allowing coordinated intersections to stay in step with one another with consistent green offsets on major street approaches. By communicating with a time server at the TMC, individual controllers are able to receive periodic updates about the actual current time, and can then make corrections to their internal clocks based on any discrepancies.

Consequently, while maintaining remote communications with individual signal controllers is important for the purposes of transmitting high-resolution controller data, there are significant additional benefits, such as proper clock synchronization, that can be realized as a result of proper connectivity.

3.2 Accessing High-Resolution Signal Controller Data

In determining what constitutes “functioning communications,” there are several levels of protocol that must be traversed in order for high-resolution data to be properly transmitted to a remote storage facility for caching. In order of necessary steps to reach the data on the controller, these protocols are as follows:

1. *IP Configuration:* The signal controller must have a valid IP address configuration. A properly configured signal controller will be reachable at a static IP address via a cellular modem.

2. *FTP Setup:* The FTP functionality on the signal controller must be set up properly with valid login credentials.
3. *Directory Setup:* A standard directory must be specified on the signal controller for the storage of high-resolution data files. This directory must also be accessible via FTP.
4. *File Generation:* The controller should be configured to generate high-resolution data files, which are written to a directory on the controller. Typically, these files contain individual phase and detector event records for 15- or 60-minute intervals, depending on the controller manufacturer and firmware.

A lack of proper configuration for any of these steps will result in the TMC-located server being unable to reach the high-resolution data files on the signal controller. Note that these steps illustrate the category of potential communications failure, but not necessarily the mechanism. For example, a signal controller may be unreachable at a recorded IP address. However, this may be due to the IP address being incorrectly programmed in the controller itself, or possibly to a malfunctioning cellular modem in the signal cabinet.

The exact determination (and rectification) of communication failure mechanisms often requires a physical trip to the signal cabinet by a systems engineer or signal technician.

3.3 Identification of Failed Communication

An automated script was created to run on the INDOT TMC servers and cache high-resolution data files from individual signal controllers. This “file ingester” script is set to run every 30 minutes, and polls the IP addresses of every signal controller from a server-side database. Once a valid IP connection has been established with a signal controller, the ingester script attempts to navigate the file directory of the controller via FTP, and download the most recent high-resolution data files. If any of the errors described in the previous section occur (e.g., an invalid IP address or incorrect FTP credentials), the ingester is unable to

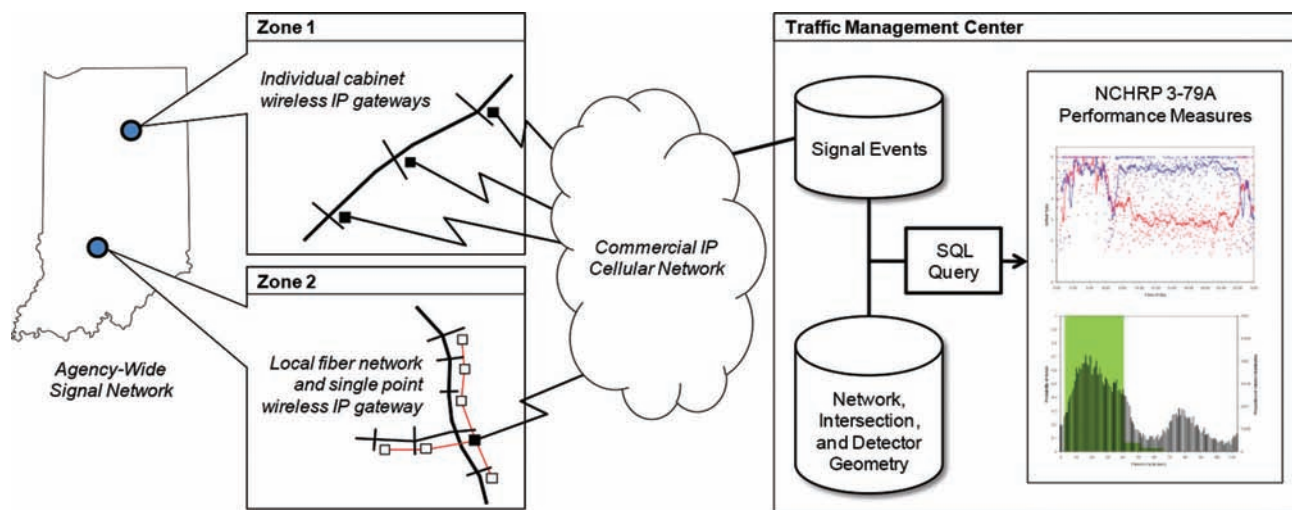


Figure 3.1 Overview of INDOT high-resolution data communications network [20,40].

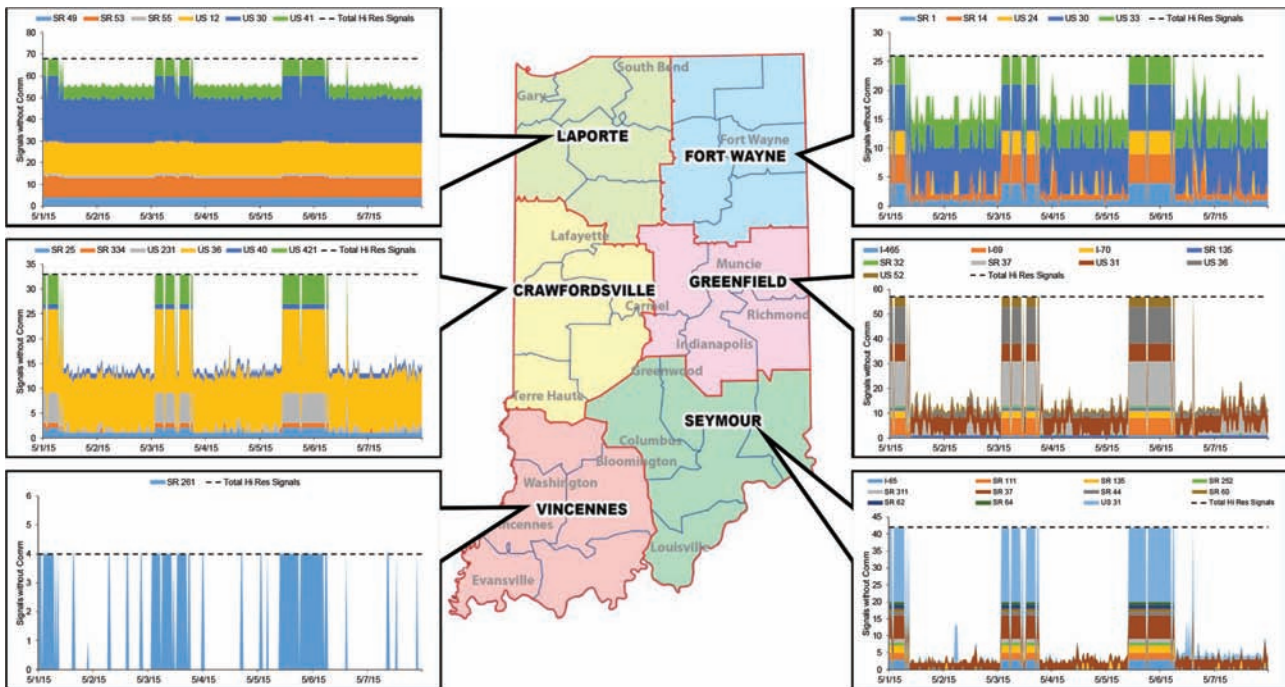


Figure 3.2 Statewide signal controller communications failures, by district and corridor.

download the high-resolution data, and an error event is recorded in a database table on the TMC server.

By using similar SQL-based query methods as for the identification of signal phases with failed loop detectors, these data ingester errors can help the agency understand which controllers have communications breakdowns. Furthermore, by cross-referencing the signal controllers in the ingester error table with a

table containing INDOT district information, a visualization of signal controller communications failures by district can be generated. An example of this visualization, where individual controller communications failures are shown over a weeklong period in May 2015, is shown in Figure 3.2. Here, the communications failures are grouped by individual maintenance districts, and further by corridor. Both Figures 3.2 and 3.3 are useful

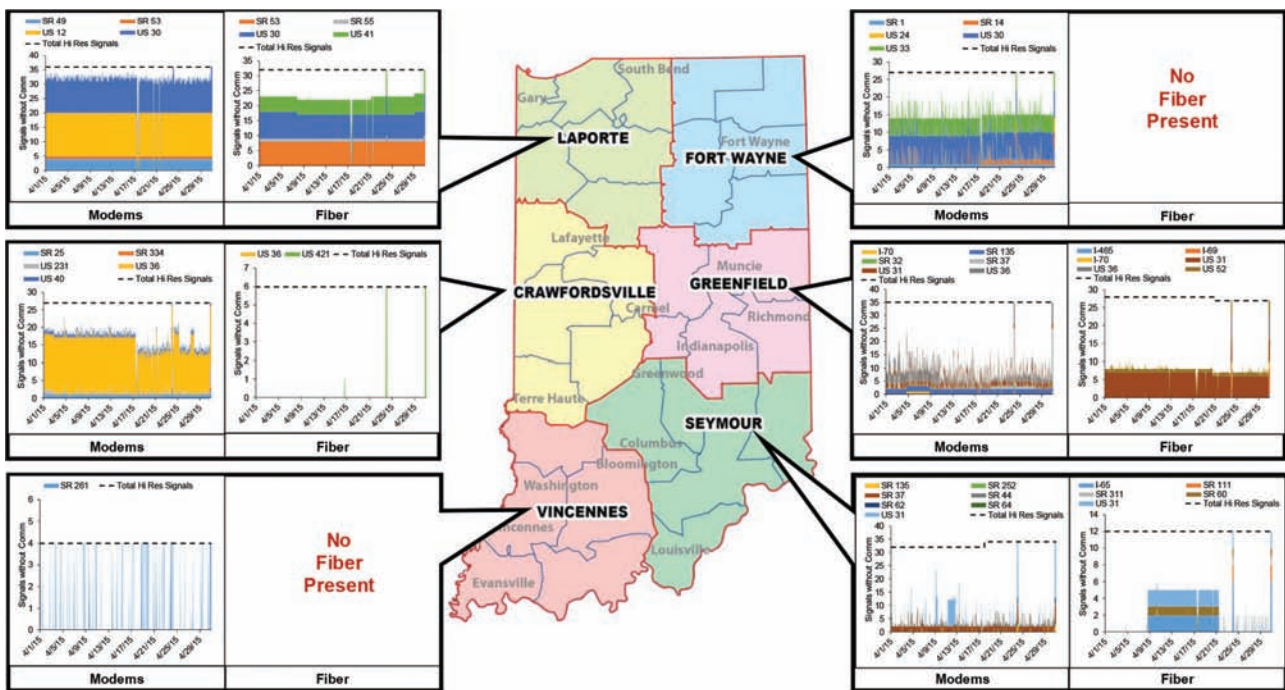


Figure 3.3 Statewide signal controller communications failures, by district, corridor, and communications technology.

for quickly visualizing signal communications failures at a high level of reporting. Failure patterns can be better understood by going one level deeper, and separating communications failures according to the method of connectivity between the signal controller and the TMC. As previously stated, these methods include direct cellular *modem* access, and *fiber* optic interconnections that transmit data for multiple controllers through a single cellular modem. Figure 3.3 also presents statewide

information on communications failures by district and corridor, except split between modems and fiber.

3.4 Long-Term Use of Signal Communications Failure Dashboard

The dashboard for visualizing signal communications failures was first presented to INDOT in April 2015, and has been utilized extensively to resolve communications

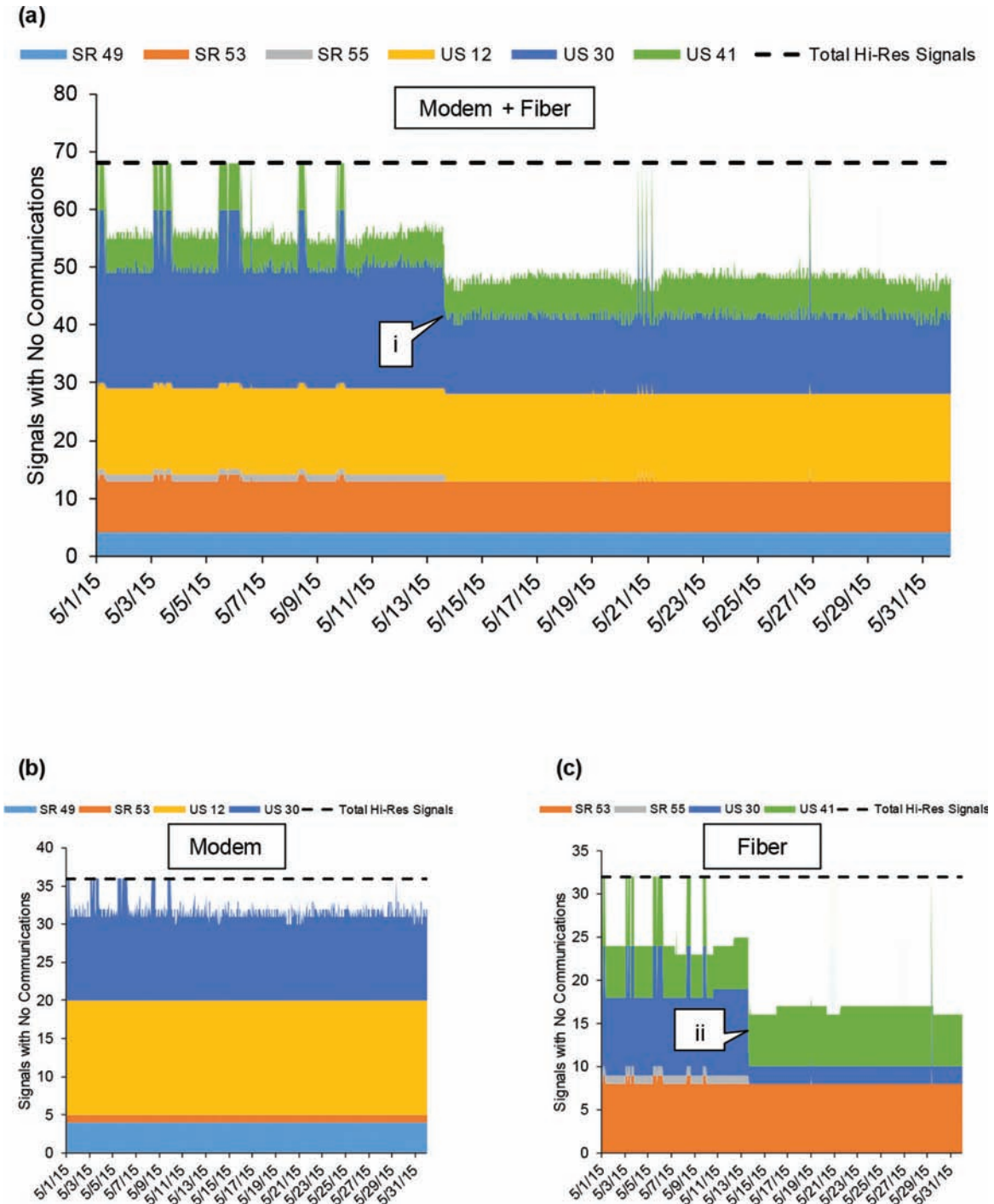


Figure 3.4 Longitudinal plot of signal communications failures for LaPorte district, May 2015.

maintenance issues for signalized corridors across the state. An example of the use of this dashboard to understand the nature of these issues is shown in Figure 3.4. This figure shows one month’s tracking of communications failures for the LaPorte maintenance district in northwest Indiana. In (a), it can be seen from the dashed line that whereas LaPorte currently has 68 traffic signals with high-resolution data capabilities (approximately 11% of all district-maintained traffic signals [61]), 55 were unreachable at the beginning of May 2015 owing to communications issues. On May 13, 2015, repairs were completed along US 30 in Porter County, which resulted in seven signals coming back online (shown by callout *i*). Parts (b) and (c) show that the repairs only impacted the fiber-connected signals along this corridor (see callout *ii*).

A longer-period view of signals communications failures for the LaPorte district shows that substantial progress has been made in addressing maintenance issues since the dashboard visualization was first implemented. Figure 3.5 presents a plot of communications failures for the district from April 1, 2015, through August 1, 2015. It can be seen that at the beginning of the plot, 55 of the 68 high-resolution traffic signals are unreachable by modem. This hits a low point of 23 signals with communications failures in late June, and finally settles at 26 signals with communications failures at the end of July. From April through August, this represents a 53% reduction in signals that are not reporting high-resolution data.

3.5 Conclusion

Hardware maintenance associated with signal controller communications is critical for ensuring the collection of high-quality event data for use in various performance measure calculations. Properly functioning communications also help to ensure that signal controller clocks remain in synchronization, which ensures good progression in coordinated signal systems. Only when a robust and reliable communications network has been established should agencies begin to focus on systemic methods of identifying other hardware issues,

such as loop detector failures. This is because a steady stream of high-resolution controller data is necessary in order for engineers to discern anomalous behavior that is caused by improperly functioning signal infrastructure.

A visual dashboard was presented that has helped INDOT to quickly identify corridors with significant amounts of failed communications hardware. By actively engaging signal technicians and systems engineers with this dashboard, it was shown that one maintenance district reduced the number of communications problems by 53% in a period of only four months.

Ensuring the proper maintenance and repair of signal communications is a critical first step toward achieving system-wide TSM&O objectives. Without proper communications, agencies are unable to access the high-resolution signal controller data that allow them to achieve a variety of other performance goals. However, communications maintenance alone is only the first step in a comprehensive ATM strategy. It is crucial that agencies consider the proper maintenance of other infrastructure elements as well, particularly for vehicle detection. The next chapter will discuss how agencies can utilize a high-resolution-based strategy to effect this goal.

4. DETECTION

Chapter 3 introduced the concept of utilizing high-resolution controller data to systemically identify malfunctioning traffic signal hardware. Once the basic level of data communications to and from the intersection has been assured, efforts can be made to identify other infrastructure deficiencies using the same data. One such area is found in identifying failed loop detectors. In the course of a signal timing intervention pursued for “capacity improvement” reasons, a basic assumption is that the signal hardware associated with the intersection(s) is in a state of good working order. The reality, however, is that various environmental factors, such as pavement freeze–thaw cycles (for pavement-embedded detection) and deferred maintenance, can result in malfunctioning traffic signal hardware. Inductive loop detectors are especially prone to failure from these external processes. These detectors are responsible for providing vehicle detection information to the signal controller, so that green time can be allocated in a demand-responsive manner. Broken detection can have significant deleterious impacts on intersection performance and vehicle progression; however, the proactive maintenance and identification of failures in loop detectors are often given little consideration by public agencies.

This chapter (which adapts material from recent papers [62,63]) presents case-based evidence of the importance of such hardware maintenance. It is herein posited that improperly functioning detection, for even a few intersections, can adversely affect mobility throughout the entire corridor. The methodology presented to identify such detector failures relies on high-resolution event-based data generated by the traffic signal controller.

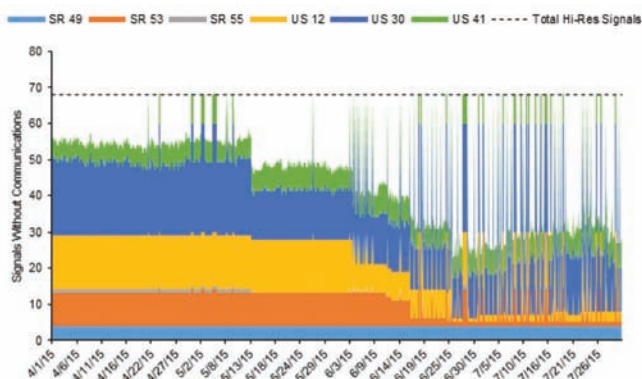


Figure 3.5 Four-month plot of signal communications failures for LaPorte maintenance district.

4.1 Background for Traffic Signal Loop Detector Repair

Ordinarily, the green indication on a traffic signal phase can terminate in one of two possible manners. In the case of undersaturated conditions, the phase will “gap out” once a certain amount of time (typically, between 2 and 4 seconds) elapses without a recorded vehicle detection. During oversaturated periods with standing queues, a phase will often be unable to provide sufficient green time during a single cycle to serve the entire queue.

A continuously extended phase will eventually be terminated by the controller after a certain amount of time, to ensure that sufficient green time is allocated to the other intersection phases. This is called a “max out” in a noncoordinated mode when the max timer causes the termination, or a “force off” during coordination when the split timer causes termination.

In the case where a loop detector for a phase is not functioning properly, the signal controller will often allocate the maximum allowable green time as a precautionary measure. Here, the phase is placed into a mode known as “maximum recall” and will be forced off by the signal controller, in a similar manner to a phase with high traffic volumes. Sometimes a “minimum recall” mode will be used instead—this ensures that a minimum amount of green time is allocated to the phase(s) with failed detection, although the effect (as a result of failed detection) is still the same: highly inefficient allocation of limited green time.

Several recent studies have examined methods to identify failed loop detectors. The Texas Transportation Institute published a comprehensive report in 2006 that made various recommendations for utilizing traffic signal data and maintenance records for ensuring quality control of signal infrastructure and detection. The report provided a state of current practice at several agencies, and found that current practices vary widely in terms of the extent to which field-collected data are utilized in the maintenance decision-making process. Generally, it was determined that although many agencies make use of limited traffic data in their maintenance processes, these data are often used inconsistently and often lack a systemic component needed to generate network-level performance assessments [64].

Methods for the automated identification of traffic signal maintenance issues were reported as far back as the 1970s, although these methods were based on analog signals communicated to a central server, and involved a fairly expensive and complex equipment setup for clusters of intersections in close physical proximity [65,66]. More recently, a number of publications have dealt with the topic of data-based methods for loop detector maintenance, but primarily within the context of freeway vehicle counting systems. For example, several publications have used detector occupancy data to identify malfunctioning loop detector units [67,68]. Other research has used real-time and historical traffic count information [69,70,71].

More recently, a 2014 publication utilized spatiotemporal loop detector counts to generate confidence intervals to identify failed detection [72]. While this method results in a highly accurate measure of failed detection, it involves a significant amount of data collection and postprocessing, which may not be tenable for many agencies. Besides these seminal publications, some public agencies publish their own comprehensive signal maintenance guides, although these tend to focus on specific repair techniques for individual signals and do not provide guidance for system-wide maintenance measures [73].

A number of manufacturers of modern traffic signal controllers have built advanced functionality into the hardware to enable the logging of various detector faults. However, these functions are often underutilized or programmed incorrectly in the field, for a variety of reasons, including a general lack of awareness about their existence. Consequently, it is incumbent on the traffic engineer to develop methods using well-known tools and data to identify problematic hardware quickly and effectively in a systematic manner.

4.2 Case Study: Corridor-Level Detection Failure

INDOT maintains approximately 200+ intersections around the state with high-resolution data capabilities. For these intersections, the number of phase force-off events was recorded over approximately 4.5 months, from October 2014 through February 2015. To ensure that the recorded force-off events would not include those resulting from legitimate high-demand traffic periods, a 1-hour period in the overnight hours, from 2AM to 3AM, was selected for the data compilation; it is highly unlikely that a phase force-off occurring during this time period would be due to high traffic volumes. Furthermore, to minimize the likelihood of such “false positives”, a threshold of five force-offs over the hour was set before a phase would be considered “in recall” due to failed detection. This threshold is somewhat arbitrary, and can easily be modified according to the comfort level of the agency.

Figure 4.1 shows a 19-week plot of phases with five or more forced terminations that occurred between 2AM and 3AM. Part (a) shows the number of phases meeting the force-off criteria for all high-resolution data signals in the state of Indiana with functioning communication. Here, it can be seen that several corridors demonstrate a significant number of phases in recall during the overnight period, including US 421, SR 14, and SR 37.

A few additional points are worth mentioning with regard to this figure. For one, corridors may show a higher number of phases in recall than other corridors for numerous reasons, such as greater data availability along the corridor. For example, if a corridor has numerous phases meeting the force-off criteria, but the signals are not equipped with high-resolution data logging, those phases will not appear in the figure. Therefore, corridors with higher numbers of phases in

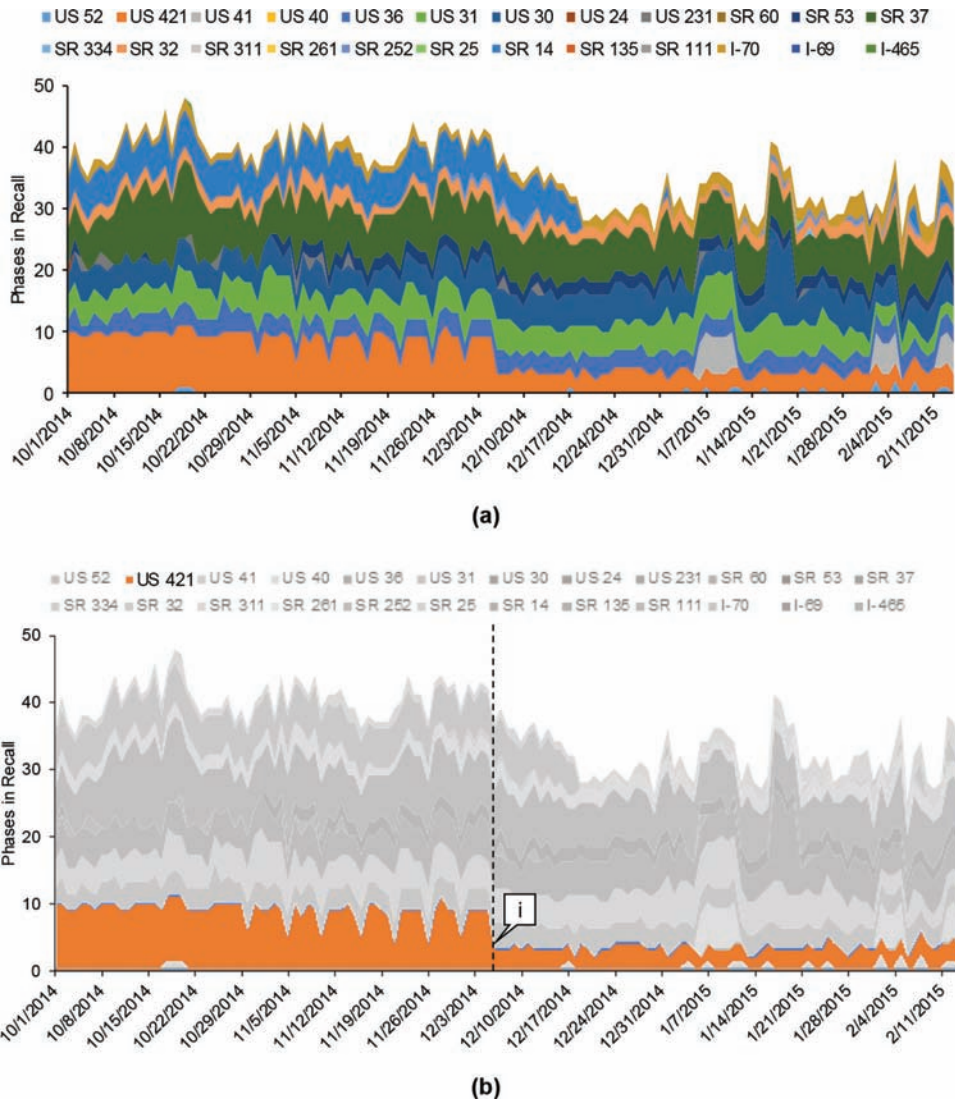


Figure 4.1 Longitudinal plot of phases from high-resolution equipped intersections with five or more forced terminations

recall should not automatically be assumed to have a greater total number of failed detectors. Additionally, a few corridors with higher traffic volumes or critical intersection movements may have phases that are purposely set in “max recall” mode during the overnight hours, ensuring that they will be forced off every cycle.

In part (b) of Figure 4.1, the US 421 corridor has been selected for more detailed analysis. Callout *i* shows that on December 5, 2014, the number of phases in recall dropped from 10 to 5, and remained largely consistent thereafter. The following section will examine the conditions surrounding the intervention that was responsible for this drop.

4.3 Case Study: US 421

Figure 4.2 provides a detailed look at the US 421 study corridor in northeast Indianapolis. This is heavily travelled urban arterial, with an annual average daily traffic (AADT) count of approximately 36,000 vehicles.

Significant minor street approach traffic is also present owing to intensive commercial development. Signals at the six intersections highlighted in the figure operate in actuated-coordinated mode from 0600 until 2200, with four different time-of-day (TOD) plans present on weekdays.

4.3.1 Visualizing Detector Errors

Once a corridor has been selected from the high-resolution data with potentially failed detection, some additional investigation is warranted to identify the specific problematic intersections and approaches.

By observing the EOG band and the pattern of vehicle detections, a traffic engineer can draw inferences about the signal timing that may be wholly or partly influenced by the failure of hardware at the intersection. Figure 4.3 shows the NB and SB PCDs for selected intersections along the US 421 corridor on December 3, 2014. It can be seen that detection

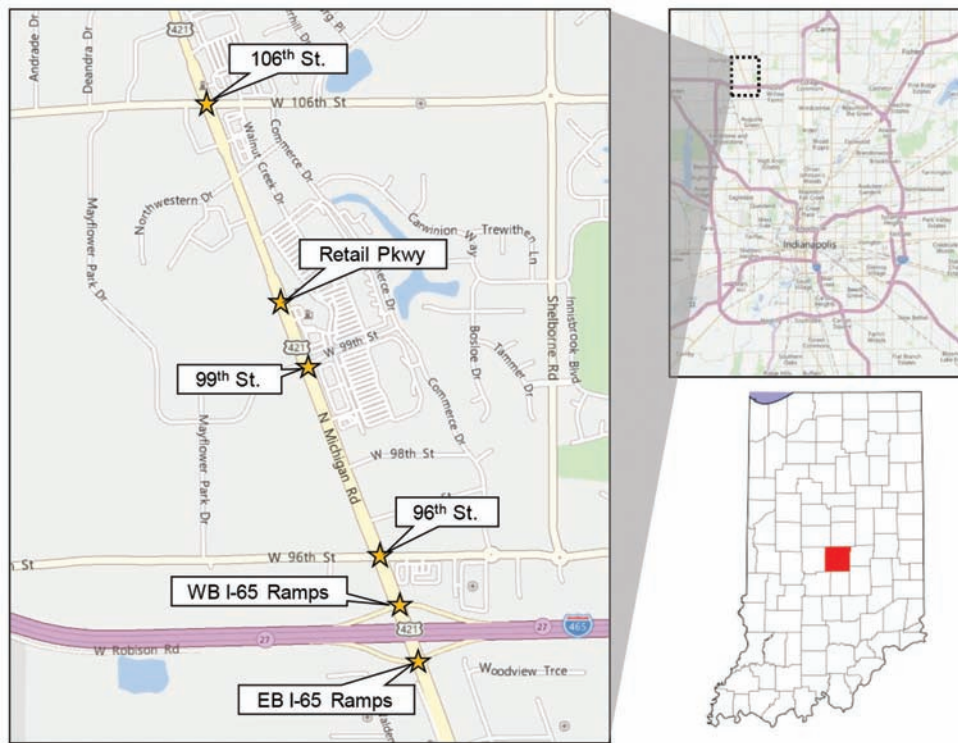


Figure 4.2 Overview of US 421 study corridor.

information is missing (callout *i*) on approach (b), for example, likely indicating a fault with the loop detectors. This is in contrast to approach (d) with fully functioning vehicle detection, as shown by callout *ii*. Callouts *iii* and *iv* on approach (b) emphasize several things pertaining to the cycle length that are also indicative of failed detection. Callout *iii* shows that the cycle length at the 106th Street intersection is nearly perfectly uniform at 100 s for most of the day. While this can often occur owing to high traffic volumes (effectively causing the phase to force-off at a fixed point for every cycle), it can also be symptomatic of a phase that is running its maximum green time for other reasons, such as failed detection.

The more telling indicator of a problem with the phase is shown in Figure 4.3 by callout *iv*. Here, even during the overnight period, the cycle length at the intersection varies between 90 and 120 s. This indicates that the signal is cycling through every programmed phase during this time period. Viewed in a different sense, the green is being diverted repeatedly from the major street to the minor street approach approaches during the entire overnight period. Although it is possible (but highly unlikely) that this is due to continuous vehicle arrivals on the minor street approaches, the far more plausible explanation is that there are failed detectors on the sides streets; in this case, the signal controller calls the minor street approach phases as a precautionary measure. A quick comparison with a properly functioning intersection confirms this; note the significantly taller EOG bands during the overnight

period at the I465 intersection, a consequence of the major street resting in green for multiple cycle lengths (owing to no calls being placed on the minor street approach phases).

It is clear from Figure 4.3 that there are most likely multiple detector failures occurring on one or more major street and minor street approach approaches at both 99th Street and 106th Street. The following discussion will focus on the 99th Street intersection only, as similar problems (and a similar resolution) were realized at both locations.

4.3.2 Repairing Failed Detection: US 421 and 99th Street

Field investigation at the 99th Street signal cabinet revealed several missing detector amplifier cards, corresponding to the malfunctioning phases. The same phenomenon was observed at the 106th Street intersection. The detector amplifiers are responsible for communicating to the signal controller the change in inductance that occurs when a vehicle drives over a loop detector; without these amplifiers, vehicle detections are not registered by the controller. The reason for the missing detector amplifiers was not immediately evident, but functioning amplifiers were installed and monitored to ensure that vehicle detections were being properly recorded.

A total of seven detector amplifiers were repaired or replaced in the course of troubleshooting this corridor. In this case, the detector malfunctions were due to missing amplifiers, although any number of other hardware

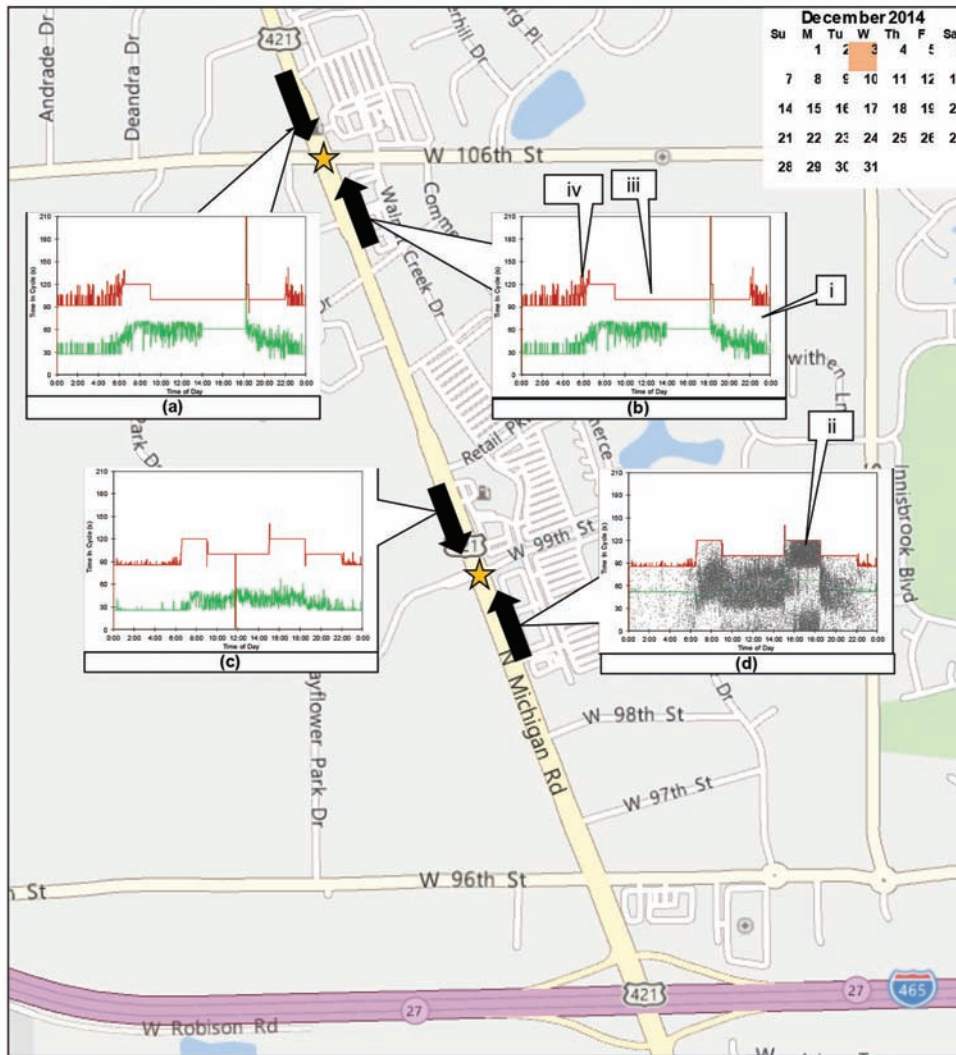


Figure 4.3 NB and SB PCDs for selected intersections along US 421, 12/3/2014.

malfunctions (e.g., broken amplifiers or damaged loop detector wiring) would cause similar observed behavior in the high-resolution data.

4.4 Detector Repair Results

It is not sufficient motivation to repair broken loop detectors simply for the sake of having functioning detection. From a business case and user satisfaction perspective, the agency must demonstrate substantive benefits from undertaking such maintenance activities. The following subsections explore some of the observed impacts on drivers on the US 421 corridor once the missing detector amplifiers were repaired.

4.4.1 Vehicle Progression

Figure 4.4 shows PCDs for the same intersections as in Figure 4.3, this time comparing PCDs before and after the detector repairs. The *Before* case is shown for December 4, 2014 and the *After* case for December 10, 2014.

Substantial differences are noted in the *After* case for both the 99th and 106th Street intersections. The functioning detection on the major street approaches is now evident, for example, as shown by callout *iii*. It can also be seen that cycle lengths at these intersections are significantly higher in the overnight period (callout *ii*), indicating that detection on the minor street approaches is now also functioning properly.

An additional performance improvement on the corridor is noted on the NB approach at 99th Street. Here, the vehicle detection was functioning correctly initially. However, a significant number of vehicles were still arriving on red (as seen by the large number of points below the BOG band in the “Before” case), owing to the poor signal timing caused by the broken detection on the SB approach. After the repair of the detector, it can be seen that vehicle progression improves substantially on the NB approach—most of the vehicles are now arriving between the BOG and EOG time periods. This is largely due to an earlier return to green on the phase, shown as a downward

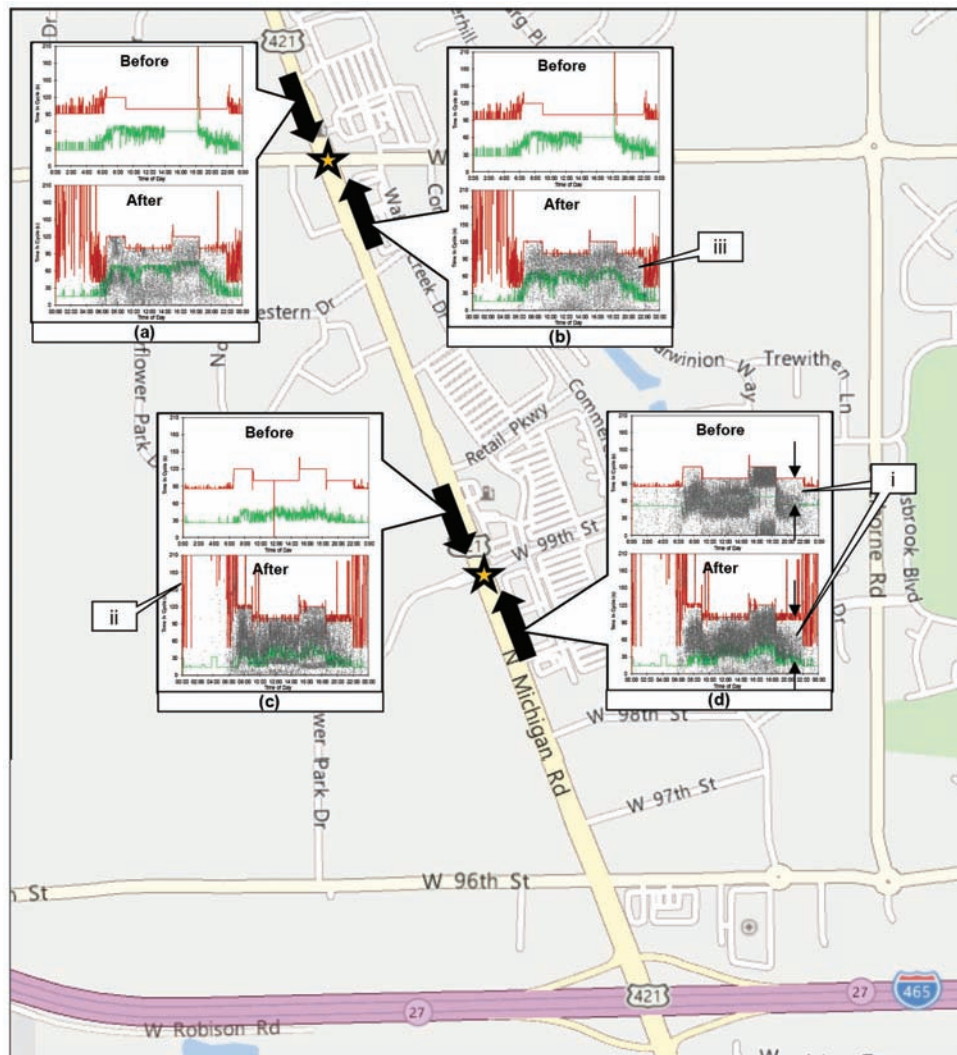


Figure 4.4 PCDs for select US 421 intersections before and after detector repairs.

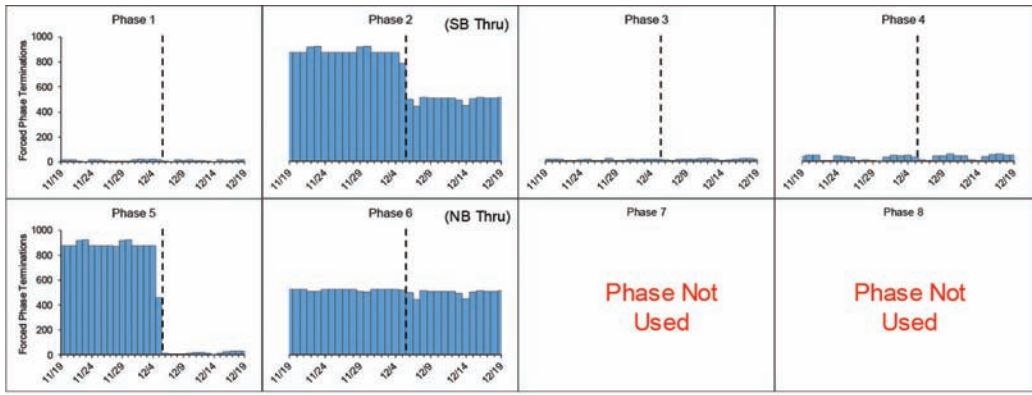
shift in the BOG band in callout *i*. Quantitatively, the POG was 56.1% before the detector repairs and 83.7% after the repairs.

The benefit to traffic on NB US 421 at 99th Street underscores the true importance of signal maintenance in a transportation systems context. This particular approach had properly functioning detection to start with, yet nearly half of all vehicles were arriving on red. By repairing detection on the SB approach, NB arrivals on green increased by approximately 50%. The dependency of individual intersection approaches on other phases and intersections in a coordinated signal system is not well understood from a maintenance perspective, and these results demonstrate the importance of focusing on systemic methods to maintain traffic signal infrastructure.

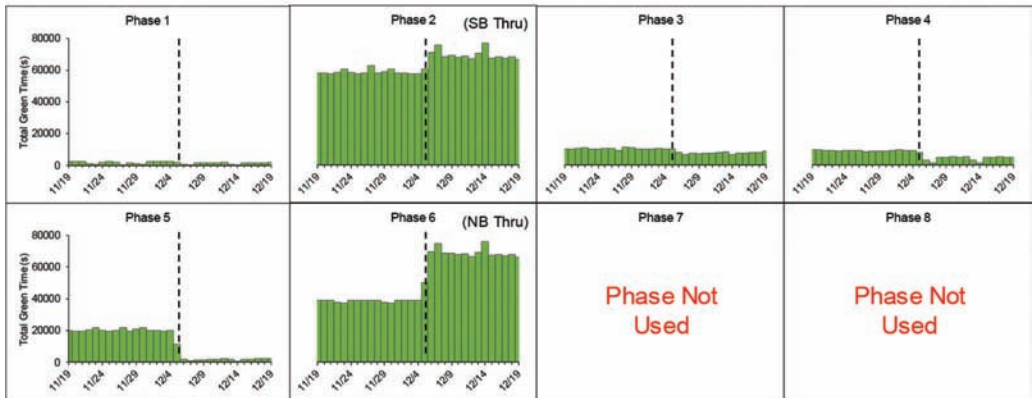
The change in green time on individual intersection approaches is better understood by examining these improvements in the context of phase termination behavior. Figure 4.5 shows a distribution of forced phase terminations and green time across different

phases at 99th Street, grouped by day, for two weeks before and after the detector repairs. In part (a), it can be seen that there are significant decreases in forced phase terminations on phases 2 and 5, which are mapped to the failed detection in question. This is due to the fact that the signal controller can now modulate the green time on these phases in a “demand-responsive” manner; that is, when low traffic volumes permit, the phases are able to gap out, instead of forcing off after running their maximum green time. A result of the drop in green time on the minor approach phases is that more green (capacity) is able to be returned to the heavy volume major street approaches. This is clearly shown in part (b), where the total daily green time on phases 2 and 6 (the major street approaches) increases considerably after the repairs.

At a daily level, it appears that the detector repairs had the effect of more efficient allocation of green time along the US 421 corridor. However, it also important to consider distribution of green time by hour of day as well. The reason for this is that minor street approach detector



(a)



(b)

Figure 4.5 Comparison of forced phase terminations and green time, two weeks before and after detector repair at US 421 and 99th Street. (a) Daily count of forced phase terminations. (b) Daily sum of phase green time.

repairs will often have the effect of allowing the major street to “rest” in green during the overnight hours, when little traffic is present to begin with. To be considered truly beneficial, it should be shown that these repairs resulted in capacity improvements on the major street approaches during periods of high demand. These improvements are illustrated in Figure 4.6, which shows hourly weekday green time for each phase, averaged over a week of observations before and after the detection repairs. It can be seen that significant reductions in green time occur during off-peak hours on several minor street approach phases, as well as the major street turning movements (phases 1 and 5). However, it is also shown that the detector repairs resulted in significant additional green time to the major street approaches during peak and midday periods (especially phase 6). In particular, callout *i* demonstrates a nearly 90% increase in green time between 5AM and 6AM (from just over 1400 s to just under 2700 s). Similar improvements are realized during adjacent hours, and are supported by the previous discussion on the shift in BOG for this phase demonstrated in Figure 4.4.

On the basis of this evidence, it can be concluded that making repairs to vehicle loop detectors not only provides benefit to drivers on approaches with the

failed detection, but can also substantially improve performance for drivers on other approaches within the same corridor. The holistic nature of the benefits realized by properly maintaining loop detector infrastructure can serve as a substantial motivator for agencies to adopt more proactive signal infrastructure management strategies.

4.4.2 Travel Time Improvements

Beside improvements in individual intersection performance, benefits can be assessed at the corridor level, namely, in terms of travel time and travel time reliability. Figure 4.7 shows cumulative frequency diagrams (CFDs) of travel time for both directions along the US 421 corridor for various times of day. Minute-level speed data are gathered for the corridor from a third-party source and converted to travel times. The travel time distributions are plotted for individual weeks from October 4, 2014, through February 28, 2014. Composite distributions for time periods before and after the repair of detection are overlaid on the weekly distributions for better visualization of the broader trends.

Significant leftward shifts in the CFDs (indicating travel time improvements) can be seen in the SB

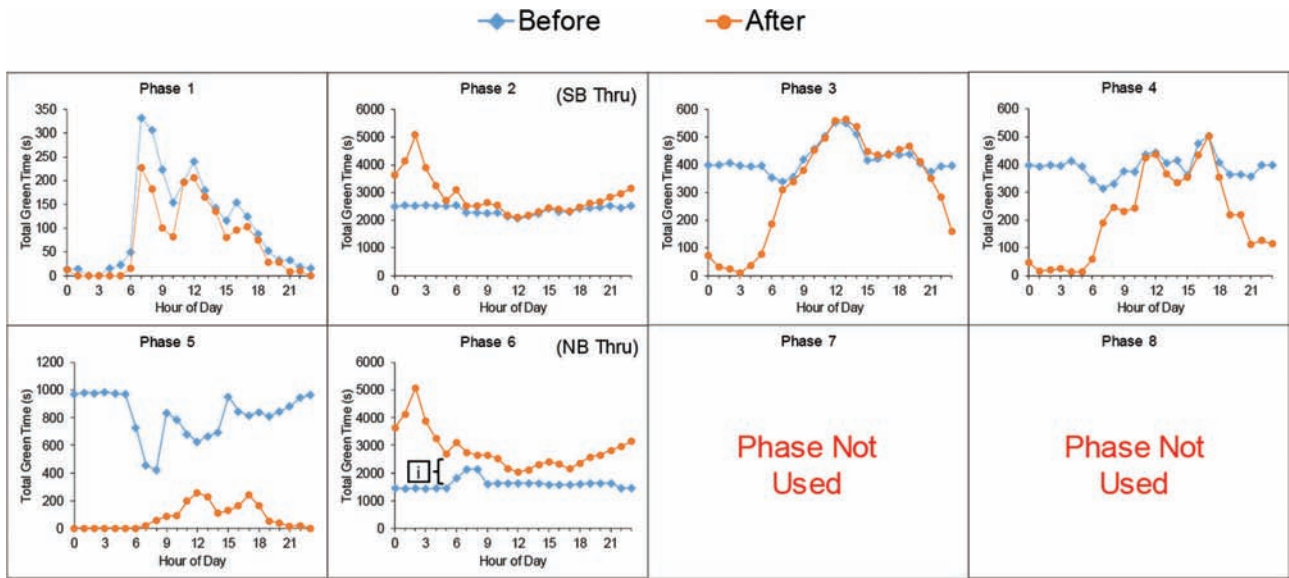


Figure 4.6 Average weekday green time by hour of day before and after detection repairs for US 421 and 99th Street.

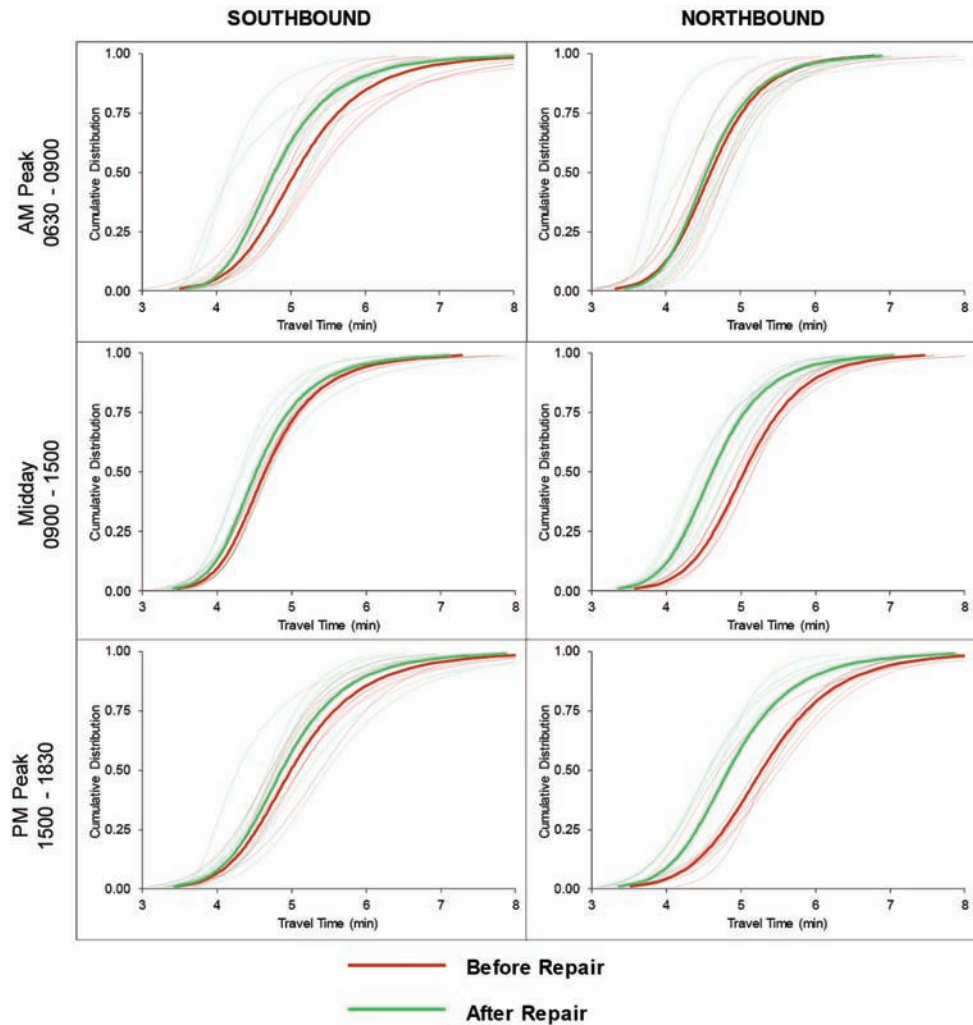


Figure 4.7 US 421 cumulative distribution functions of travel time, by direction and TOD.

TABLE 4.1
Changes in median corridor travel times (in seconds) as result of loop detector maintenance.

	Southbound				Northbound			
	Before	After	Δ	% Δ	Before	After	Δ	% Δ
AM Peak	296	280	-16	-5.4%	271	269	-2	-0.7
Midday	279	272	-7	-2.5%	303	277	-26	-8.9%
PM Peak	296	288	-8	-2.7%	324	293	-31	-9.6%

direction during the AM Peak TOD plans, and in the NB direction during the Midday and PM Peak TOD plans. These improvements correspond to the predominant directional traffic volumes during these time periods, suggesting that most drivers experienced significantly lower travel times as a result of the detector repairs. Specifically, median travel time decreased by 6% (16 seconds) in the SB direction during the AM Peak, 9% (26s) in the NB direction during the Midday period, and 10% (31s) in the NB direction during the PM Peak period. A complete tabulation of median travel time changes can be found in Table 4.1.

It was also verified that traffic volumes along the corridor did not change substantially during the 19-week analysis period, which could also affect travel times. The general shape of the travel time distributions does not appear to change, suggesting that most of the travel time benefits are realized as a result of average travel time reductions, rather than a change in travel time variance or reliability.

4.5 Benefit/Cost Analysis of Detector Repairs

As a means of further illustrating the business case for carrying out data-based maintenance of traffic signal infrastructure, a benefit/cost ratio (BCR) analysis was conducted for the US 421 intervention. Figure 4.8 presents annualized estimates of user benefits due to travel time improvements along the US 421 corridor.

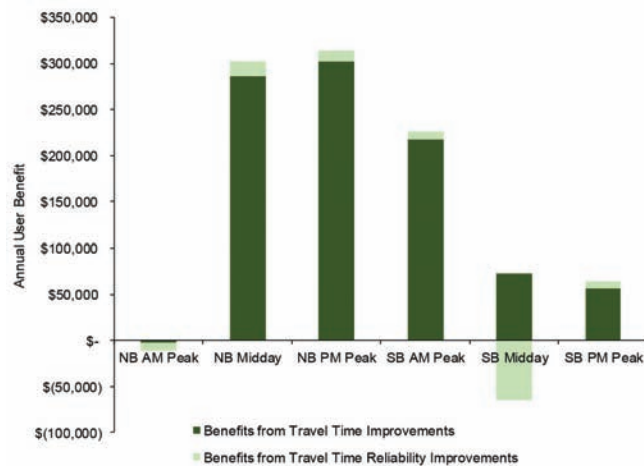


Figure 4.8 Annualized user benefits from travel time improvements due to detector repairs.

It can be seen that the largest annual benefits are realized by those directional and TOD plan combinations that showed the largest shift in the travel time CFDs in Figure 4.7. In particular, the NB Midday and PM Peak TOD plans, as well as the SB Midday TOD plan, all realize between \$200,000 and \$300,000 in annual user travel time benefit. The SB PM Peak TOD plan sees modest annual benefit due to average travel time improvements, while the SB Midday TOD plan sees a net benefit close to zero, owing to average travel time improvements being offset by a worsening of travel time reliability. Similarly, the NB AM peak period does not realize a significant positive or negative benefit from the detection repairs.

With the exception of the SB Midday TOD plan, it appears that the impacts of detector repairs on travel time reliability were quite small; nearly all of the user benefit along the corridor comes from changes to average travel time. In some sense, this is to be expected, as weekday traffic volumes along the corridor are very consistent from week to week; furthermore, with the failed detection due to missing amplifiers, it is extremely unlikely that there would be any intermittent functioning of the broken detectors. Thus, any travel times realized by vehicles along the corridor, regardless of detector functionality, would either be consistently good or consistently poor.

To conduct the BCR analysis, information was gathered from INDOT to determine the approximate costs associated with the loop detector repairs. Total repairs costs were estimated at \$4,000, with \$3,000 for parts and \$1,000 for labor. From this information, a 1-year BCR was calculated, under the assumptions that travel time improvements were the only significant monetizable benefit and that negligible additional maintenance cost was incurred as a result of the repairs over the analysis period. Because of the inherent uncertainty in estimating costs and benefits, a range of BCRs was

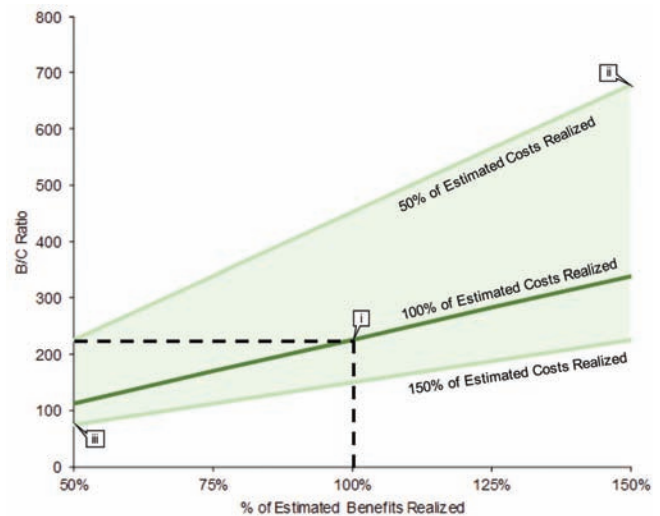


Figure 4.9 BCRs for detector repairs along US 421. The shaded areas correspond to actual costs between 50% and 150% of those estimated.

computed. Figure 4.9 shows the range of these BCRs. On the benefits side, it was assumed that actual benefits could be as high as 150% of those estimated, and as low as 50% of estimated benefits. A similar range (50–150%) was used for estimated costs. The shaded region of Figure 4.9 illustrates this range of cost estimates, while the solid lines correspond to actual costs being 100% of estimated costs.

If the benefits and costs associated with the detector repairs were to be approximately realized as they were estimated (callout *i*), the BCR would be 226. The best case 1-year BCR (callout *ii*, assuming actual benefits are 150% of those estimated, and actual costs are 50% of estimated costs) is approximately 680, while the worst case 1-year BCR (callout *iii*, 50% of estimated benefits, 150% of estimated costs) is 75. The shaded area of Figure 4.9 provides BCRs for a range of realized costs and benefits between these scenarios. Regardless of the actual benefits and costs realized from the detector repairs, it can be seen that such maintenance activities are overwhelmingly worth pursuing.

4.6 Conclusion

The maintenance of loop detectors is an important part of ensuring effective overall management of traffic signal infrastructure, and has significant impacts on reducing driver delay and maximizing system performance. This research used high-resolution signal controller event data as a means to systemically identify intersection phases with failed detection around the state of Indiana. Several specific findings resulted:

1. The use of forced phase terminations in low-volume periods can be an effective alternative detection failure identification method, compared with more advanced loop detector monitoring tools, which are often unutilized by public agencies. The intuitive nature of this methodology improves the likelihood that field maintenance personnel will believe in and trust this measure.
2. Minor problems in a traffic signal hardware system (such as several missing detector amplifiers) can cause significant delay for drivers on a corridor level. In the US 421 case study, where seven missing amplifiers were replaced (an estimated cost of \$300/detector card, plus labor costs), approximately \$900,000 in user benefits were obtained. The relatively low cost of repairs, coupled with significant user benefits, resulted in a 1-year BCR of between 75 and 680.
3. The repair of problematic detection resulted in substantial reductions in average travel times for several directional and TOD plan combinations, as well as modest improvements in travel time reliability. The most significant improvements were seen in the NB direction during the midday and PM peak periods, which saw median travel time reductions of 9% and 10%, respectively.
4. Agencies rarely measure the impact of maintenance activities. With the ability to use modern signal systems to both identify maintenance opportunities, and the ability to measure impact, it is important that practitioners and agencies develop integrated dashboards that allow agency staff to share success stories with elected officials and decision makers that set maintenance budgets.

Given these results, it is evident that the proactive maintenance of traffic signal infrastructure is a highly cost-effective method of reducing traffic congestion, travel times, and driver dissatisfaction. Agencies would be well served by utilizing data-based methods of signal infrastructure management, and should consider the investment in such technologies as a necessary component of their long-term active traffic management strategies.

5. EVALUATION OF LOCAL SIGNAL CONTROL

5.1 Overview

This chapter presents uses of performance measures and visualizations based on high-resolution data to assess the quality of local control. This concerns the distribution of green times and trends in their initiation and termination at each intersection. The scope of this chapter encompasses performance aspects of individual intersections. The associated performance measures are then aggregated to system- and network-level views to show how to spatially and temporally locate “hot spots” of poor performance. Therefore, the problem of local control is considered both from a “local” and from a “system” perspective. The problem can be rather complex. There are more approximately 74,000 total changeable parameters in a modern signal controller, of which about 2,000 are likely to be changed from default [74]. While such optimization can enable elaborate capabilities and complex timing plans, it increases the likelihood of programming errors, even before getting to the problem of addressing correctly programmed control parameters that are suboptimal.

System control is discussed thoroughly in Chapter 6. System control imposes constraints on local control; this chapter will touch on these from the perspective of its local impact on the distribution of capacity. The next chapter will then explore the outcomes of system control from the perspective of signal progression.

5.2 Basic Phase Service Concepts

The purpose of signal control is to share time between conflicting rights-of-way, which involves the safe and efficient control of when green and walk intervals begin, how long they are served, and how they terminate. Therefore, a good starting point to discuss how to evaluate aspects of this endeavor is to break it down into its various processes.

Figure 5.1 provides a block diagram showing phase initiation and termination processes for actuated signal controllers. Six steps are identified, as follows:

- *Determine call for service on phase.* The controller watches for demand using detectors and registers calls for individual phases based on various settings. These demands are then served in a sequence determined by the controller programming. Phases that have no detection typically have constant calls, and a fixed-time controller has constant calls on all phases.
- *Determine min/max green duration.* Prior to service of the phase, the controller makes some choices regarding the

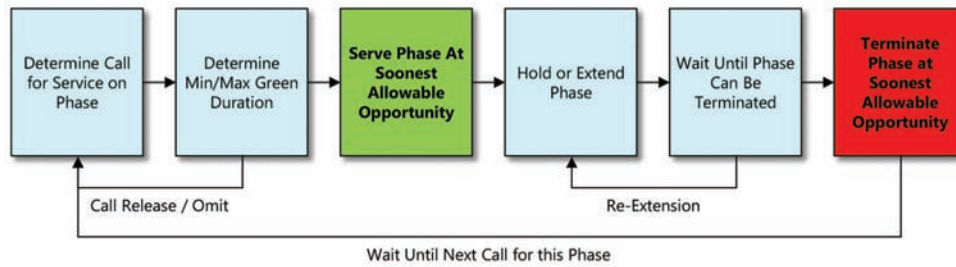


Figure 5.1 Phase initiation and termination processes.

phase service based on various settings, depending on the type of control being used. The simplest example is the selection of minimum and maximum green from the phase table during fully actuated operation. During coordination, force-off times are determined. Other forms of control may use other constraints negotiated between the local controller and the system. During this process, calls could potentially be released or omitted, which sends the controller back to the first step of waiting for calls for the phase.

- *Serve phase at soonest allowable opportunity.* The controller will begin service on the phase as soon as there are no other higher priority phases in sequence that must be served first. Some methods of control enable phase sequences to be dynamically adjusted, which would influence this process. At the end of this process, the phase will be in service: that is, in a green and/or walk state.
- *Hold or Extend Phase.* After the start of green, the controller will set up one or more timers that will ultimately cause the phase to terminate, if it is extended by detectors past the minimum green. For example, a max timer will begin counting as soon as there is a conflicting call and a split timer will count down according to the current time in cycle. Besides extensions, a hold may be placed on the phase by a coordinator or some other form of control.
- *Wait until Phase Can Be Terminated.* The duration of green on a phase is held until the controller determines that it is allowed to end it. Constraints such as the minimum green cannot normally be violated. The phase may also need to wait for other phases (e.g., in other rings) to end service before they can terminate. There must also be some reason for the controller to terminate the phase. When there are no other demands to serve, the signal will *rest* in the current phase unless instructed to go

to all-red. If simultaneous gap-out is used, the phase will be re-extended and will return to the prior step.

- *Terminate Phase at Soonest Allowable Opportunity.* Once the termination conditions are met, the phase will end. The controller transitions the phase into its clearance states, which may be modified according to conditions. This ends service on the phase and returns it to the first step while the controller moves on to serve other phases.

Each process brings some opportunities for analysis. The presence of vehicles in detection zones and of pedestrian calls (as well as other modes), the time when calls are registered, the time waited for service, the duration of service, the utilization of the service interval, and the process of termination can all be examined in microscopic detail. The results can be aggregated over time and across intersections to broadly evaluate performance and pinpoint operational deficiencies. Also, microscopic data enable diagnostics relevant to controller programming, as well as debugging of logic in the controller.

The impact that phase timing has on vehicle traffic is illustrated by Figure 5.2. This shows a time-space diagram with trajectories of vehicles that are stopped at a signalized approach. Two cycles of red indications are shown. In this particular example, vehicles arrivals are randomly distributed over time. Each vehicle comes to a stop at some point, and its trajectory becomes horizontal; these form queues that are released during green. This example also shows a case where there is more traffic than can be served by the green time during the cycle illustrated, a condition that we will call a *split failure*. The two vehicle trajectories affected by that

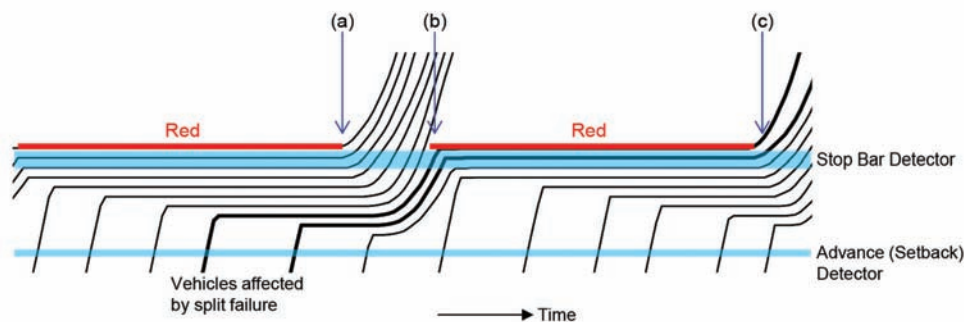


Figure 5.2 Illustration of local control of a phase, including an example of a split failure.

failure are represented by thicker lines than the rest. Every split failure represents a potential public complaint call from a motorist stating that they didn't receive enough green time.

Labels (a), (b), and (c) in Figure 5.2 correspond to video stills in Figure 5.3 showing conditions at different times in the cycle. The images show a double left turn lane. In Figure 5.3a, there is a long queue of vehicles waiting for service at the beginning of green. Note the two trucks in the queue at this time, particularly the second truck, which is highlighted. Figure 5.3b shows the vehicles at the *end* of green; the second truck has advanced forward but was not served, along with several other vehicles. It is not until the next green (Figure 5.3c) that these vehicles have a chance of passing the intersection.

The record of vehicle trajectories contains practically all of the information that we need to know regarding the quality of service: the arrival time of each vehicle, its departure time, the total delay time, the amount of time spent in a queue, its queue position, number of cycles

needed to traverse the intersection, and so forth. However, this information is almost never known by the analyst. Some high-resolution vehicle trajectories might soon be retrieved from connected vehicle and probe vehicle data, but these are unlikely to represent a *complete* record of traffic for many years.

Detector information, on the other hand, is readily available at many locations. The two basic types of detectors most commonly used are illustrated in Figure 5.2 by the blue shading that passes over the trajectories. The presence of any lines across this region represents vehicle presence in the detector. The stop bar detector, approximately two vehicles in length, is constantly occupied, while the advance or setback detector, which is shorter in length and away from the queues, is intermittently occupied, with the blips representing arrivals. This detector data offers the opportunity to report various types of performance measures, as Chapter 2 discussed in more detail. At the outset, it is clear that oversaturation is related to high occupancy and, potentially, a high vehicle count. As the

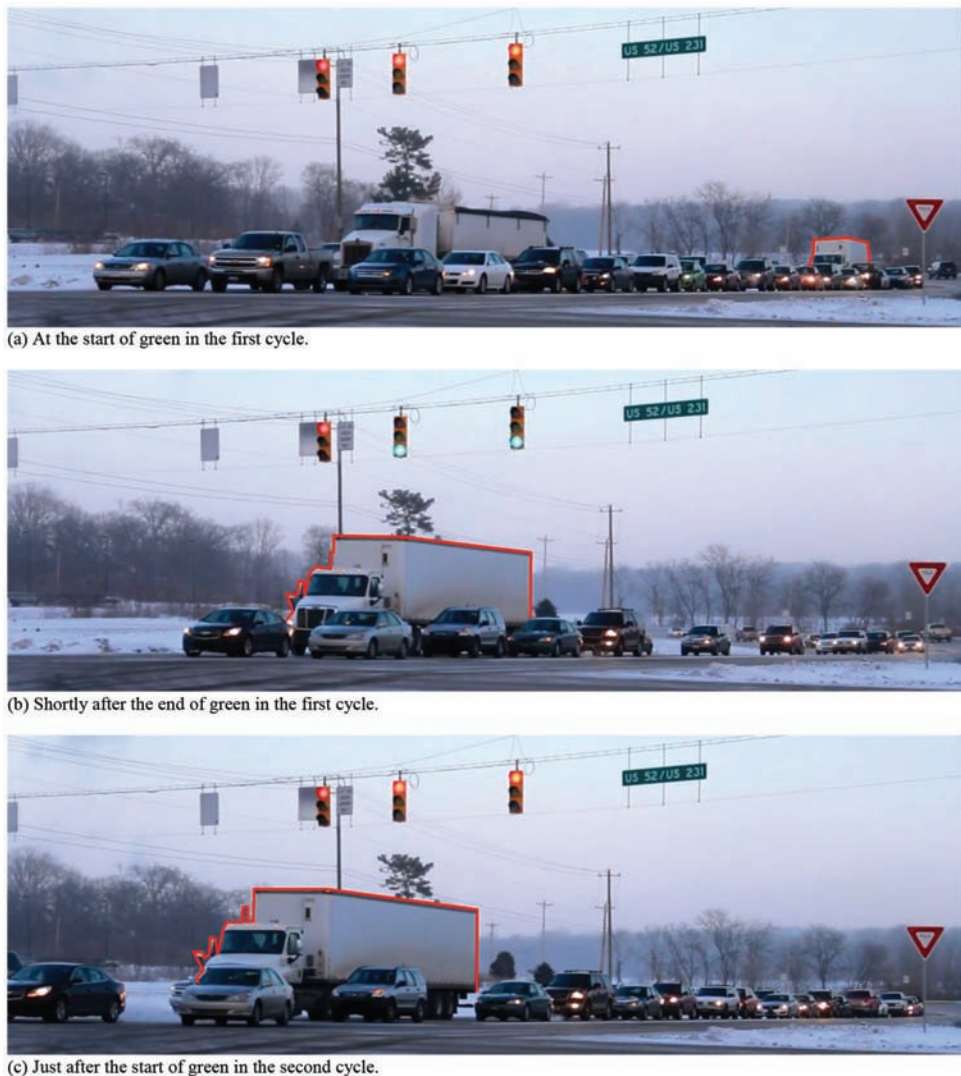


Figure 5.3 Illustration of a split failure. The highlighted truck is the same vehicle in the three images.

following discussion will show, this is indeed true, but the detector data can be put to better use when coupled with phase status.

5.3 Flow Rates

Perhaps the first use of high-resolution data that comes to mind is to measure the amount of traffic passing through the intersection. Traffic volumes, particularly AADTs and turning movement counts, are the basic data needed to for virtually all of the traditional supporting tools for the business processes of planning, design, and operations. The use of existing detectors at intersections has long been recognized as a convenient way to *automatically* collect this data [43], which traditionally has required manual counting. For example, INDOT has a longstanding program of collecting traffic counts that started in the early 2000s. High-resolution data provides another means of obtaining these volumes.

Figure 5.4 shows vehicle flow rates arranged by phase as seen in the INDOT performance measures website. Each of the eight plots shows the total volume for the 24-hour period and a trend in the volume throughout the day. There are characteristic peaks in the volumes corresponding to AM and PM rush hours that are very typical of most signalized intersections on weekdays. Such views of traffic volumes have been extensively examined previously [2,10,15], so this report does emphasize it further. However, such information is essential.

The development of signal timing plans would be currently possible by entering the volume data into design software, although one could imagine full integration into such software in future. The data can also be used more directly to develop signal timing plans using relatively simple calculations [75].

Although they provide a great deal of valuable information for many different applications in traffic engineering, flow rates alone provide only a limited amount of useful information about the quality of operations at a signalized intersection. They do not directly tell whether adequate green time is provided, but, rather, the analyst would need to enter the values into a model along with the programmed signal settings and make a rough estimate. Those models tend to work fairly well for fixed-time control, but actuation and more dynamic types of control are more difficult to accurately model in such a way. For this reason, this report does not discuss volume data further, but instead moves on to performance measures that relate to the allocation of green times and how these interact with the traffic.

5.4 Evaluating Local Timing in Detail

A more complete picture of intersection operation can be developed when information about the actual allocation of capacity is combined with detailed information about the utilization of that capacity. This allows the evaluation of how well the capacity is being distributed, as well as various other aspects of the performance.

Ideally, an actuated phase will receive as much green as needed to serve the demand, with the phase terminating shortly after the demand has cleared in order to expediently begin service on the phase. There are numerous strategies that control this process, varying by the type of control (fully actuated, actuated-coordinated, or adaptive) and the type of detection. The type of detection also determines the sort of performance measures that are possible for evaluating the quality of service on the phase. Table 5.1 shows a list of various

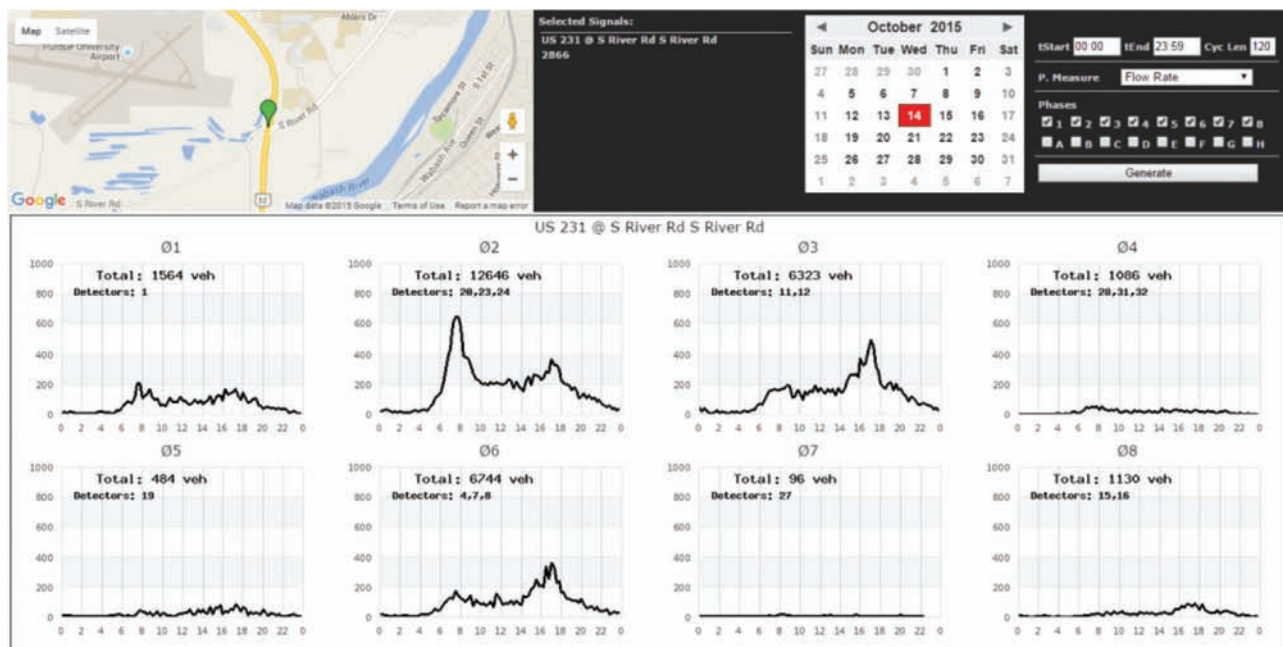


Figure 5.4 Flow rates for US 231 and River Road intersection.

TABLE 5.1
Performance measures for evaluating capacity utilization.

Performance Measure	Compatible Detection Types	Criteria for Poor Service	Note
Cycle Length, Red Time, Green Time	N/A	Inferred by high values.	Can serve as a very rough estimate of likely delay.
Volume-to-Capacity Ratio	Any (with count detection)	Indicated by high values, particularly those greater than saturation (> 1.0).	Requires assumptions about saturation flow rate and effective green time.
Rate of Force-Off / Max-Out	Any	Indicated by high recurrence of force-off / max out.	Sensitive to actuation settings. Not accurate for simultaneous gap phases. Cannot be used for phases in max recall.
Green Occupancy Ratio (GOR)	Stop Bar	Indicated by high values.	Sensitive to stop bar detection zone length. Tends to reach a high value relatively early. Utility may be limited by actuation behavior
Red Occupancy Ratio and Green Occupancy Ratio (ROR/GOR)	Stop Bar	Indicated by high green occupancy combined with high red occupancy.	Sensitive to stop bar detection zone length. Sensitive to choice of red duration interval length.
Time to Service	Stop Bar	Indicated by high values.	Sensitive to stop bar detection zone length.
Maximum Vehicle Delay	Stop Bar	Indicated by high values.	Sensitive to stop bar detection zone length.
Queue Service Time	Stop Bar	Indicated by high values.	Sensitive to stop bar detection zone length.
Average Delay	Stop Bar	Indicated by high values.	Requires assumptions about arrival profiles if they are not known.
Spatial and Temporal Oversaturation Severity Index	Advance	Indicated by high values.	Requires assumptions for determining unusable green time. Well suited for analysis where queuing and downstream blockage are factors.

performance measures that have been used to evaluate how well the capacity of an individual phase is being used.

5.4.1 Cycle Length

The first item in this list consists simply of phase times, without any vehicle information: the cycle length, and the amounts of red and green time that each movement experiences. Cycle length represents the amount of time needed for the signal to rotate through all of the phases for which there is demand; it can be directly programmed, as in traditional coordination, or it can simply result from the signal timing. That is, there is *always* an “actual” cycle length regardless of whether or not there is a “design” cycle length. Because this number represents the amount of time needed for each movement to receive green, it can serve as a very rough estimate of delay at the intersection. The longer the cycle length, the longer many vehicles will have to wait for service, especially those that arrive on minor movements.

Figure 5.5 shows an example plot of actual cycle lengths over 24 hours for two different days at the intersection of US 31 and 126th Street. During the overnight period (22:00–6:00), the cycle lengths vary considerably owing to actuation, with the longer cycles occurring as the signal dwells in green for the major through phases. During the rest of the day, the cycle length is constant within each time-of-day (TOD)

pattern, with occasional higher values occurring near the transitions from one TOD pattern to another.

Comparison of cycle lengths between different intersections can help ensure that all of the intersections in a system are changing their cycle lengths to the same value at the same time. A wrongly entered cycle length or pattern change time can lead to operational problems, as later examples will demonstrate.

5.4.2 Duration of Green and Reason for Termination

Similar information can be obtained from an analysis of green times. Many older central systems have had the ability to display statistical properties of the actual green times. The high-resolution data enables the cycle-by-cycle green times to be seen as well. Quite a bit of detailed information can be gleaned from this information, if needed.

Figure 5.6 shows a detailed plot of cycle-by-cycle green times for two different phases, the southbound (phase 5) and northbound (phase 5) left turns at SR 37 and SR 32 in Noblesville, Indiana. The graph examines 2 hours of operation spanning a TOD plan change. The plot reveals that phase 5 receives more green time than phase 1, and also that there is very little variation in phase 5. From this one can infer that phase 5 forces off or maxes out in each cycle. Only during transition and one early termination (gap out) does it differ. Phase 1,

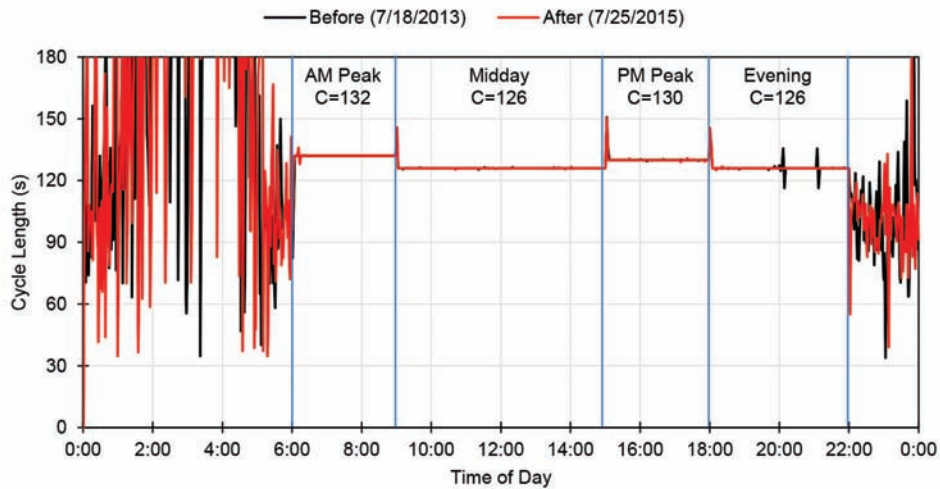


Figure 5.5 Cycle length, US 31 and 126th Street.

in contrast, ranges between 7 and 12 seconds depending on the number of vehicles, and is often skipped.

Figure 5.7 shows the same plot for eight different phases at US 31 and 126th Street before and after a change to the splits for the midday pattern. The impact of actuation is clearly visible in the varying duration of green time from one cycle to the next, and it is also possible to correlate the increases or decreases in the splits (as indicated) with movement in the green line plot between the before and after datasets. Another way of looking at the same data is by viewing cumulative green times, as in Figure 5.8.

Combining the duration of green with other information increases the utility of the metric. Figure 5.9 shows an example of the “split monitor” from the UDOT Signal Performance Measures website [76], which combines a plot of the total phase duration (green, yellow, and red clearance), the programmed splits, duration of concurrent pedestrian phases, and cause of phase termination (gap-out, max-out, or force-off). Summary statistics are also shown per TOD Plan at the top of each image, such as the average split duration, percent of cycles forcing off, and percent of cycles in which the phase is skipped.

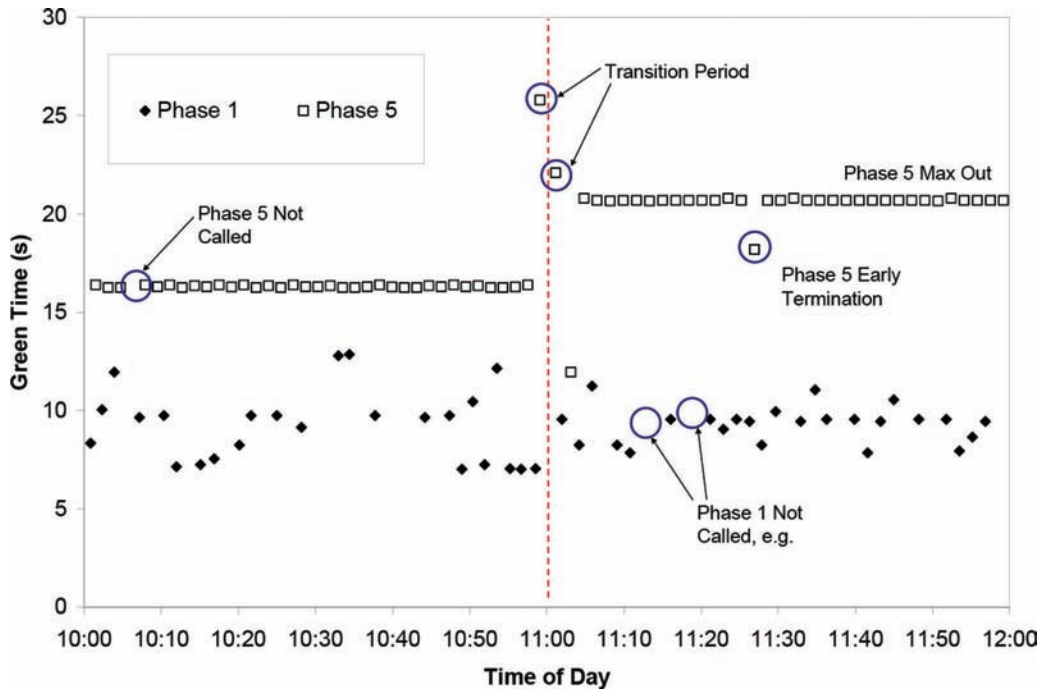


Figure 5.6 Artifacts of signal timing that can be observed from a close-up view of green duration [15]. Data shown for SR 37 and SR 32 in Noblesville, Indiana.

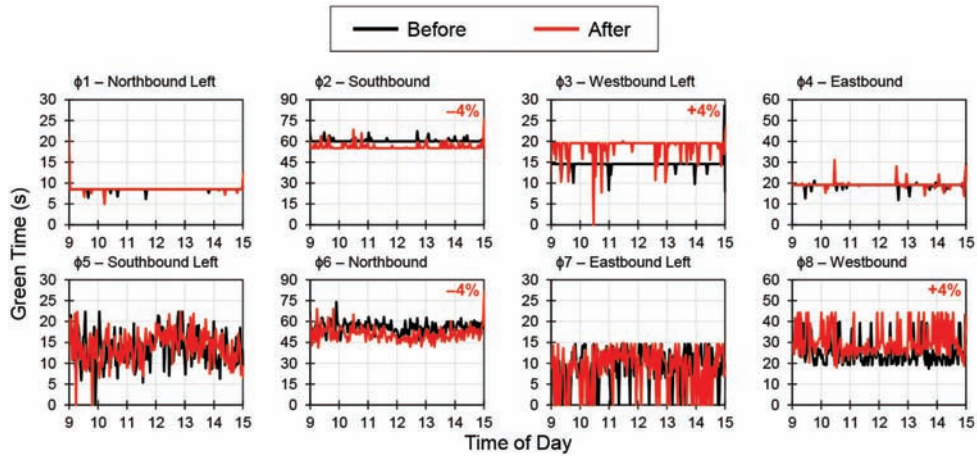


Figure 5.7 Cycle-by-cycle green times, US 31 and 126th Street.

Figure 5.9a (top) shows data for a coordinated phase; almost all of the phase durations exceed the programmed split, the phase forces off during every cycle, and there is also pedestrian activity throughout the entire day, suggesting that the pedestrian phase comes up automatically—either by recall or because of configuration of phase 2 as a CNA phase. (CNA, i.e., “call to non-actuated,” phases will, when also designated as coordinated phases, have their walk interval extended so as to fully utilize the portion of cycle where the coordinated vehicle phase is held in green.) Figure 5.9b (bottom) shows an actuated phase, which typically gaps out, except for about half of the cycles during the PM peak. Most of the time, the phase duration is less than the programmed split, but occasionally it is higher, which illustrates the impact of using fixed force-off for this phase.

The phase termination graphics can be summarized rather succinctly for all of the phases [20], as shown in Figure 5.10. These plots graph the number of force-offs and max-outs that occur per phase by time of day. Figure 5.10a shows the occurrences of all force-offs, whereas Figure 5.10b shows only cycles that ended in a third (or more) consecutive force-offs. Phases 2 and 6

exhibit a number of single force-offs but there are considerably fewer occurrences of three in a row. Phases 1, 3, 4, and 8, in contrast, are forced off during nearly every cycle in the analysis period.

The UDOT implementation calls this chart the “Purdue Phase Termination” graphic, and additionally layers the occurrence of pedestrian phases on it. Figure 5.10c shows an example showing individual force-off occurrences. Force-offs are indicated by blue dots, gap outs by green dots, and occurrences of pedestrian intervals by a yellow dot slightly above the vehicle phase symbol. What is particularly striking about this plot is the relatively low number of occurrences of the left-turn phases (1, 3, 5, and 7) throughout the day, which is likely due to detector configuration that avoids calling the protected left-turn phase until a substantial queue has built for those movements.

These data are relatively simple, yet can help the engineer seek an objective of providing signal timing that is consistent and predictable. The minimum and maximum values, their central tendency, and the degree of variation can all be ascertained from analysis of this type of data, verifying that things are operating as expected.

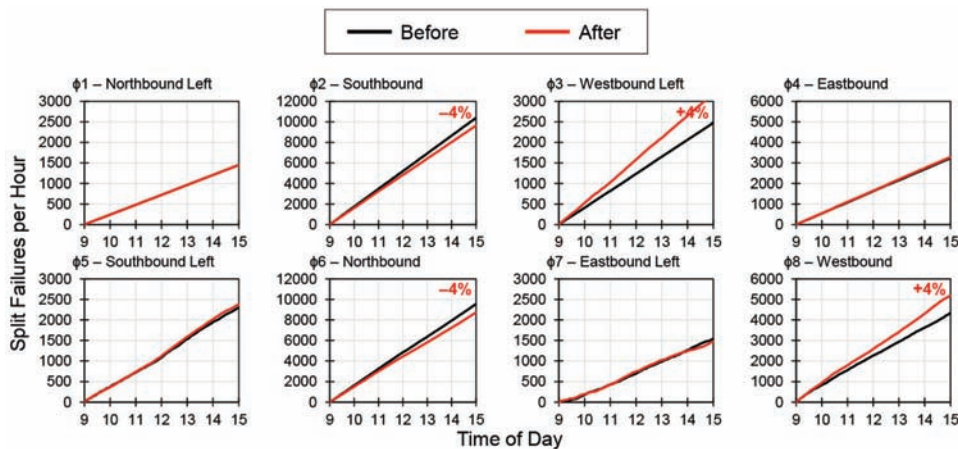


Figure 5.8 Changes in cumulative green time at US 31 and 126th Street, with changes to splits noted.

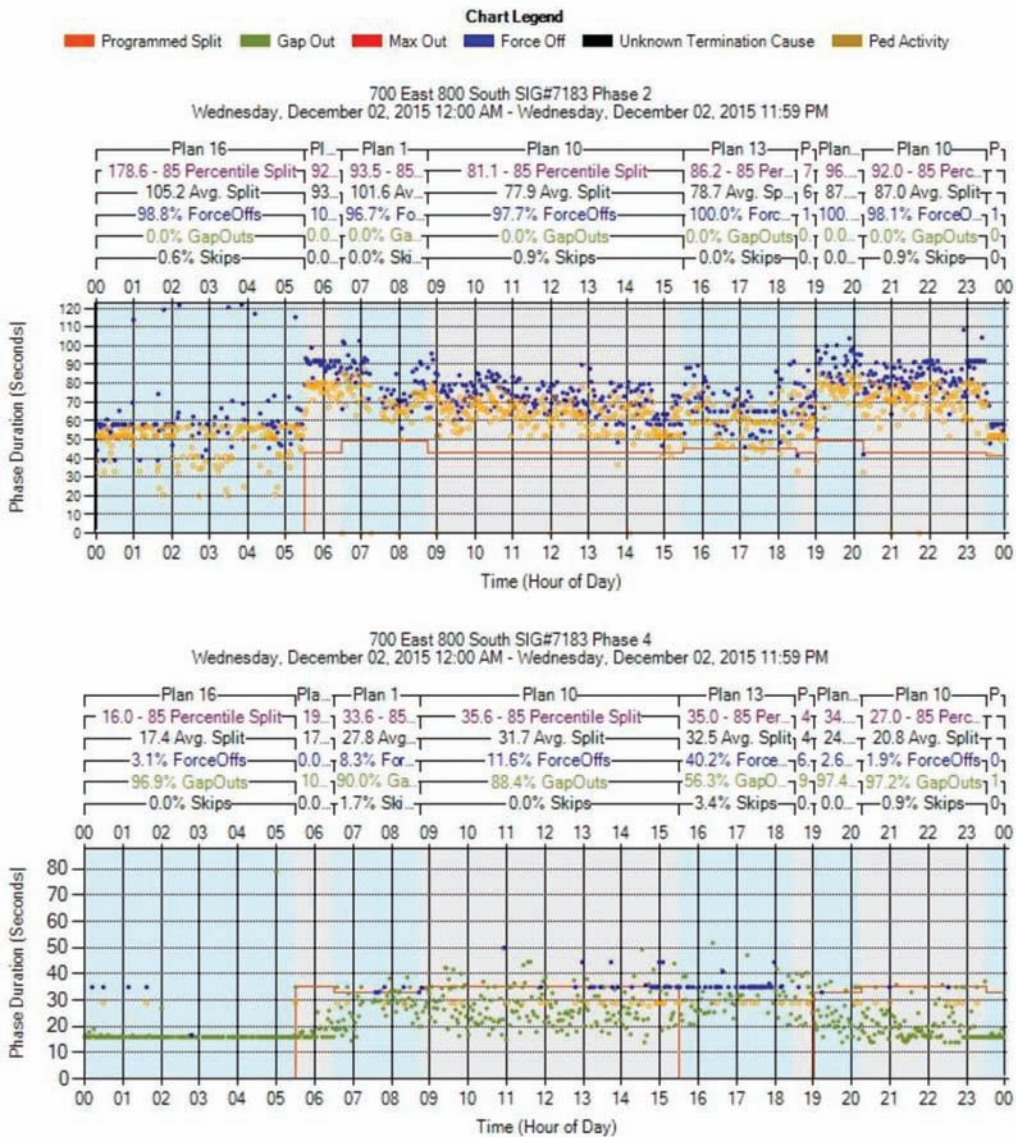


Figure 5.9 Split monitor plot from UDOT Signal Performance Measures website [76]. Data shown for 700 E at 800 S, 12/2/2015, for phases 2 and 4.

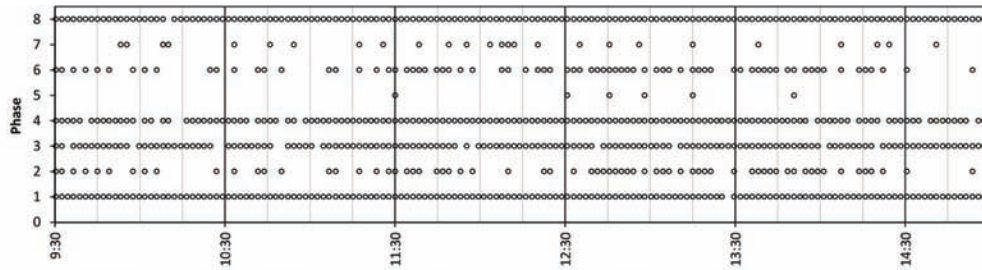
Note that, at this point, we have not yet considered the details of the intersections being discussed—a general knowledge of how dual-ring eight-phase timing works is sufficient to understand the dynamics.

Some inferences can be made about the quality of service, if it can be assumed that force-offs and max-outs correspond to high demand. The data show when an actuated phase is reaching its limits, such as when it is always serving at least the programmed split plus any additional time it can absorb from other phases. Similarly, excess green (or lack thereof) on coordinated phases is a good indicator of the overall intersection demand. Since those phases tend to absorb any green time given up by minor phases that end early, when the minor phases are using all their splits, the coordinated phases will shrink to their programmed splits.

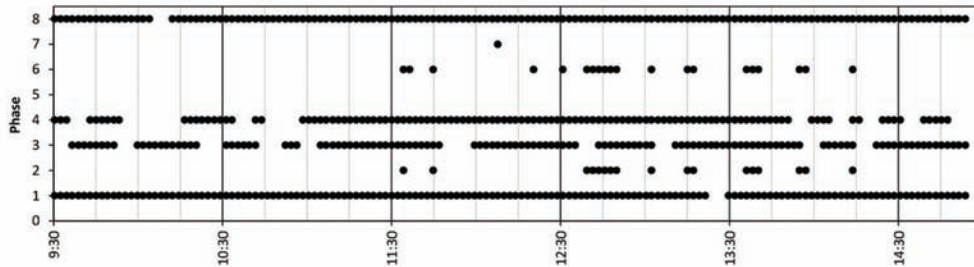
5.4.3 Capacity Utilization Metrics

Combining information from vehicle detection with phase durations allows a more thorough analysis of how much capacity is being used. As shown in Figure 5.2, depending on the type of detection (setback or stop bar), it is possible to measure either the individual vehicle arrivals or the overall service time.

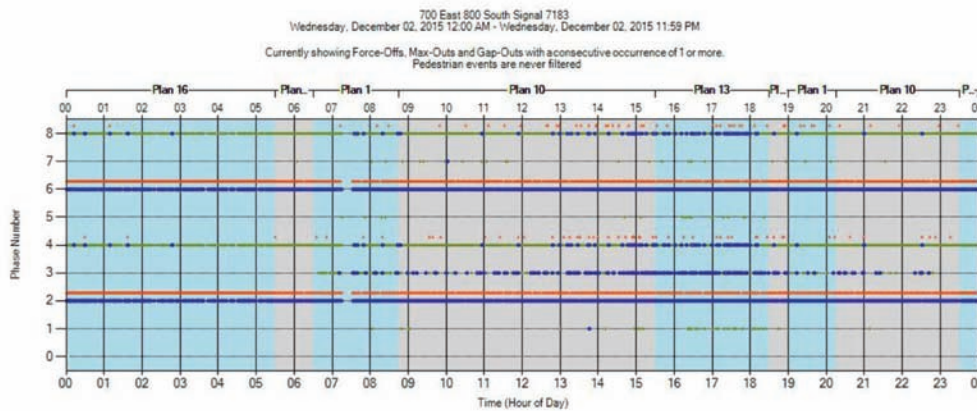
To illustrate alternative approaches, we will examine detailed information from an arterial intersection for which a split adjustment was made. This is the intersection of US 31 and 126th Street. Figure 5.11 shows a map of the intersection, showing the layout of detection zones and with a few specific detectors called out. These will serve as an example for demonstrating a way to evaluate the quality of service on a phase using stop bar detector data.



(a) Pilot implementation at US 31 and Carmel Drive, 6/18/2012: force-off occurrences by phase [20].



(b) Pilot implementation at US 31 and Carmel Drive, 6/18/2012: three consecutive force-offs by phase [20].



(c) UDOT implementation [76]. Data shown for 700 E at 800 S, 12/2/2015.

Figure 5.10 Purdue Phase Termination plots.

Figure 5.12 shows examples of phase events for the undersaturated eastbound left turn (Figure 5.12a) and the oversaturated eastbound through movement (Figure 5.12b). Each graphic includes two video still images that correspond to the conditions just after the beginning of green and several seconds after the beginning of red. Beneath this, the detector state and phase state are shown during the cycle represented in the video images, including the green occupancy ratio (GOR), which gives the percent of the green interval in which the detector was occupied; and the red occupancy ratio (ROR), which gives the percent occupancy of the first five seconds of red. The occupancy during the yellow interval is ignored.

- Phase 7, the westbound left turn, is undersaturated (Figure 5.12a). Shortly after the start of green, the two queued vehicles leave the detection zone, and shortly after this, the phase gaps out. No additional vehicles

arrive within the first 5 seconds of red. This cycle has a GOR of 67% and an ROR of 0%.

- Phase 4, the westbound through, is oversaturated (Figure 5.12b). At the beginning of green, there is a very long queue, which is able to advance forward, but after the end of green, there are still multiple vehicles remaining that were part of that queue. The red outline highlights the same vehicle in both images. The green interval was fully occupied (combining the occupancy of both lanes together), giving a GOR of 100%, whereas the first 5 seconds of red had 90% occupancy.

GOR alone tends to reach the ceiling of 100% somewhat before saturation [77]. It is possible for all of the queued demand to be served during a phase and for GOR to be rather high. Also, if the extension time is very small, there will be a natural tendency for the resulting GOR values also to be high. Combining this information with ROR measures both high utilization

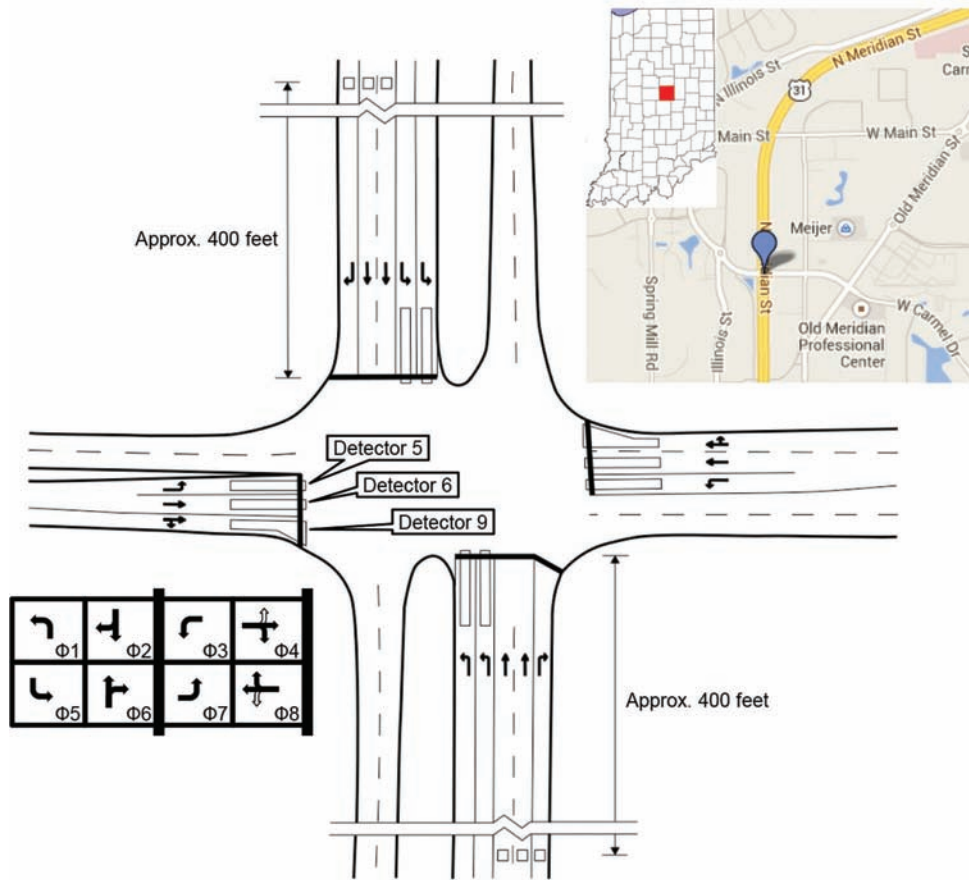


Figure 5.11 US 31 and 126th Street, Carmel, Indiana. [21]

of green and the presence of leftover demand beyond the end of green.

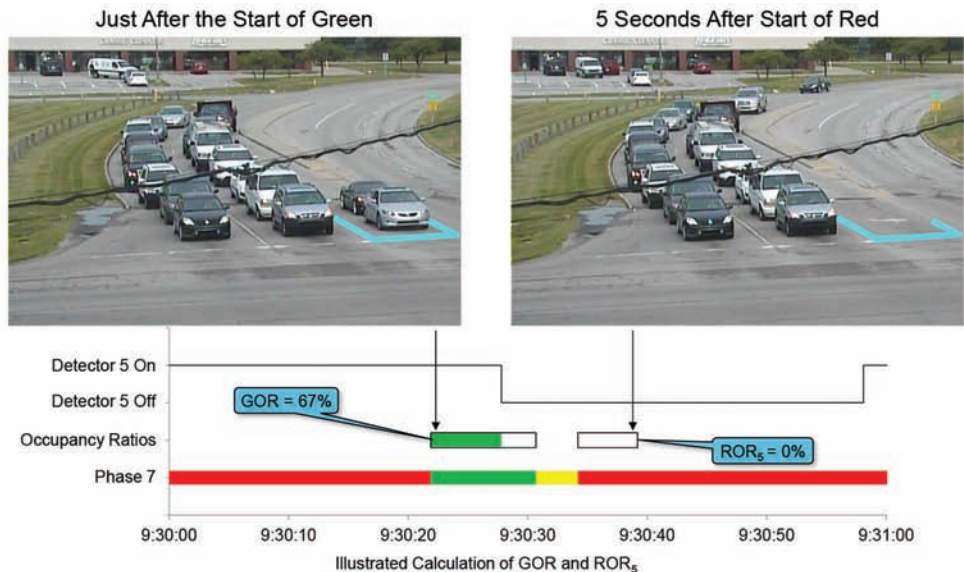
Figure 5.13 shows scatterplots of the cycle-by-cycle values of ROR versus GOR for 8 phases before (Figure 5.13a) and after (Figure 5.13b) the change to the splits at US 31 and 126th Street. Data are shown for two Thursdays for the 09:00–15:00 time period. Phases 2 and 6, which do not have stop bar detection, are excluded. Each chart contains a red box in the upper left-hand corner corresponding to $ROR > 0.8$ and $GOR > 0.8$, along with a number indicating the number of cycles having values within those ranges. These represent cycles that are likely split failures, having had their green interval completely utilized, and with demand remaining after the end of green. Therefore, the count of cycles with ROR and GOR both greater than 0.8 can be taken as the number of split failures. Phases 3 and 8, which received additional green time because of split increases, both saw large reductions in the number of split failures. The other phases did not see any substantial changes.

The two phases from which time was taken, phases 2 and 6, have setback detection rather than stop bar detection. This is a typical INDOT configuration. For these phases, we can use the volume-to-capacity (v/c) ratio based on the counts from the setback detectors [21].

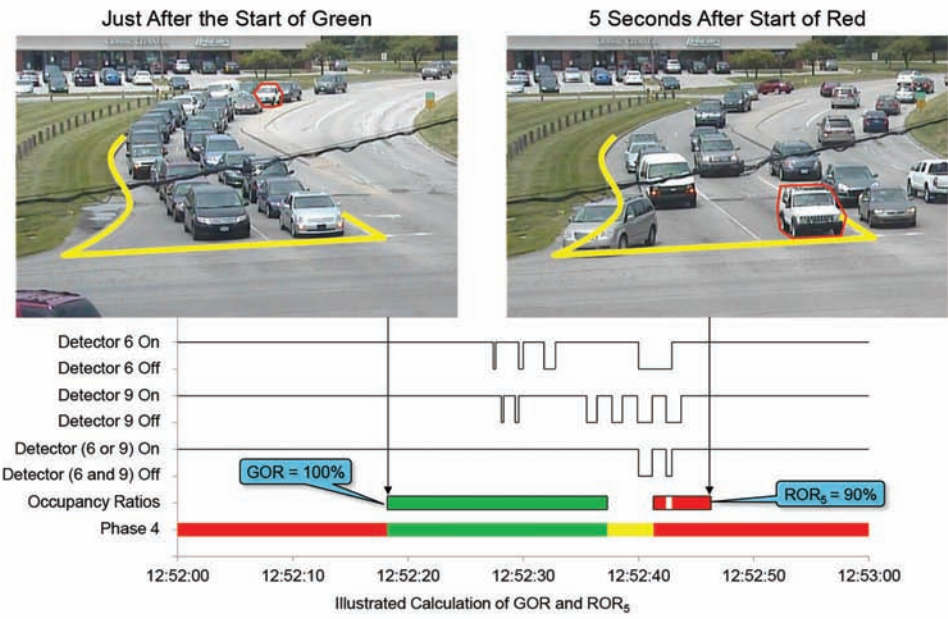
Figure 5.14 shows 24-hour plots of the cycle-by-cycle values of v/c ratio and 10-point moving averages for southbound phase 2 (Figure 5.14a) and northbound phase 6 (Figure 5.14b). A 10-cycle moving average is also shown to indicate the central tendency of the point cloud. For the midday period, both phases see a slight upward trend in the v/c ratio during the midday period, corresponding to the split decreases.

Another way of viewing the same data is with a cumulative frequency diagram of the v/c ratio, as shown in Figure 5.15. In both cases, the “after” curve makes a rightward shift from the “before” curve, indicating an increased value. The cycle-by-cycle data provide a basis for statistically comparing the two datasets; in this case, a T-test finds that the differences in means between the “before” and “after” sets are statistically significant for both the northbound and southbound v/c ratios.

Values of v/c ratio above 1.00, indicating higher volume than the capacity (assumed in this case to be 1900 vehicles per hour per lane), are indicative of split failures [77]. These numbers can be combined with the split failures determined using ROR/GOR to get an overall picture of the intersection performance, as shown by Figure 5.16. This is a bar chart showing split failures per hour, using two series to represent before and after conditions. The reduction of splits on phases 2



(a) Undersaturated phase.



(b) Oversaturated phase.

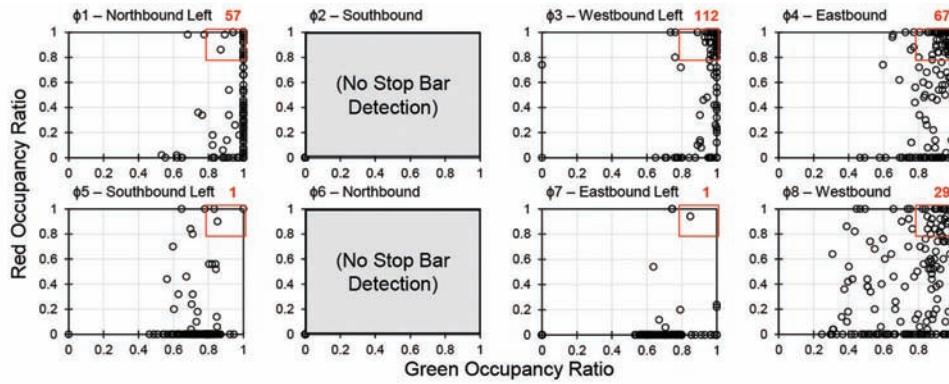
Figure 5.12 Illustration of the calculation of ROR and GOR [21].

and 6 led to no change for phase 2—there were no split failures in either case—and an increase for phase 6, mostly in the 15:00–16:00 hour. This illustrates the tradeoff inherent in the split adjustment, and provides a means of deciding whether a handful of additional split failures on phase 6 are worth a drastic reduction on the other phases.

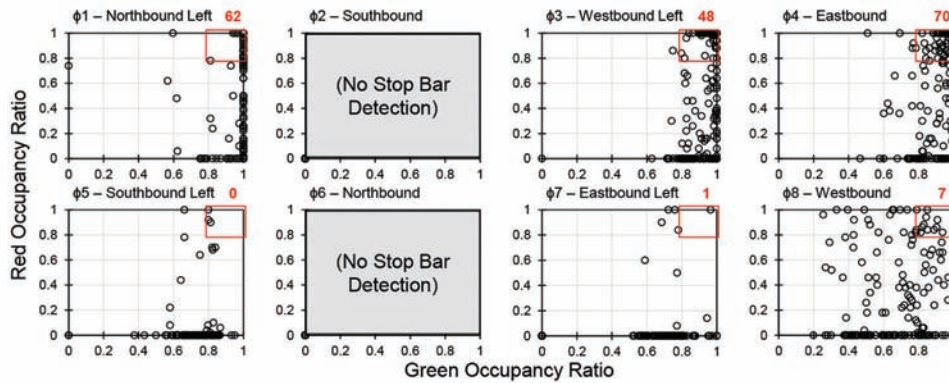
These results are for a pair of days before and after the change, but it is also desirable to know the impact throughout different days of the week. Figure 5.17 presents the same information as Figure 5.16, this time showing the counts of split failures per phase in before/after pairs by day of week, with the labels indicating the

total number of split failures for the intersection in each case. Each day of the week saw a reduction in that number, and overall the trends are much the same, with phases 3 and 8 having a visible decrease and phase 6 having a slight increase. Interestingly, Thursday has the highest number of split failures of all the days of week considered, rather than Friday as one might expect to be a slightly busier day.

Yet one more way to view split failures is by considering their severity. A single split failure that leaves behind demand, yet is corrected in the next cycle, poses less of a problem than a series of consecutive split failures indicating that the phase is not recovering over



(a) Before split changes: Thursday, 7/18/2013, 0900-1500.



(b) After split changes: Thursday, 7/25/2013, 0900-1500.

Figure 5.13 Red Occupancy Ratio (ROR) versus Green Occupancy Ratio (GOR) at US 31 and 126th Street. The numbers indicate the cycles within the red region where ROR and GOR are both greater than 0.8.

a long time period. This can be visualized using the cumulative number of split failures per cycle. This is a counter that increments during each split failure that follows one in the previous cycle, but resets to zero as soon as one cycle passes in which there was *not* a split failure, which indicates that demand was cleared at that time.

Figure 5.18 shows a plot of cumulative split failures for phase 3, which saw the most dramatic reduction in the split failure count. The “before” data show several instances where there were a number of split failures in a row. At approximately 12:10, for example, there had been 19 cycles in a row in which the phase repeatedly failed. One cycle after this saw demand clear, but this was followed by an additional 8 cycles that failed. There are many occasional breaks, but more often than not the phase failed. The “after” data still exhibits a number of failures, but there are rarely more than 2 in a row. The area under the curve in each case provides an index that gives a summary quantitative value to this metric.

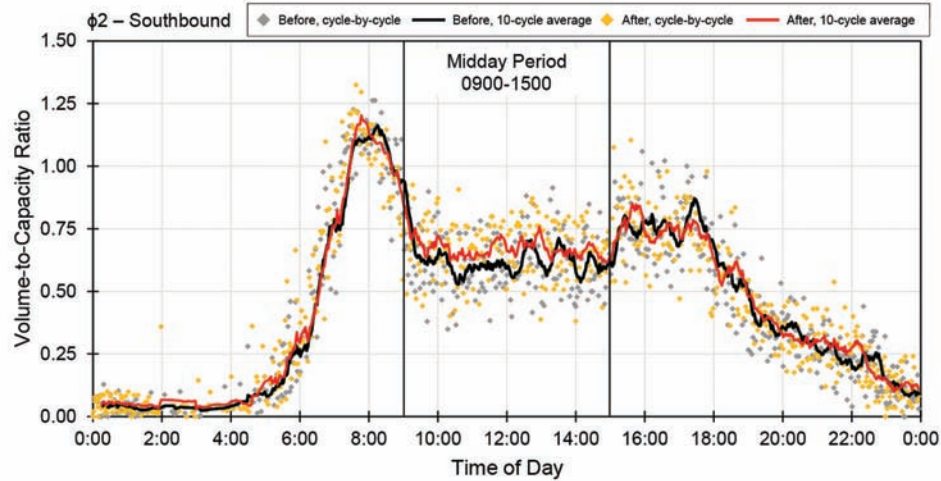
Figure 5.19 expands this to an eight-phase view, which gives some perspective to the overall intersection condition. The number in the corner of each chart shows the area under the curve before and after the split change. Phases 3 and 8 both saw substantial decreases. While phase 6 saw an increase, there was never more than one cycle in a row having a failure, which adds

some additional information about the impact on that phase. Phases 1 and 4 both saw increases. Since the splits of those movements did not change, it is more likely that there were transient spikes in demand in the “after” analysis periods on those dates. Phase 4 in particular seems to be a candidate for additional green time, especially in the 12:00–13:00 hour.

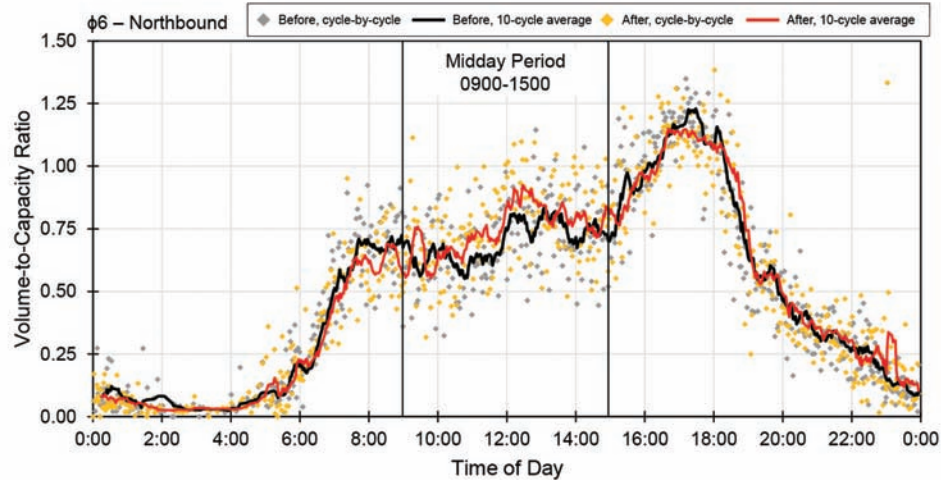
These views help facilitate the objective of providing efficient operation by enabling split failures to be indicated for competing movements. The impact of an action such as rebalancing splits can be directly determined by watching how such metrics move when comparing before and after conditions. Lastly, the data themselves can suggest potential amendments to the distribution of green times by indicating which movements need more time and which can give it up.

5.4.4 Estimated Delay

Vehicle delay is a commonly used metric for evaluating signal timing. The Highway Capacity Manual uses it as the basis of the Level of Service on movements, approaches, and intersections. However, it tends to be costly to measure, as it requires the tracking of vehicles from their approach to the intersection through to their departure from it, or matching of vehicle observations between the two points to obtain a



(a) Southbound phase 2: Thursday, 7/18/2013 versus Thursday, 7/25/2013.



(b) Northbound phase 6: Thursday, 7/18/2013 versus Thursday, 7/25/2013.

Figure 5.14 Cycle-by-cycle volume-to-capacity ratios for mainline through phases 2 and 6 at US 31 and 126th Street.

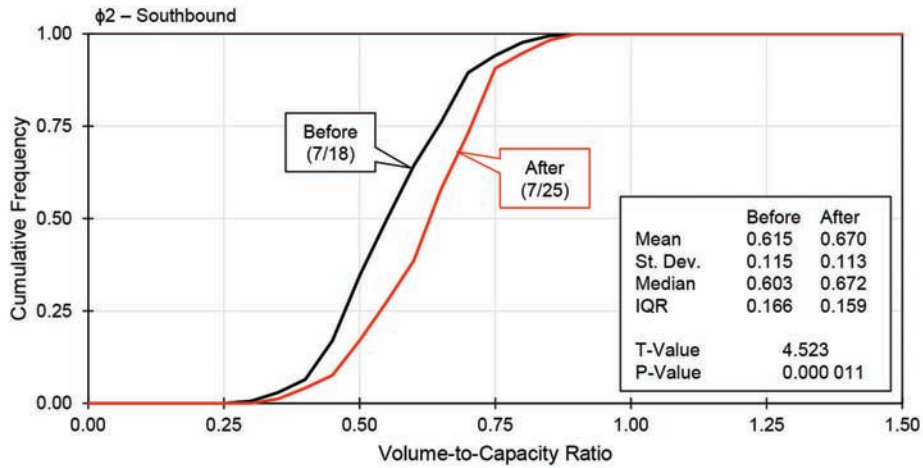
travel time, from which delay can be found. These are still difficult datasets to obtain, although one day connected vehicle data may make such measurements possible.

It is still possible to make *estimates* of the delay based on simple bulk queuing theory based on arrivals and departures. Depending on the type of detection that is available, two types of delay estimates can be made:

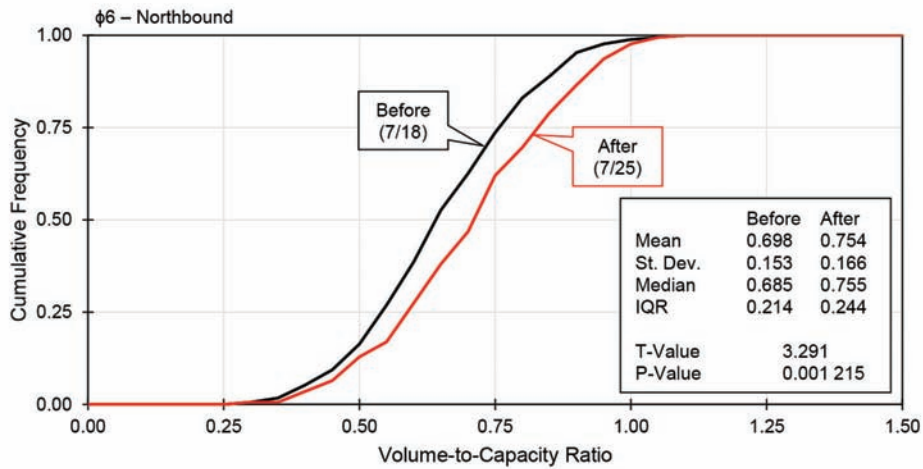
For advance detectors, the arrival profiles can be measured directly from the detector actuations while the green time represents when departures take place. The combination of these two enables a queue polygon to be constructed; this method is called “input–output” delay [78]. Figure 5.20 shows an example of the construction of a queue polygon based on this method. The blue vertical lines indicate vehicle arrivals. As these occur during red, the queue accumulates. After the beginning of green, the queue starts to shrink as vehicles are estimated to depart from the front of the queue, until it reaches zero and the process begins again

in the next cycle. The area under the curve is the total delay. Figure 5.21 shows a view of the approach delay from the UDOT Signal Performance Measures website, which shows both the total delay (delay per hour) and the average delay (delay per vehicle). There are AM and PM peaking characteristics as one might expect.

For stop bar detectors, the arrival profile is not known. Only the arrival time of the first vehicle can be measured, since it rests in the detection zone between its arrival and departure time. In the meanwhile, other vehicles arrive behind it. However, the delay of the first vehicle provides a means of knowing the base of a triangular queue polygon, and the total count of vehicles that are served during the cycle gives its height. Assuming random arrivals, these can be distributed throughout the cycle to construct arrival and departure curves. Figure 5.22 shows how this is done; the method is described in detail elsewhere [2,78], but is summarized as follows. The arrival times consist of the measured first arrival (A_1) and randomly distributed arrival times



(a) Southbound phase 2: Thursday, 7/18/2013 versus Thursday, 7/25/2013.



(b) Northbound phase 6: Thursday, 7/18/2013 versus Thursday, 7/25/2013.

Figure 5.15 Cumulative frequency diagrams of volume-to-capacity ratio for mainline through phases 2 and 6 at US 31 and 126th Street.

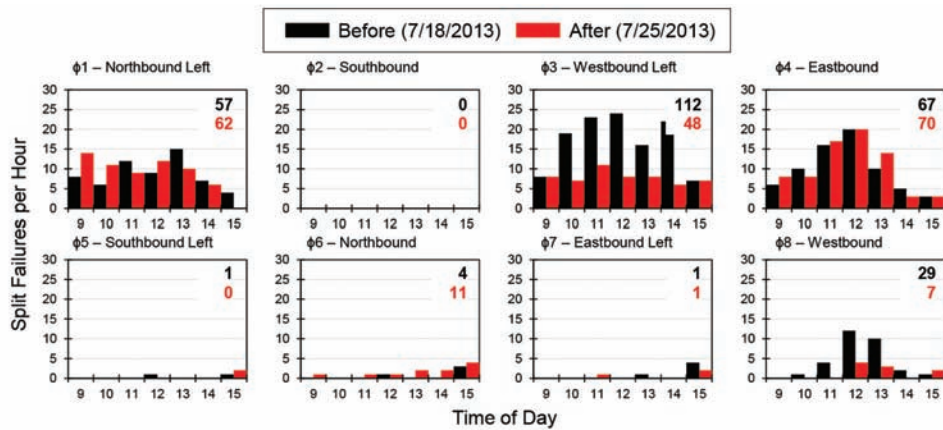


Figure 5.16 Number of split failures before and after split adjustment at US 31 and 126th Street: Thursday, 7/18/2013, versus Thursday, 7/25/2013.

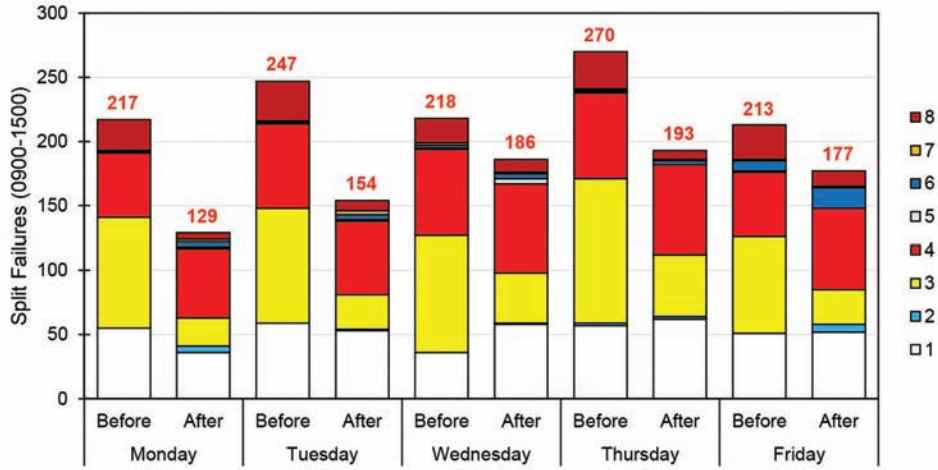


Figure 5.17 Split failures by phase, by day of week, before and after split adjustment at US 31 and 126th Street: week of 7/15/2013, versus week of 7/22/2013.

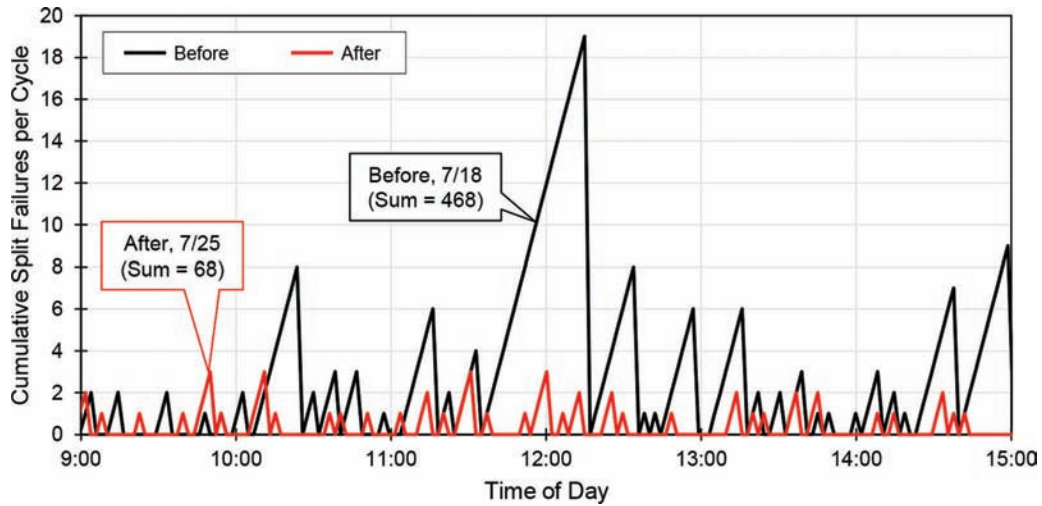


Figure 5.18 Cumulative split failures on phase 3 at US 31 and 126th Street, before and after a split adjustment.

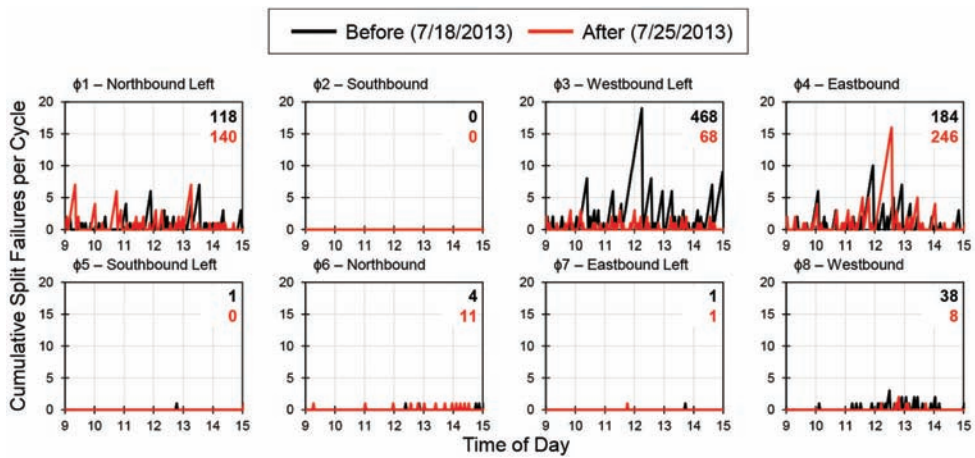


Figure 5.19 Cumulative split failures on 8 phases at US 31 and 126th Street, before and after a split adjustment.

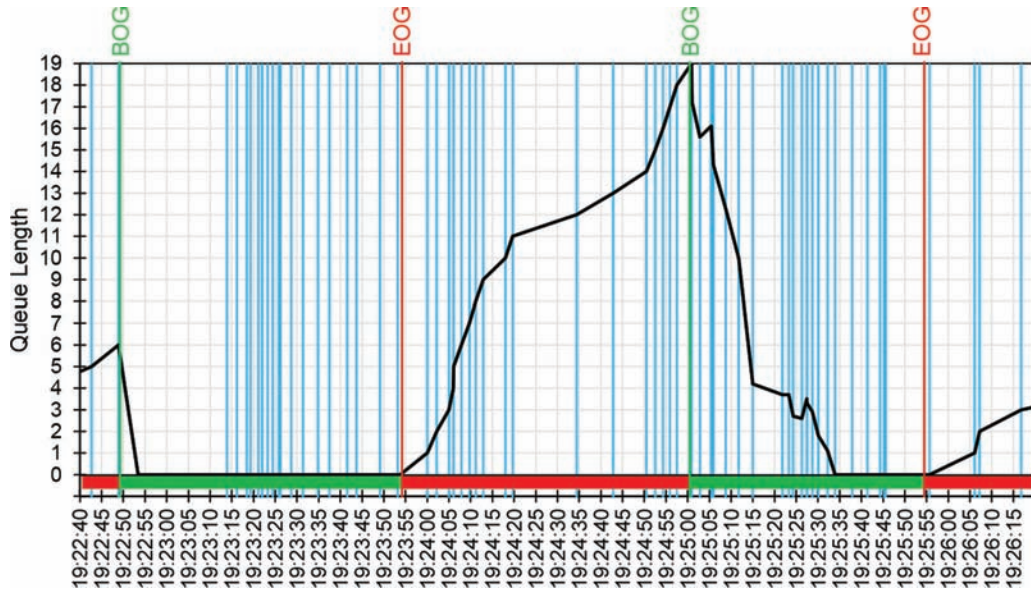


Figure 5.20 Input-output delay (advance detector).

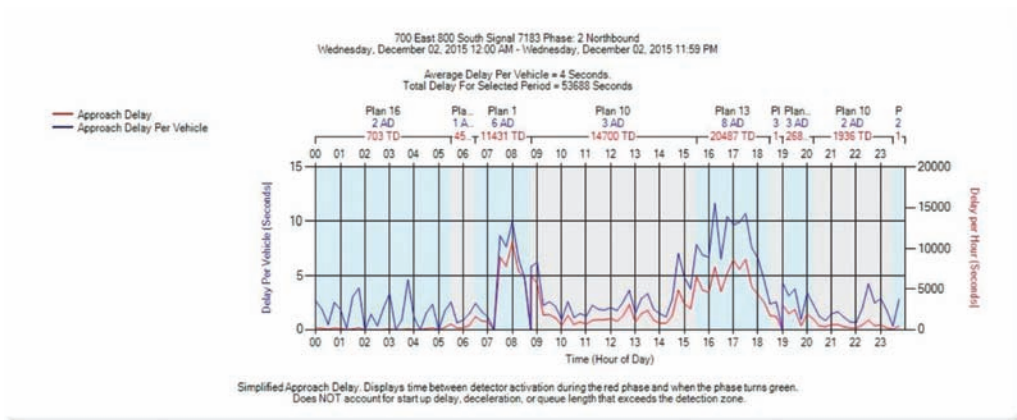


Figure 5.21 Approach Delay plot from UDOT Signal Performance Measures website [76]. Data shown for 700 E at 800 S, 12/2/2015.

occurring after that, up to the end of green. The departure times are assumed to occur after the start of green with headways determined from the saturation flow rate. This process yields another queue polygon, the area of which is equal to the total delay.

Figure 5.23 shows a bar chart of the total delay per intersection that compares free versus coordinated operation on a 5-intersection corridor [75]. Figure 5.23a presents the total delay for the mainline phases (2 and 6) only, and Figure 5.23b shows the total delay for the minor phases only. The implementation of a coordination plan yields reductions in the total delay for all the coordinated phases (the results for State Street are limited to the southbound direction because the northbound detectors were not working properly at the time), while it increases the total delay on all of the side street phases. This is another example of a direct tradeoff between alternative objectives, with implemen-

tation of coordination judging that the decrease in coordinated phase delay is worth the increase in side street phase delay.

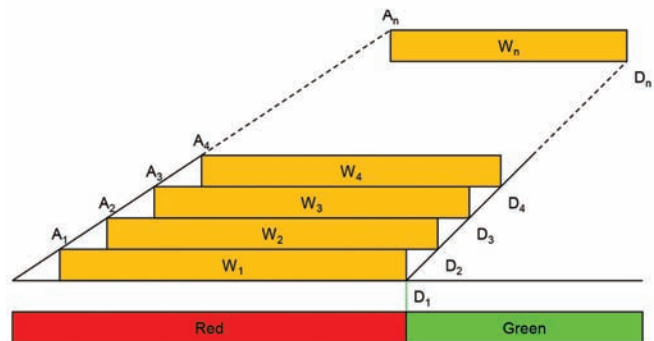


Figure 5.22 Random-arrival delay (stop bar detector).

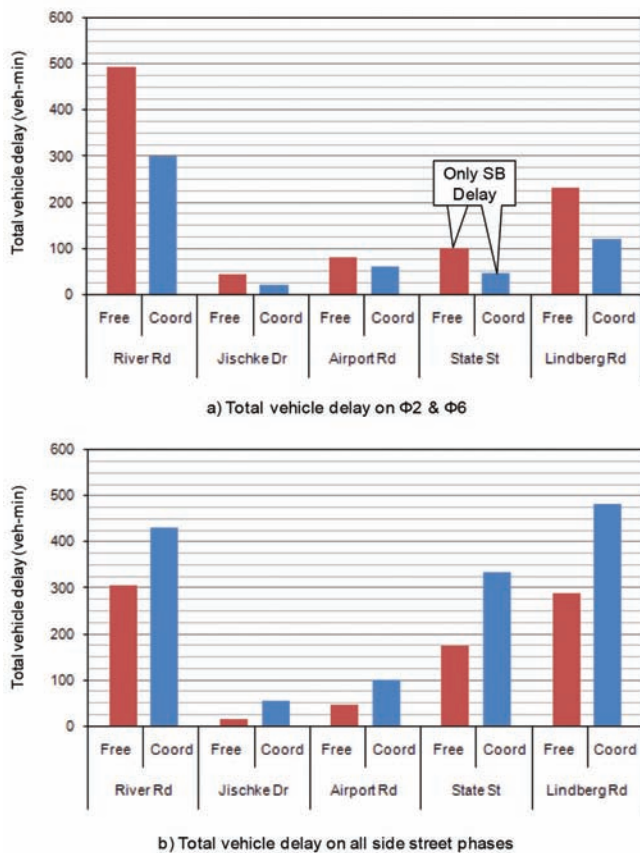


Figure 5.23 Analysis of total delays for free versus coordinated operation for US 231 during the AM Peak (06:00–09:00) during timing plan design in 2014 [75].

Dividing the total delay in a cycle by the number of vehicles served in that cycle yields the average delay. Figure 5.24 shows cumulative frequency diagrams of the average delay for US 31 and 126th Street before and after the split change. Each series contains five traces for five representative weekdays for the before and after

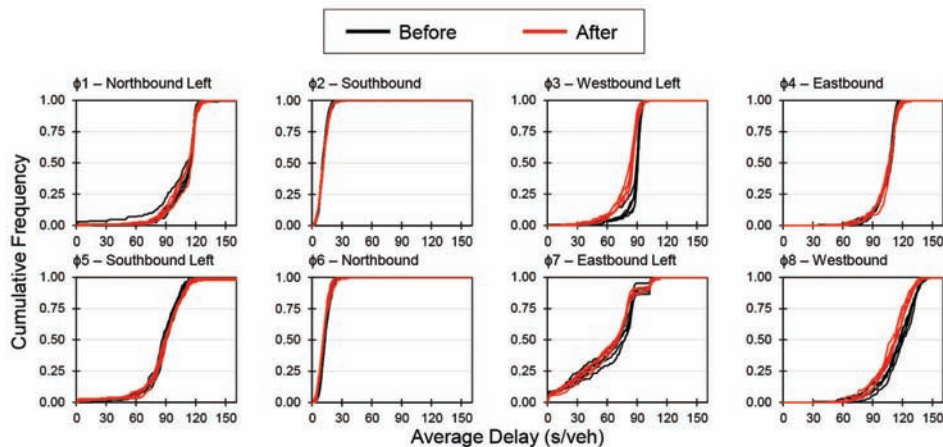


Figure 5.24 Cumulative frequency diagrams of average delays per cycle for 5 “before” and 5 “after” days at US 31 and 126th Street, before and after a split change.

periods. Phases 3 and 8, which received split time, both see movement of the average delay CFDs to the left, indicating decreased delay; the other movements do not exhibit any change at all, including phases 2 and 6, which gave up split time.

While delay can measure efficiency to some degree—the less, the better—the objective that is perhaps best characterized with this metric is that of equitability. While it is difficult to provide the same amount of delay time to all users on the approach to an intersection, it is often possible to provide what could be called a “fair” amount of delay time, in accordance with agency policy. The tradeoff between mainline and side street phases is the quintessential decision point for signal coordination, and for the choice of parameters such as the cycle length. Assistance by delay estimates from the actual vehicle performance would better assess the outcomes of those decisions.

5.4.5 Red Light Running

The performance measures examined thus far have focused on mobility. Safety is perhaps an even more important aspect of signal performance. Occurrences of red light running (RLR) are especially problematic because of the potential severity of RLR crashes, particularly at intersections with high-speed roads. To date, most effort in applying technology to this problem has concentrated on enforcement tools such as red light cameras. These systems are often only deployed at limited locations; despite reported crash reductions, the use of red light cameras is rather controversial, and not every agency is authorized to use them.

Another possibility is to use existing detector data to *proactively* evaluate RLR occurrences to locate areas where actions might be done to prevent future crashes. In Utah, radar detection on intersection approaches are used to determine vehicles passing through a detection zone without stopping after the start of red, and their

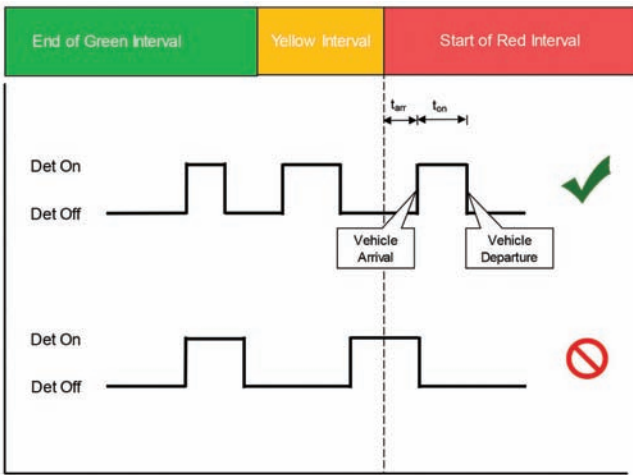


Figure 5.25 Conceptual overview of RLR detection using loop detector and phase event data [22].

measured speed is used to categorize whether the vehicle ran the red light.

A slightly more general analysis method can develop the RLR identification logic for other detector types. Taking

inspiration from the previous work by UDOT, a method using stop bar detection to measure RLR was recently developed as part of this project [22]. The logic is explained in Figure 5.25, which contains two detector on-off traces along with the concurrent phase state. The upper graph shows a trace in the detector presence that includes an off-on transition after the start of red, and an on-off transition shortly after this while the light is still red. This represents a likely RLR vehicle. The lower graph shows a detector presence trace that has an off-on transition *prior* to the start of red. This is excluded, as it is more likely to represent a vehicle entering during yellow.

Figure 5.26 presents an example of the RLR identification method using video stills and the trace of the detector and phase events as they occur at the intersection. The detection zone is shown by the red circles in the figure (callout *i*). The first image (Figure 5.26a) shows the unoccupied detector; the detector then turns on as a vehicle passes over it (Figure 5.26b, callout *ii*) and back off again when the vehicle clears the detection zone (Figure 5.26c). At the start of yellow (Figure 5.26d), we can observe a vehicle in the opposing left turn lane getting ready to execute the

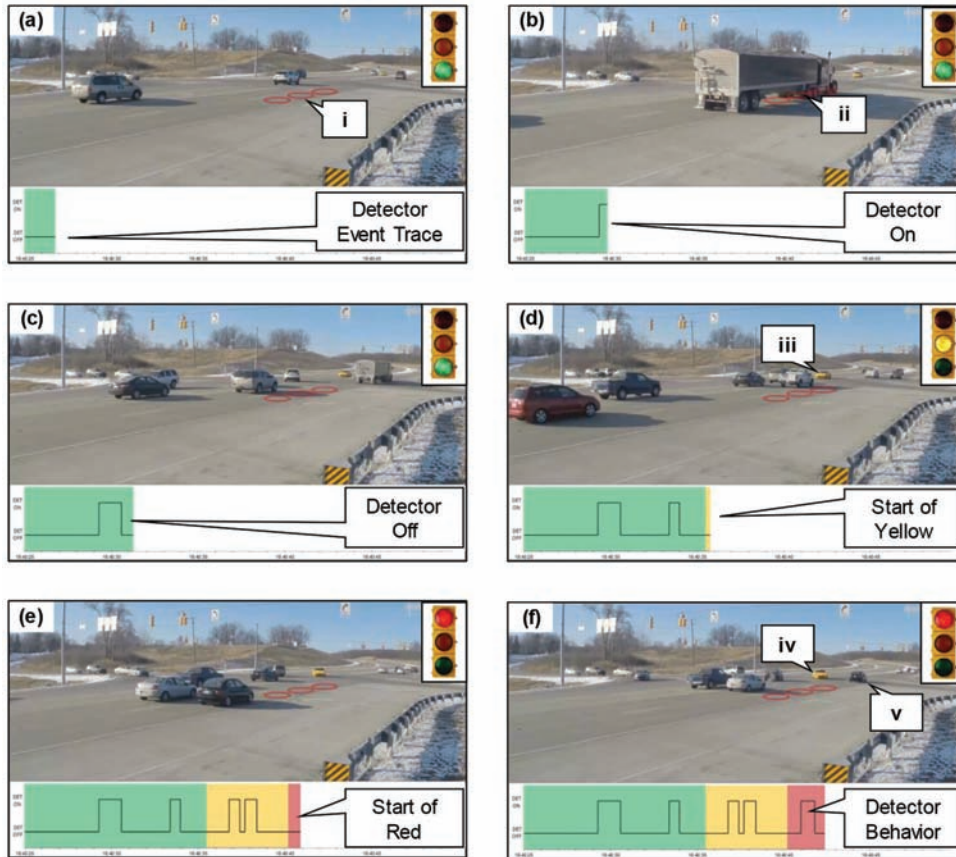


Figure 5.26 Example video validation of RLR data at US 231 and State Street (facing northbound US 231) [22]. (a) Normal green progression, detector zone unoccupied (callout *i*). (b) Detector zone occupied (callout *ii*). Note the corresponding “detector on” event from the high-resolution data. (c) Detector zone unoccupied, with corresponding “detector-off” event. (d) Start of yellow. Note the vehicle waiting to turn left on southbound US 231 (callout *iii*). (e) Start of red. (f) The vehicle in the near through lane (callout *v*) was flagged as a RLR in the data. Note the near-miss with the left-turning vehicle (callout *iv*).

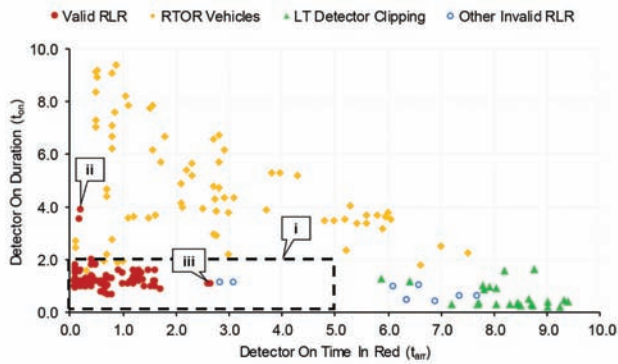


Figure 5.27 Comparison of RLR data events and video validation, along with safety risks for valid RLRs [22]. Callout *i* shows the refined search area for RLR events, which captures all valid RLRs, except for two large semi-trucks (callout *ii*). Callout *iii* identifies particularly high-risk RLR events, which occurred nearly 3 s after the start of red.

turn on yellow (callout *iii*). However, it is prevented from doing so by the approach of an oncoming vehicle (Figure 5.26e). Finally, Figure 5.26f shows the condition a few seconds afterward, with the left turn vehicle still waiting to proceed (callout *iv*), while the RLR vehicle has just cleared the intersection (callout *v*). This is a near-miss; had the left-turn vehicle been less attentive, it might have easily become a collision.

Returning briefly to Figure 5.25, note the two time intervals t_{arr} and t_{on} . These are respectively the time between the start of red to the vehicle arrival and the duration of the detector on time. Not all of the detector on-off events that occur during red will be RLR vehicles. The detector could see similar events occurring because of right turns on red (RTOR), “clipping” of the detection zone by left-turning vehicles from the crossing street, and other events such as detector chatter or lane incursions. However, video validation of the RLR performance measure considering t_{arr} and t_{on} found that a simple filter will suffice as a means of separating valid and invalid RLR detections. Figure 5.27 shows the results of the validation procedure; in Figure 5.27a, each detected RLR vehicle is displayed on a plot of t_{on} versus t_{arr} . Each of these has been identified as a valid or invalid RLR using a video recording of the site. Threshold values (callout *i*) of t_{on} and t_{arr} were found to be sufficient to isolate most of the valid RLR vehicles. The only missed vehicles were large two semi-trucks that had a longer t_{on} value owing to their greater length (callout *ii*). It is also worth noting the two particularly high-risk RLR events at callout *iii* that occurred nearly 3 s after the start of red.

Besides viewing RLR events at a particular movement, the method can be applied to archived data to evaluate whether changes in signal timing can affect the safety performance. We will return one last time to the example of the split adjustment at US 31 and 126th Street to see whether there were any such observable effects from that change.

The outcomes of the RLR analysis over 3 months before and after the split change are shown in Figure 5.28. The date of the split change was 7/29/2013. Figure 5.28a shows the daily RLR count over time on phase 8 at US 31 and 126th Street, and Figure 5.28b shows the same data in terms of the rate of RLR per thousand entering vehicles. In both plots, a reduction in RLR vehicles is visibly evident. From this data, the increase of the split by 4% of the cycle corresponds with an average decrease in RLR counts by 34% [22].

Table 5.2 shows results of an odds ratio test comparing the odds of RLR for the two different splits. This indicates that with the 20% split condition, a vehicle is about 1.6 times as likely to run a red light as with the 24% split condition. Estimation of a 99% confidence interval finds that the lower bound of the range is well above 1. A value of 1 would indicate no difference in the odds of between the two splits. From this, we can conclude that the split increase is associated with a significant reduction of RLR activity, for the time periods evaluated.

5.4.6 Pedestrian Utilization

So far, this discussion has focused on the vehicle mode. Vehicle demands are relatively easy to measure using detection. Pedestrian demand is more challenging

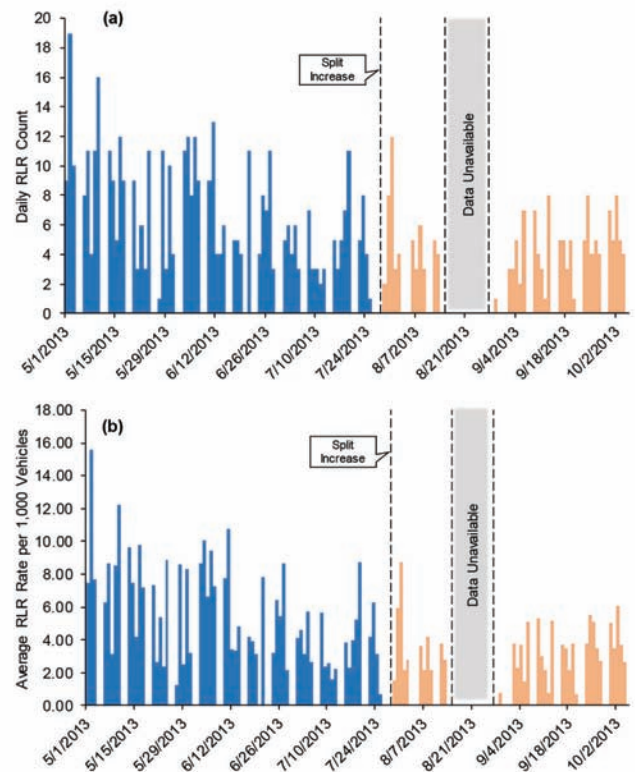


Figure 5.28 Effect of phase 8 split increase on red light running (RLR) vehicles during the 09:00–15:00 TOD plan at US 31 and 126th Street. (a) Change in daily RLR counts before/after split increase; (b) Change in average daily RLRs per 1,000 entering vehicles before/after split increase [22].

TABLE 5.2
Odds ratio test for RLR rates per 1,000 entering vehicles before/ after split adjustment [22].

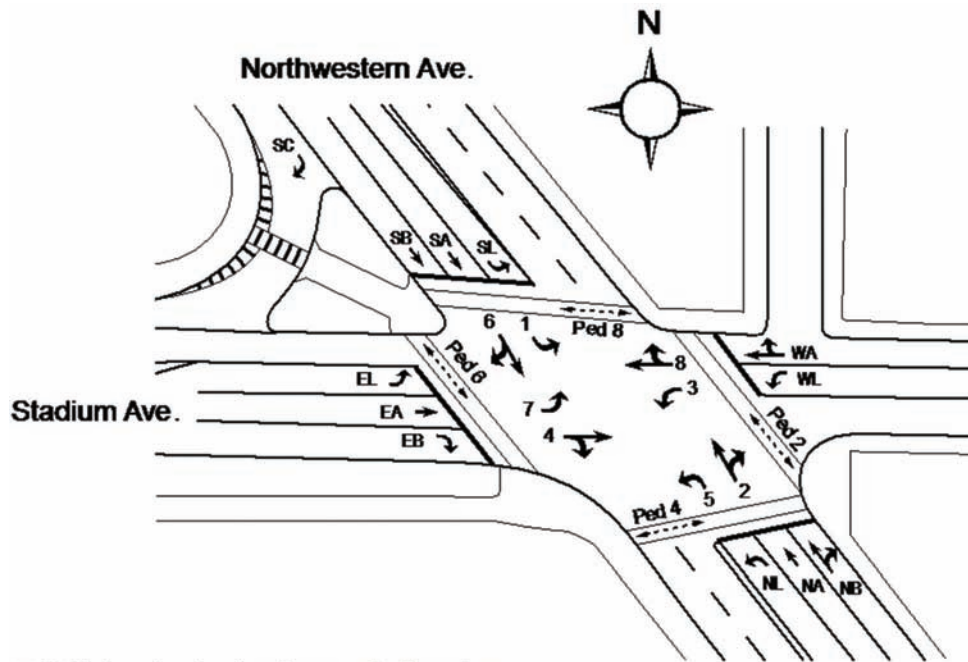
	Odds Ratio	
	20% Split	24% Split
RLR Odds	0.006	0.003
Odds Ratio	1.672	
Upper 99% CI	1.707	
Lower 99% CI	1.637	

to characterize, and automatic pedestrian detection is not yet common. However, at intersections where Ped buttons are used, the actuation of the pedestrian phase provides an opportunity to monitor its usage, which can serve as a proxy measure for the pedestrian demand [79].

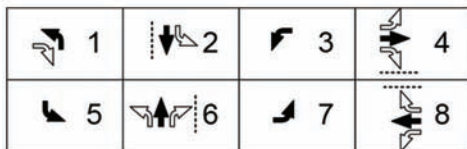
Figure 5.29a shows a map of the intersection of Northwestern Avenue and Stadium Drive in West Lafayette, Indiana. Two ring diagrams showing signal operation

with alternative pedestrian phasing options are also shown. Figure 5.29b shows the signal phasing before 2009, which has conventional pedestrian phases that are parallel to traffic. In March 2009, an exclusive pedestrian phase was added to the signal timing scheme to mitigate conflicts between the eastbound right turn and pedestrian traffic crossing Northwestern Avenue. Previous research had found a high number of conflicting vehicles crossing the phase 4 pedestrian movement, as Figure 5.30 shows [80]. This shows the equivalent flow rate of vehicles making the eastbound right turn in cycles when pedestrian phase 4 was called. Three different dates around the beginning of fall semester on campus are included. The data shows that there are many cycles within each day where conflicting flow rates are very high for the single turning lane, meaning that pedestrians are very likely being cut off by turning traffic at times.

Figure 5.29c shows the signal phasing after 2009, with an exclusive pedestrian phase (phase 13). Somewhat



(a) Vehicle and pedestrian phase numbering scheme.



(b) Ring diagram with conventional adjacent pedestrian phases.



(c) Ring diagram with exclusive pedestrian phase.

Figure 5.29 Intersection phase configuration at Northwestern and Stadium, West Lafayette, Indiana [79].

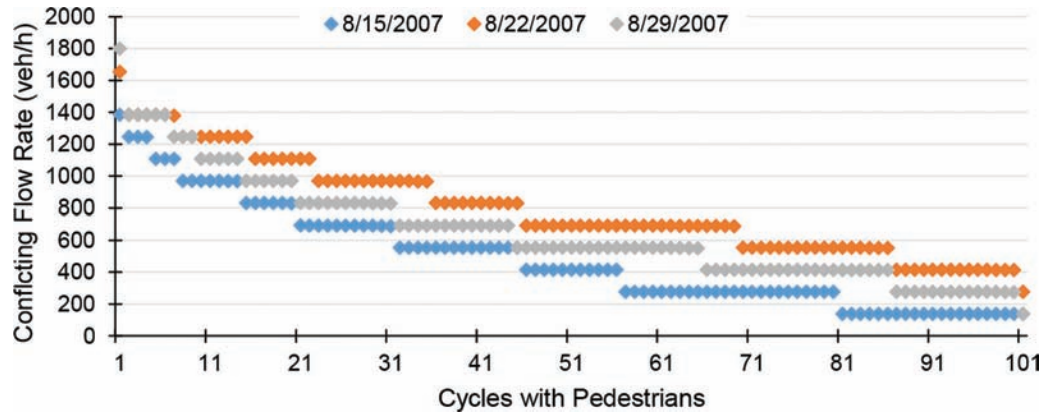


Figure 5.30 Count of the number of vehicles conflicting with pedestrian phase 4 per cycle [80].

unusually, the pedestrian phases were retained on phases 2 and 6, so that pedestrians wishing to cross Stadium Avenue would not trigger the exclusive pedestrian phase. Pedestrians must actuate a Ped button for a movement that crosses Northwestern Avenue to call phase 13.

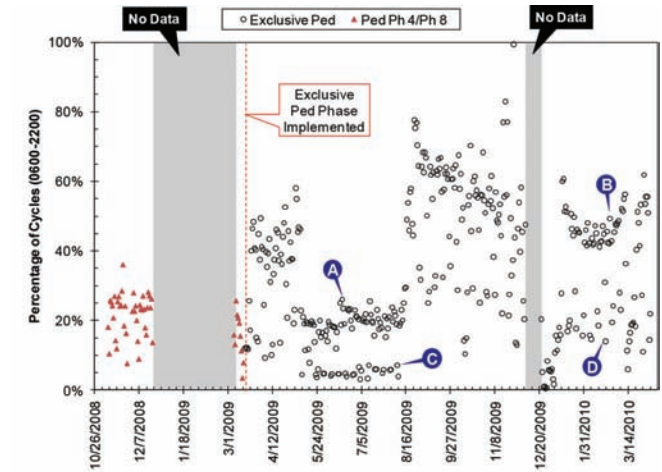
At the time, it was very common for pedestrians to cross the street along the entire length of Northwestern Avenue south of this intersection. This was largely eliminated by a reconstruction of the median in 2014. However, at the time of the study, it was very common for pedestrians to cross one set of vehicle lanes, stand on the median strip (which is very narrow near the intersection), then wait for a gap in the other direction.

Figure 5.31 shows the percentage of cycles during the 06:00–22:00 time period, during which the east–west pedestrian phases were called. Data were collected from October 2008 through April 2010, with some data outages in the winter of 2008–2009 and again briefly in December 2009. The two symbols indicate whether these were served by traditional phases (phase 4/8) or the exclusive phase. Figure 5.31a shows all of the data, whereas Figure 5.31b filters the dates to include only weekdays where academic classes were in session.

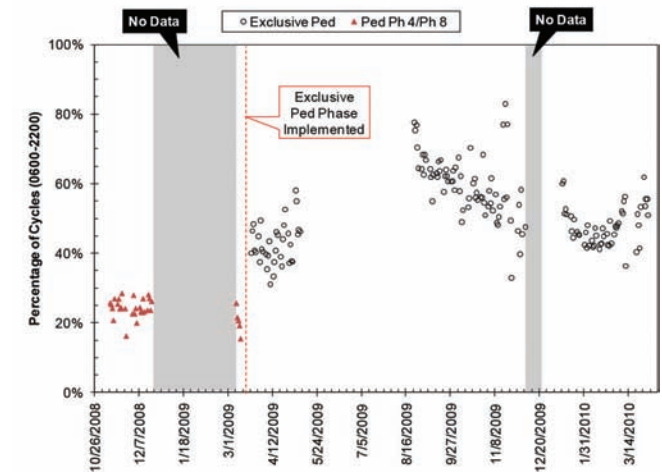
In Figure 5.31a, the dramatic reduction in pedestrian phase utilization can be seen during the summer months, when most students have gone home. There is also some apparent bimodality in the data, which is attributable to there being higher pedestrian phase utilization during weekdays (callouts *A* and *B*) compared with weekends (callouts *C* and *D*). In Figure 5.31b, where the weekends and summer days have been removed, some seasonal trends are clearer. The increase in pedestrian phase actuation after implementing the exclusive pedestrian phase is also apparent.

Figure 5.32 shows the rate of pedestrian phase utilization by day of week between November 2008 (with conventional phases) and November 2009 (after the exclusive pedestrian phase had been in use for eight months). The data in the graph are limited to the 06:00–22:00 hours, and exclude special events, university holidays, and inclement weather. The average pedestrian phase for weekdays increased from 45% to 63%, and for weekends increased from 17% to 30%. The implementation

of the phase induced a substantial and obvious increase in utilization. This was an increase of about a fourth for weekdays, but the weekends saw utilization nearly double.



(a) All days when data were collected.



(b) Only days in the academic year with classes in session.

Figure 5.31 Percentage of cycles with pedestrian phases (during 06:00–22:00) crossing Northwestern Avenue [79].

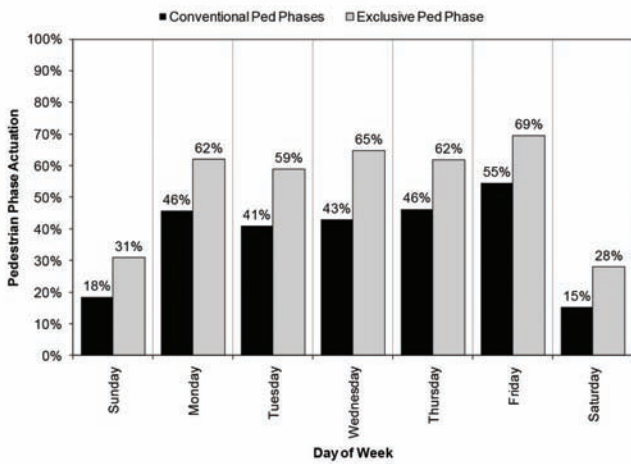


Figure 5.32 Percentage of cycles with pedestrian phases crossing Northwestern Avenue before and after exclusive pedestrian phase implementation (data from November 2008 and November 2009, 06:00–22:00, excluding special events, breaks, and inclement weather) [79].

Besides a change in operation, it is also possible to characterize seasonal variations. Figure 5.33 shows pedestrian utilization across several different comparison sets. Figure 5.33a and Figure 5.33b respectively show weekdays and weekends, with the two series representing the academic year and the non-academic year. Utilization is much higher during the academic year

year for obvious reasons, but there remains a spike in utilization around noon during weekdays during the non-academic year (Figure 5.33a), which can be attributed to university employees crossing the street around lunchtime. We can also see relatively high utilization in the early morning on weekends during the academic year (Figure 5.33b), which is higher than the weekdays. This reveals an interesting activity pattern likely related to the large student population.

Figure 5.33c examines weekly variations, comparing the academic year with the non-academic year. The rate of utilization is very consistent throughout the week in both cases, and there is somewhat more activity on Saturday than on Sunday. The data are tabulated from month to month in Figure 5.33d, which compares weekdays with weekends. The academic year runs from August through April. Pedestrian activity appears to peak in September, but is generally 50% or more of cycles during the academic year on weekdays, falling to about 20% during the non-academic year. Similar trends occur for the weekends, with lower overall rates of utilization.

The pedestrian utilization data were cross-referenced to historical weather data from a nearby weather station to determine how much the two were related to each other. Figure 5.34 shows some comparisons of pedestrian activity with temperature and precipitation; all data are limited to the academic year after implementation of the exclusive pedestrian phase. The trend in

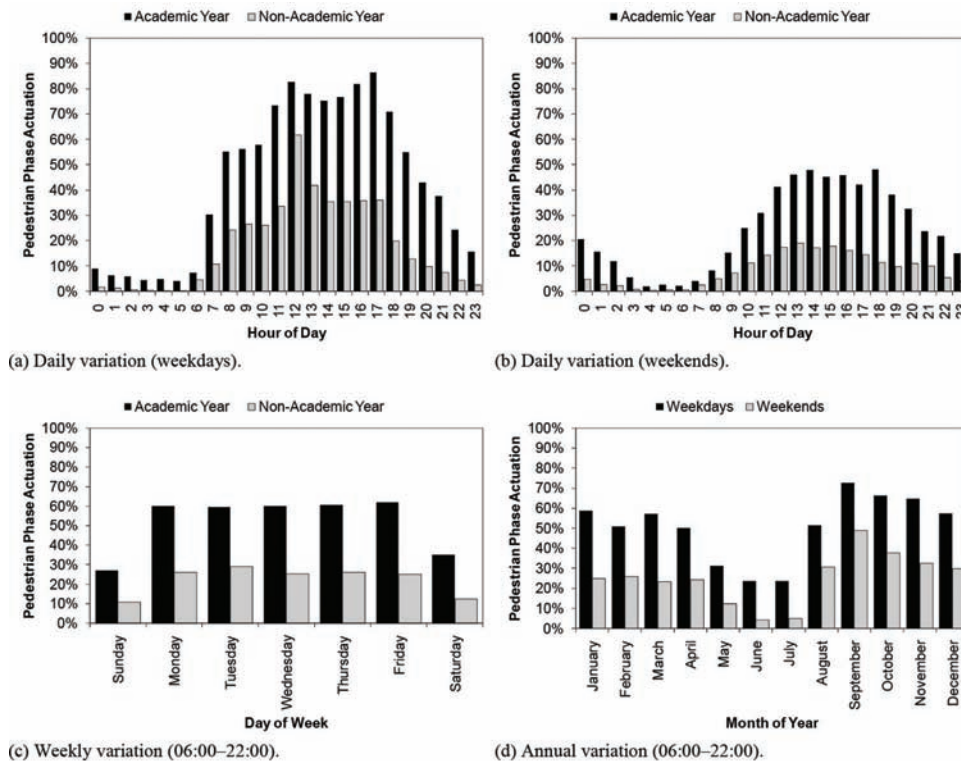
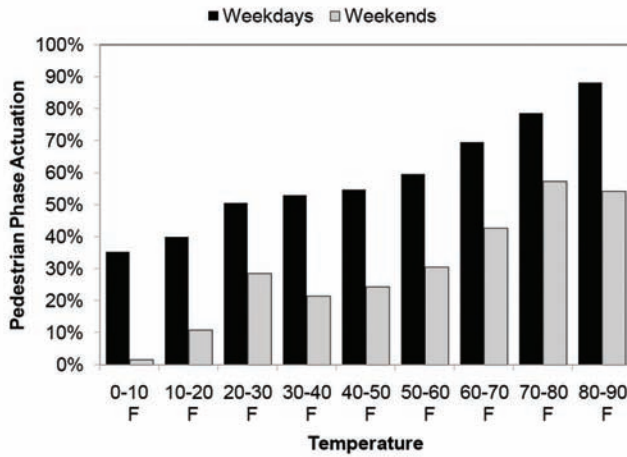
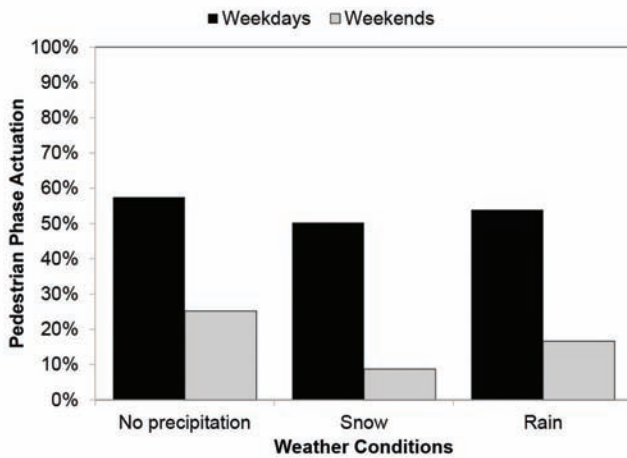


Figure 5.33 Daily, weekly, and annual variation in pedestrian phase actuation (all data from after implementation of exclusive pedestrian phase) [79].



(a) Temperature.



(b) Precipitation.

Figure 5.34 Variation in pedestrian phase actuation due to weather effects (all data from 06:00–22:00 during academic year after implementation of exclusive pedestrian phase) [79].

temperature (Figure 5.34a) is clear: there are fewer pedestrians during colder weather, on both weekdays and weekends. Utilization is especially low on very cold weekends. Precipitation causes a very slight decrease in activity on weekdays (Figure 5.34b), but snowy weekends have less than 10% of cycles with pedestrian phase actuation.

Although this performance measure does not quantify the number of pedestrians, these graphics show that the relative demand can be observed. The results are unlikely to be transferable to other locations, because of the particular characteristics of the location, as an intersection on a university campus. However, this in itself underscores the importance of doing such measurements on site, as a “typical” rate of utilization for a more general location would not capture such characteristics. This also demonstrates the outcome of implementing the phase. The higher rate of pedestrian phase utilization shows that the feature was indeed used by pedestrians. It also implies that more pedestrians are crossing at the

intersection rather than executing less safe crossings at various locations elsewhere along the street.

5.4.7 Special Operational Diagnostics

High-resolution data can serve as a means of validating whether signal control performs as expected. This can be particularly useful for diagnostic purposes for special control functions, such as preemption or advanced control.

A simple example of using the event times to evaluate preemption is the timeline of events in emergency vehicle (EV) preemption [81]. Figure 5.35 shows a sample phase event timeline for a location where EV preemption is used. The timeline begins with the initiation of phases 1 and 6; these terminate and service moves on to 2 and 5. Clearly, this intersection features a lagging left turn (phase 5). At 20:37:04.6, the preempt becomes active. At the same exact time, phases 2 and 5 are terminated and the controller begins green on phases 4 and 8. At 20:37:33.8, the preempt input becomes inactive; very shortly after, phases 4 and 8 terminate, and the controller returns service to phases 2 and 6. This is the beginning of the preempt exit interval, as recorded by the final event at 20:37:40.1. Figure 5.36 shows these events in a phase status timeline. The duration of the preempt is clearly observable, as well as the time between the preempt input and the start of the

Time Stamp	Event Type	Phase
20:36:07.4	Green	Ø1
20:36:07.4	Green	Ø6
20:36:20.4	Yellow	Ø1
20:36:24.0	Red	Ø1
20:36:26.1	Green	Ø2
20:36:51.0	Yellow	Ø6
20:36:55.3	Red	Ø6
20:36:57.0	Green	Ø5
20:37:04.6	Yellow	Ø2
20:37:04.6	Yellow	Ø5
20:37:04.6	Preempt On	Ø4+Ø8
20:37:08.2	Red	Ø5
20:37:08.9	Red	Ø2
20:37:10.6	Green	Ø4
20:37:10.6	Green	Ø8
20:37:33.8	Preempt Off	Ø4+Ø8
20:37:34.0	Yellow	Ø4
20:37:34.0	Yellow	Ø8
20:37:37.6	Red	Ø4
20:37:37.6	Red	Ø8
20:37:40.0	Green	Ø2
20:37:40.0	Green	Ø6
20:37:40.1	Preempt Begin Exit Interval	Ø4+Ø8

Figure 5.35 High-resolution phase and preemption data, US 36 at Ronald Reagan Parkway, 10/14/2014 [81].

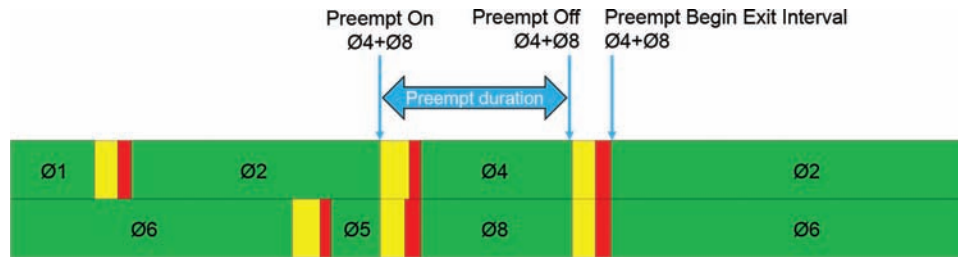


Figure 5.36 Phase and preemption data, US 36 at Ronald Reagan Parkway, 10/14/2014 [81].

prioritized phases, as well as the impact on non-prioritized phases, such as the truncation of phases 2 and 5.

Railroad preemption is an extremely important controller function for intersections close to railroad crossings. A common railroad preemption strategy is to use a track clearance phase to clear queues from the crossing area, before going on to limited service.

Figure 5.37 shows a diagram explaining the sequence of events that are preferably well choreographed during railroad preemption. Preempt begins when the advance warning becomes active, which triggers the preemptor. The train is still somewhat upstream of the intersection, and the crossing itself is not active. The time needed for the signal to begin serving track clearance is the right-of-way transfer time (RTT). This is highly variable, as it strongly depends on the current phases in service; the controller cannot violate absolute minimum greens or yellow and red clearance times.

What should happen next is that the railroad crossing becomes active and the crossing gates begin to descend. Ideally, the gates will be fully down before the end of the track clearance phase, after which the controller begins limited service. Ideally, the track clearance phase will remain green until after the gates have fully descended. If the gates are still up when the track clearance phase ends green, queues could build on top of the railroad tracks. Because the signal would be serving some other movement, the queued vehicles might not be able to make an emergency maneuver to avoid a collision. This is especially true of long vehicles such as buses, combination trucks, or vehicles with trailers. In other words, extension of track clearance green past the gate down event is critical.

Figure 5.37 shows track clearance green (TCG) as affixed portion with a subsequent extension. The extension option is possible when a gate down circuit exists. This is a connection between the railroad warning equipment and the signal controller that indicates when the gates are fully descended. This is not always available. One possibility is to use a very long fixed TCG interval, but because the amount of variation in RTT might not always be known, the upper bound of this value is uncertain.

Figure 5.38 describes a location where high-resolution data was used to measure preemption activity to evaluate the variability in RTT and the performance of TCG [82]. This is the intersection of Pendleton Pike and Carroll Road in Indianapolis, Indiana. The railroad crossing is situated

on the southbound lanes; the detector configuration is shown in Figure 5.38a, and the corresponding ring diagram is shown in Figure 5.38b. This intersection features presignals before the railroad crossing. However, motorists have a tendency to ignore these and proceed through to the second set of signal heads, even when steerable signal heads are used [82,83].

Because all of the relevant controller events can be measured (start of preemption, beginning of TCG, and end of TCG), it is possible to begin measuring what is occurring in the field.

Figure 5.39 shows a distribution of measured RTT for thousands of individual preemption events in 2008 and 2009 [82]. The longer the RTT, the less likely the TCG interval will sufficiently clear traffic before the arrival of the train. Figure 5.39a shows the distribution for about 2,000 preemptions under the existing preempt entry logic, and Figure 5.39b shows 2,600 preemptions after some changes to the that logic. Figure 5.39a shows that RTT is often as small as 3 seconds, but varies and can be as long as 16 seconds. About 2% of all the observed preemptions had a RTT as long as 21 seconds. Figure 5.39b shows the distribution of RTT with improved logic; in this case, the most common RTT was 6 seconds, but the upper bound decreased to 13 s.

The impact of the TCG settings is measured in Figure 5.40. This shows the difference between the end of TCG and the gate down time, which is here called the “residual time” [82,84]. Negative numbers represent the situation where TCG ends before the gates are down, which, as mentioned before, is highly undesirable. The existing preempt settings are shown in Figure 5.40a (over 3200 observations); the fixed TCG interval length was 15 seconds. The distribution shows that the residual times have a 35-second range of variation. Not only does TCG end prior to gate down for over half of the preemption events, but it sometimes ends over 10 seconds prior to gate down—which is ample time for vehicles to queue in front of the tracks.

Figure 5.40b shows the distribution of residual times after extending the fixed TCG interval from 15 to 20 seconds (1760 observations). The distribution has shifted to the right substantially. There are now fewer than 10% of preemptions where TCG ended prior to gate down, and the residual times have been decreased substantially. However, although the TCG deficiencies have been reduced, performance is still not acceptable;

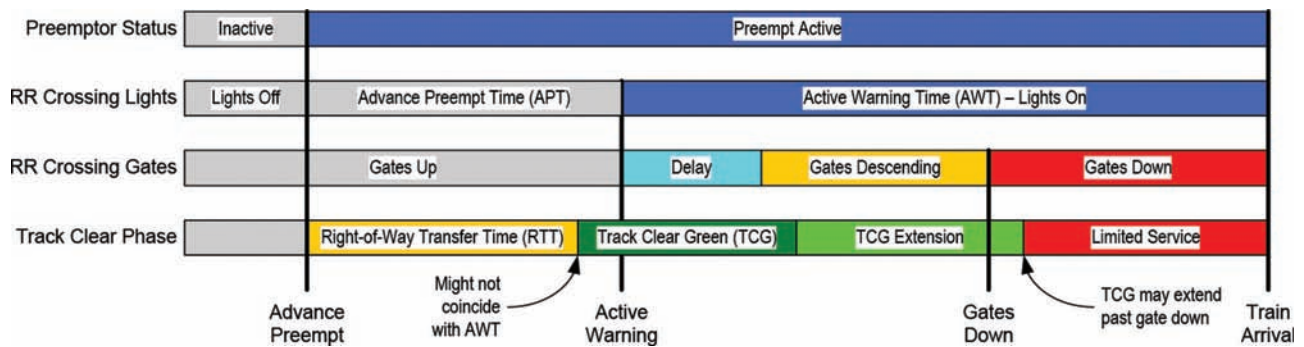


Figure 5.37 Railroad preemption events.

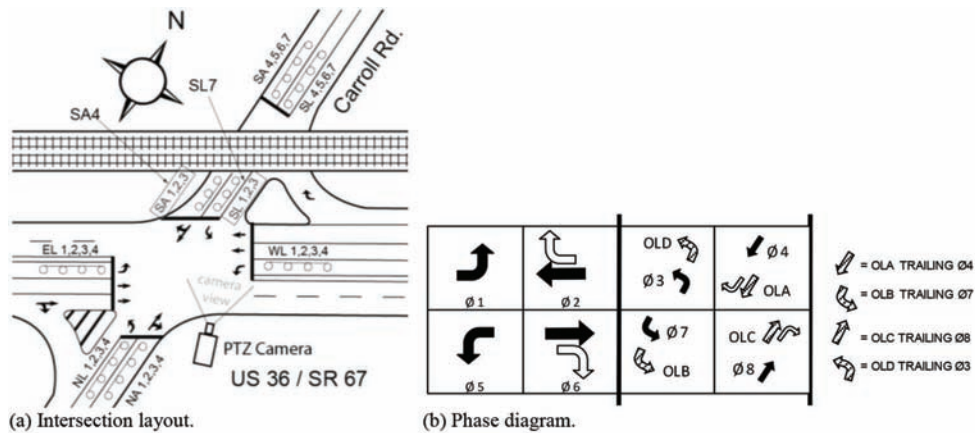


Figure 5.38 Evaluation of track clearance phase timing during signal preemption [82].

even one event with a negative residual time represents a potentially dangerous situation.

Figure 5.40c shows the distribution of residual times after implementing a gate down circuit to extend TCG until gate down. The overall shape of the distribution is similar to Figure 5.40b, except that the right tail has been cut off; the minimum residual time is now zero seconds; much of the spike in the distribution at zero represents the previously existing tail. There are now zero preemption events where TCG occurred before gate down. The data in this plot represent over 2600 observed preemption events.

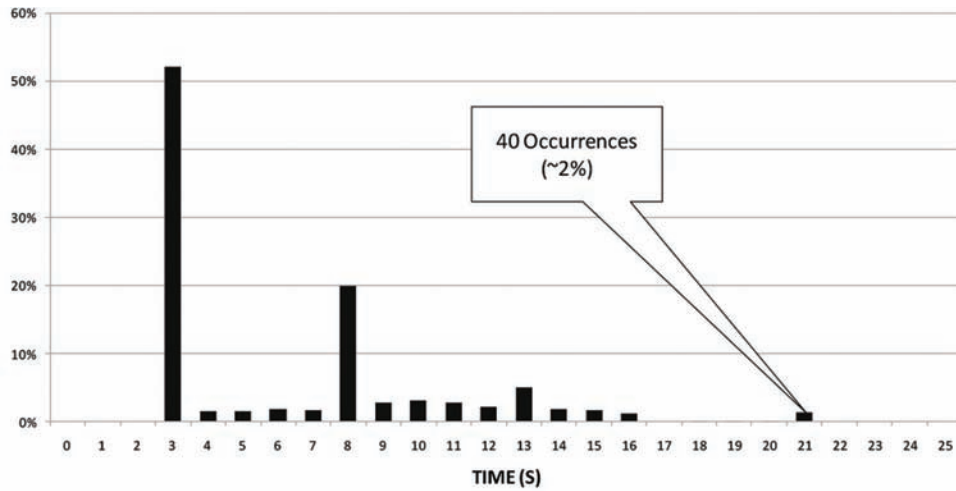
This example demonstrates the utility of high-resolution data to evaluate whether a critical safety function is performing as expected. While it is indeed possible to manually observe whether the configuration is working properly, such observations can only cover a limited amount of time when the engineer is present in the field. The use of automatic data collection enables thousands of events to be evaluated over many months, which is a far more powerful means of validating the operation.

Another control application that can be evaluated in a similar fashion is the dynamic adjustment of phase sequences. This is done by advanced control algorithms; one such feature is called “phase reservice,” which adjusts the permissive periods during coordination. The idea is that if the controller is resting on the coordinated

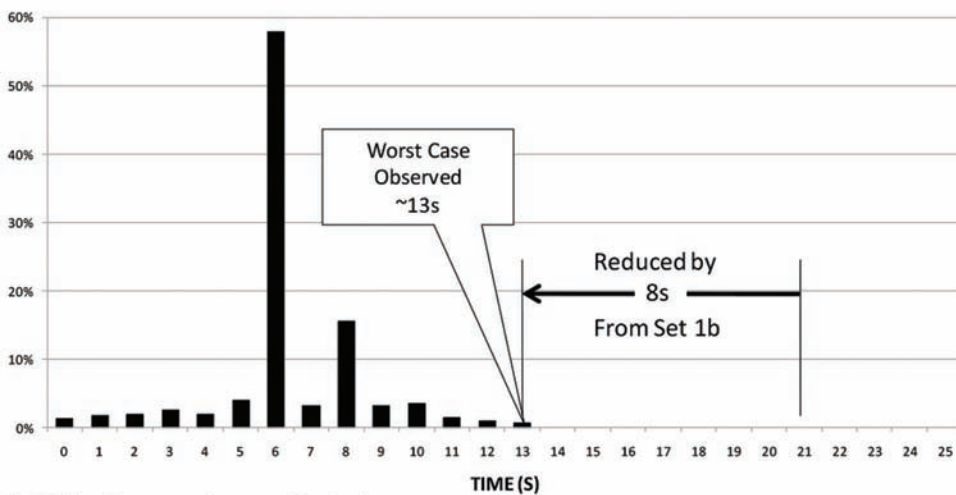
phases but is not yet into the coordinated phase split, if a call arrives on a noncoordinated phase, the controller has the option of leaving the coordinated phases to serve the noncoordinated phase early, even if that phase’s permissive period would not normally be open. This requires a dynamic change in the phase sequence.

Figure 5.41 shows two conceptual example event timelines that compare operation before and after implementation of phase reservice [75]. The coordinated phases are 2 and 6, while phase 8 is a noncoordinated phase. In Figure 5.41a, the coordinated phases cannot end until after the yield point at 9:15:50. The coordinated phases terminate and phase 8 is served shortly after, beginning at 9:16:55. After the end of green at around 9:16:20, it is not possible for phase 8 to be served again until the next cycle. The earliest possible beginning of green is 9:17:30. In Figure 5.41b, it is possible for the controller to leave phases 2 and 6 and begin service on phase 8 again at 9:16:40, meaning that it would be possible for vehicles waiting on phase 8 to be served earlier than usual.

Figure 5.42 shows some results that were measured after implementing phase reservice at the intersection of US 231 and Martin Jischke Boulevard in West Lafayette, Indiana. For reference, a map of the intersection is presented in Figure 5.43 showing the detector locations and phase assignments. Phases 2 and 6 are coordinated and phases, whereas phases 1 and 8 are noncoordinated.



(a) Using existing controller logic.



(b) Using improved controller logic.

Figure 5.39 Measurement of right-of-way transfer time (RTT) during signal preemption [82,84].

Figure 5.42 shows an analysis of phase events per 30 minutes between 9:00 and 15:00 on 2/2/2014. Figure 5.42a shows the v/c ratio for phase 2 during the same interval. Although there are some cycles in which phase 2's v/c ratio exceeds 50%, it is well below 100% during the entire analysis period. Figure 5.42b indicates the number of cycles in which phase 1 was served, and whether it was served once or twice in the cycle, and Figure 5.42d shows the same information for phase 8. Figure 5.42c, arranged between the two, shows the cycles where *both* noncoordinated phases were served, and whether they were served in or out of sequence.

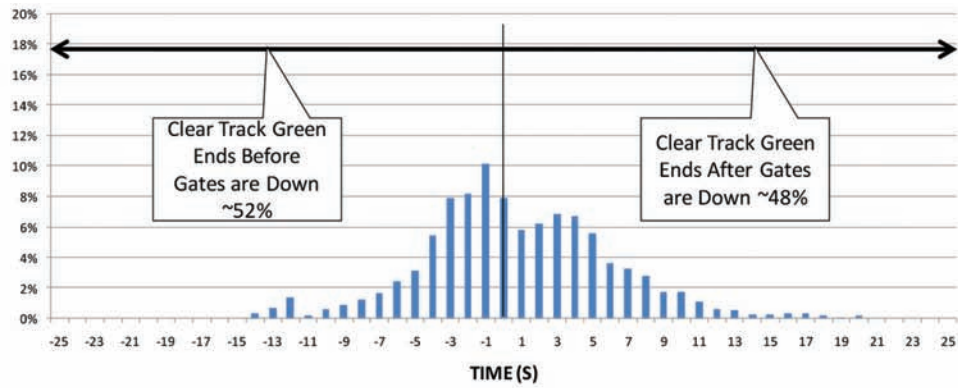
The impact of using the feature can also be evaluated using performance measures. Figure 5.44 shows distributions of the maximum vehicle delay, which is an estimate of the likely vehicle delay measured by taking the time difference between the vehicle arrival during red and the clearance of the queue during the subsequent green [75]. These delays are compared for

phase 1 (Figure 5.44a) and phase 8 (Figure 5.44b) for several different signal control options:

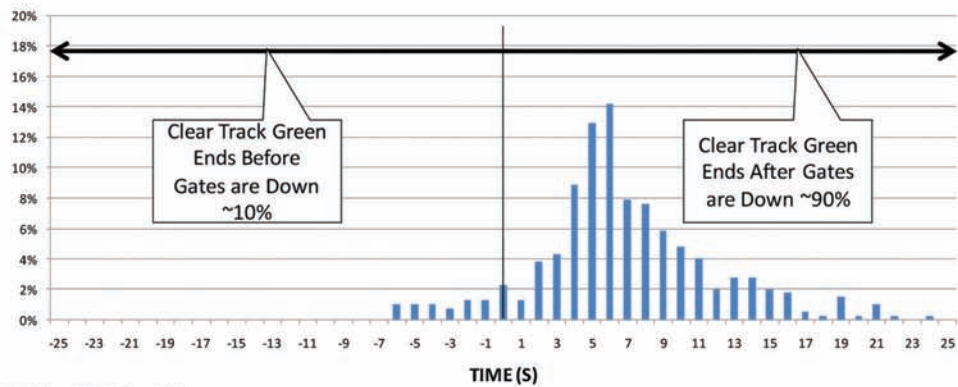
- Free (fully actuated, noncoordinated);
- “20%”: coordinated with early yield using 20% split extension;
- “20% Ph. Res”: the same as the above, but with the phase reservice feature enabled;
- “10%”: coordinated with early yield using 10% split extension;
- “0%”: coordinated without early yield.

Early yield, also called the “actuated coordinated” feature, divides the coordinated phase into a nonactuated portion and an actuated portion. The split extension is the percentage of cycle used to determine how much of the coordinated phase should be actuated.

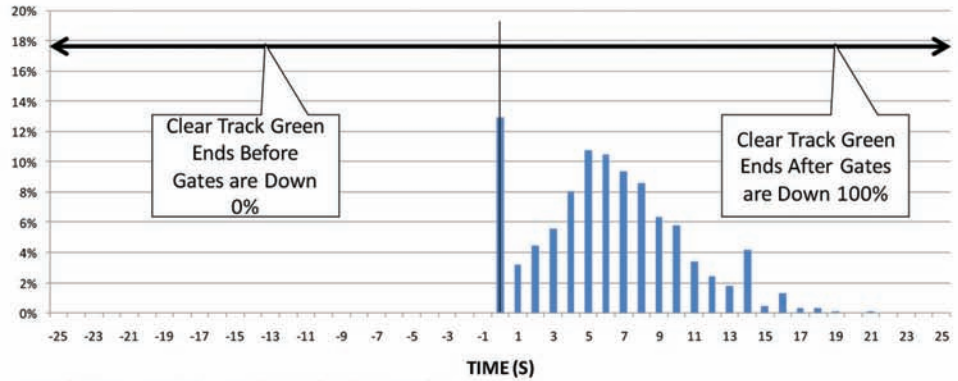
For each time of day, free operation has the lowest delay for both phases 1 and 8. This is not unexpected, since coordination holds green on the mainline, forcing



(a) Fixed TCG = 15 s.



(b) Fixed TCG = 20 s.



(c) Variable TCG with gate down circuit extension.

Figure 5.40 Evaluation of track clearance green performance [82,84]: time difference between the end of TCG and gate down confirmation.

phases 1 and 8 to wait. During the midday plan, the 20% with phase reservice option has almost the same delay as free operation for phase 1 (Figure 5.44a) during the midday and early night. This demonstrates a strong positive impact of reservice on phase 1. However, the same benefit is *not* seen for phase 8 (Figure 5.44b). Here, only marginal improvements are achieved in comparison with the “20%” option (without phase reservice). The results show that the phase reservice feature tends to strongly benefit phase 1 rather than phase 8, at least in its implementation at this intersection. Additionally,

the results show that, for almost every time of day, use of early yield lowers delay for phases 1 and 8. That is, delays tend to be highest for the 0% option compared with 10% or 20%. This is consistently the case for phase 8, but phase 1 shows some variation.

These examples show potential ways in which advanced control features can be examined in detail and their impacts directly measured. As such features become increasingly common in signal controllers, it will be important to determine whether they operate as intended and that their use creates a benefit for stakeholders.

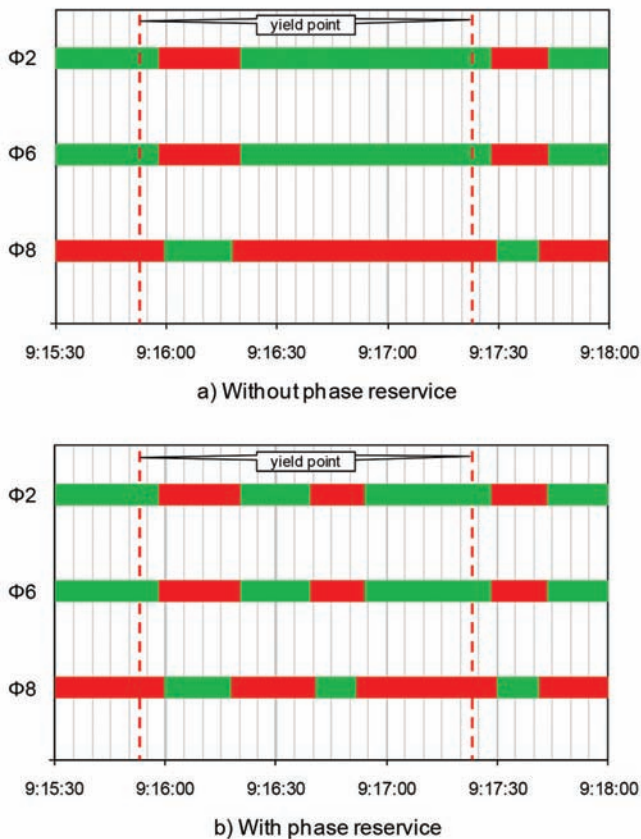


Figure 5.41 Phase green and red times with and without phase reserve [75].

5.5 Case Study of a Single Intersection during a Freeway Diversion

5.5.1 Overview

This section presents observations from a scenario where freeway traffic was diverted onto an arterial route during a month-long emergency bridge repair on the freeway. Most agencies have parallel arterial routes parallel to freeways that take on the freeway traffic during incidents. The objectives of integrated corridor management as a means to cooperatively manage the two parallel facilities would be facilitated by the use of signal performance measures as a means of evaluating the conditions in real time and assessing incident outcomes. This section presents a detailed analysis for a single intersection to demonstrate the use of performance measures to facilitate operational comparisons.

Figure 5.45 shows a map of the affected region, which is a span of I-65 between Indianapolis and Chicago. The bridge closure affected the northbound bridge over the Wildcat Creek on the northeast side of Lafayette, Indiana. Because of extensive road work occurring on alternate routes in Lafayette, traffic was instead diverted onto US 52 and US 231, as shown by the black line. All of the traffic signals along US 231 were affected by the diversion, and signal timing was

adjusted to accommodate it. The signal timing was then readjusted for normal conditions after the end of the detour. This section will focus on operations at US 231 and River Road, with the overall corridor being discussed later in this report.

A plan of US 231 and River Road is presented in Figure 5.46, which shows the detector configuration and the phase scheme used at this intersection. The intersection operates as an eight-phase intersection. However, phase 3 and phase 7 (the westbound and eastbound left turns) are physically incompatible, so a lead/lag sequence is used for the east-west street (River Road) to keep them separated. As is typical of most INDOT intersections, the detection includes setback detectors on the mainline through movements (northbound, southbound) and stop bar detectors on all of the minor movements. This intersection also includes, rather untypically, advance detectors on the side street approaches and an array of detectors on the through lanes that are 100 ft upstream of the stop bar. These “untypical” detectors were not used in the following analysis.

5.5.2 Vehicle Volumes

Northbound I-65 has an AADT of approximately 25,000 vehicles per day on the closed section. The closure necessitated the diversion of that traffic to the detour route shown in Figure 5.45. Not all of the usual traffic went that way, however. Advance warning of the detour caused many drivers to seek alternate routes that bypassed the entire region, and other motorists used alternate, unsigned detours through the region (such as continuing along US 52 through Lafayette).

Figure 5.47 shows the measured northbound volumes at US 231 and River Road for several different dates in the summer of 2015. Normal volumes are shown by the 7/22/2015 dataset. There is a northbound peak in the AM close to 8:00, and the rest of the day the volumes are substantially lower. During the middle of the night the volumes are particularly low. The detour volumes are shown by the 8/12/2015 and 8/19/2015 datasets. Volumes increased for all times of day. The AM peak at 8:00 has about the same amount of volume in both cases, but during the rest of the day, volumes were doubled or more. The total volume on 7/22/2015 was 27,472. This increased to 56,592 on 8/12/2015 and even further, to 60,308, on 8/19/2015. After the detour ended, volumes reverted to normal, as shown by the 9/9/2015 and 9/16/2015 datasets.

Volumes for all eight phases at US 231 and River Road are shown in Figure 5.48. This plot shows the same observation dates as Figure 5.47. At a glance, the eight-phase view shows that the northbound movement was by far the heaviest affected movement; there are no significant differences on the other phases, except for an intermittent spike on phase 7 that occurred on 8/19/2015. In summary, the I-65 diversion caused volumes on the northbound movement to double while the other movements did not see any change in their volumes.

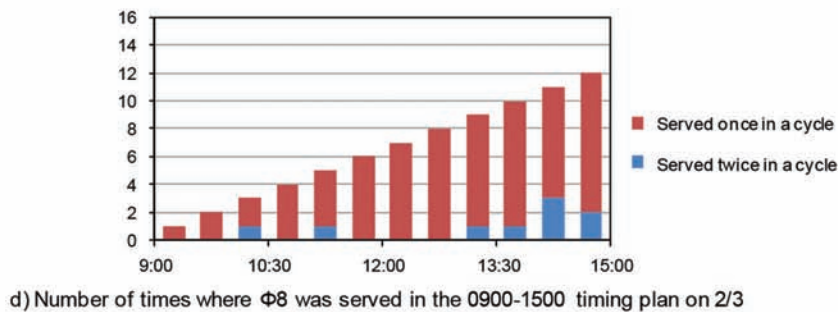
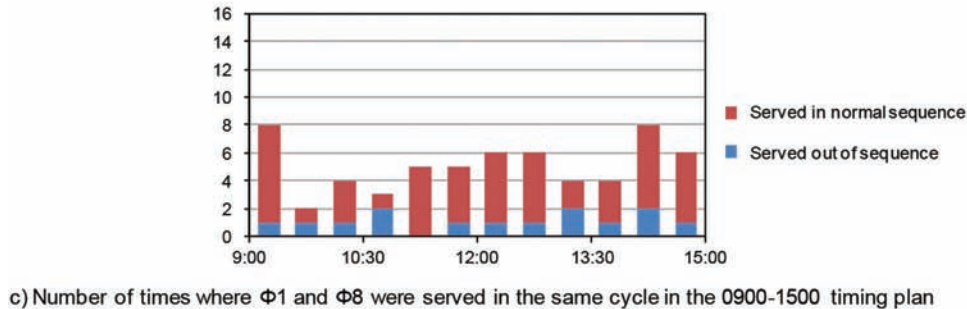
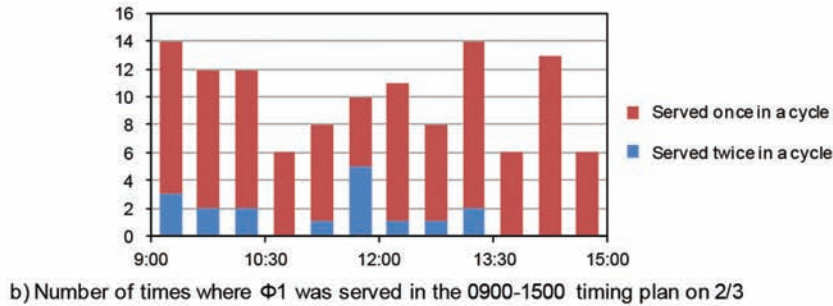
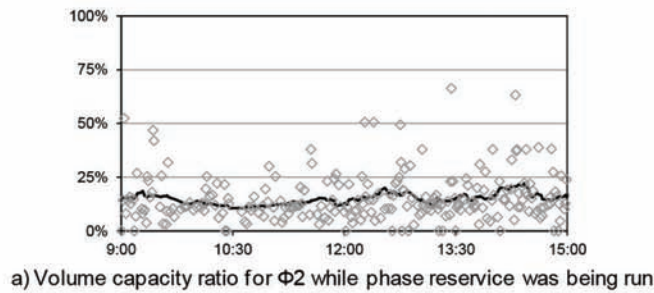


Figure 5.42 Analysis of signal operations with phase reserve [75]. Data are shown from US 231 and Jischke Boulevard on 2/3/2014.

5.5.3 Cycle Length

Figure 5.49 includes several plots of actual cycle lengths at US 231 and River Road on different dates in the summer and fall of 2015. The cycle lengths changed several times over this time period because of the closure of a nearby interstate route that diverted volumes in one direction onto northbound US 231. At a very high level, the plots of cycle length tell almost the entire story of the I-65 diversion, at least in terms of how the signal operation changed on US 231:

- Before the summer of 2015 (Figure 5.49a, 7/22/2015), the corridor ran free, except for a 90-second cycle in the AM peak (6:00–9:00). During the rest of the day, the cycle length hovers around 75 seconds with some spikes during the PM peak close to 17:00. The overnight periods see long cycle lengths as the signal rests in green along the mainline when there is no demand for other movements.
- At the end of the first week of August, the interstate was closed, diverting traffic to US 231 (Figure 5.49a, 8/12/2015). The greatly increased amount of traffic caused the cycle length to substantially increase throughout the entire day except for the AM peak, which initially retained its programmed cycle length of 90 seconds.

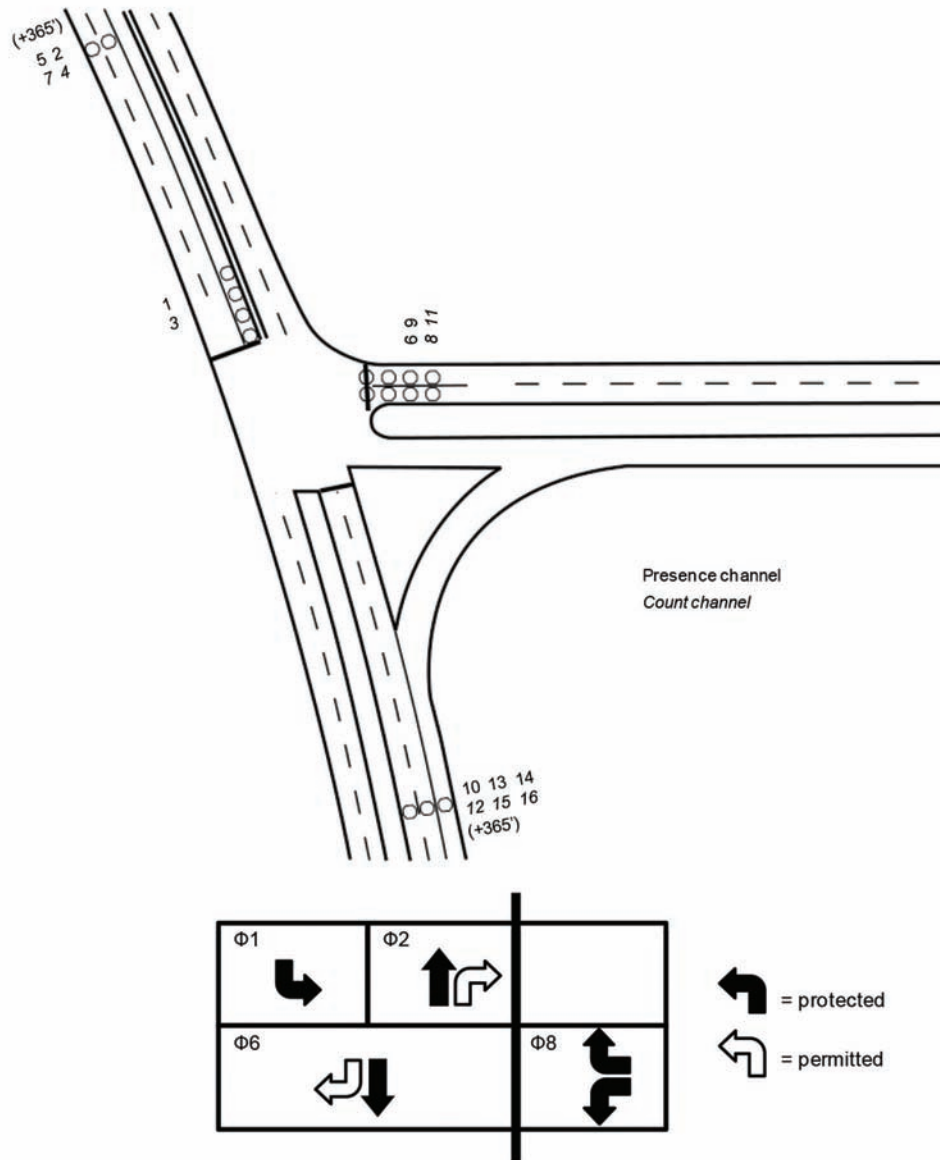


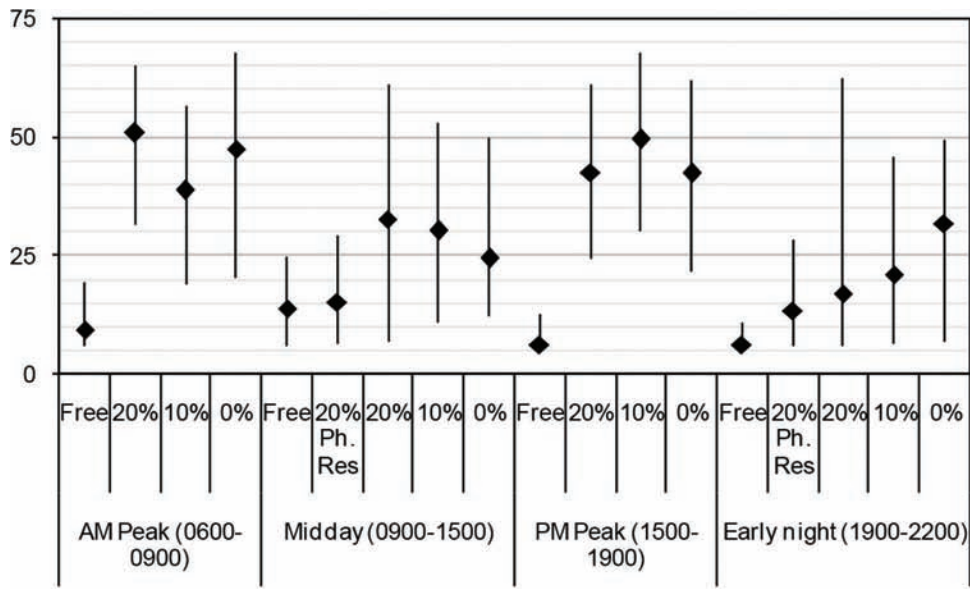
Figure 5.43 US 231 and Jischke Boulevard, West Lafayette, Indiana [75].

- By the end of the second week of August, a timing plan had been devised for the corridor under the diversion volumes, the goal of which was to create a very large northbound green band to shunt the traffic through the area. A cycle length of 150 seconds was used (Figure 5.49a, 8/19/2015), which persists throughout the entire day, including the overnight period. Small variations between cycles occur owing to the use of actuated coordinated phases.
- The 150-second cycle length is used for several weeks. The interstate was reopened in the second week of September, but the TOD plan was not immediately changed (Figure 5.49b, 9/9/2015).
- In about a week's time, new volumes were collected and a new timing plan was designed for "ordinary" conditions. It was decided to implement coordination for the entire

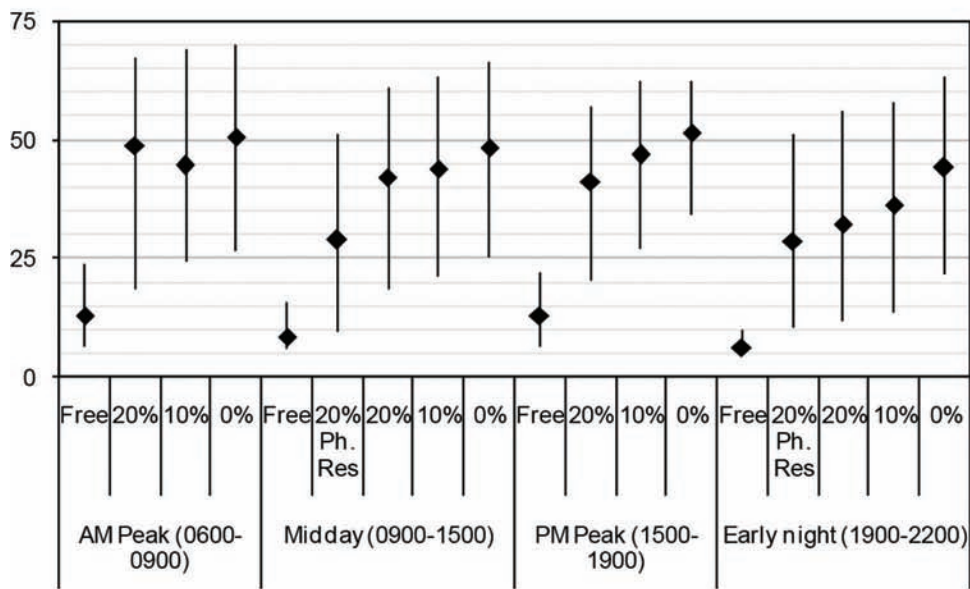
day between 6:00 and 22:00. Following typical agency practices, the initial TOD plan was conceived as consisting of an AM peak pattern, a PM peak pattern, and a default pattern to run during the midday, evening, and weekends (Figure 5.49b, 9/16/2015). A cycle length of 116 seconds was used for the peaks and 82 seconds for the default pattern.

- After another month of fine-tuning the timing plan, the TOD plan was changed to better accommodate the traffic patterns, with cycle lengths changing around somewhat (Figure 5.49b, 10/21/2015).

From this it is possible to guess at how the side street delays changed over the time period, with cycle lengths nearly doubling during the diversion, and recovering afterward to "normal" levels.



a) $\Phi 1$ (SBL)



b) $\Phi 8$ (WB)

Figure 5.44 Max vehicle delay at Jischke Drive with phase reservice and different split extensions [75]. The dot shows the median value and the bar shows the interquartile range spanning the 25th and 75th percentiles.

5.5.4 Occurrence of Split Failures

The introduction of a massive amount of traffic into the northbound direction on US 231 profoundly impacted the signal operation, as engineers were forced to introduce a coordination plan with a very long cycle length into an area where for most of the day the intersections ran free, and usually with relatively short cycles. Besides the most obviously affected northbound

movement, it is also desirable to know how the changes affected the other movements.

One way to evaluate the quality of the local control for those phases is to consider the number of split failures for each one. As mentioned before, each split failure represents a potential public complaint call about not receiving enough green time. The distribution of split failures, and the concentration of a large number of any on a particular movement or group of

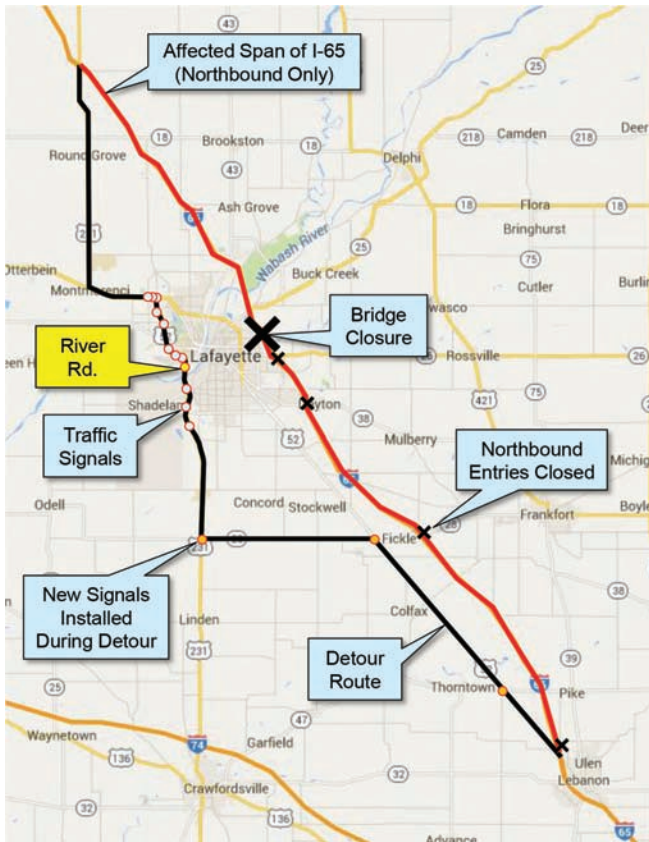


Figure 5.45 US 231 diversion route.

movements, is a good indicator of an inefficient allocation of capacity.

Figure 5.50 shows the number of split failures per phase at US 231 and River Road for several different dates during the detour route. These are color-coded by phase number. The split failures are determined by finding the number of cycles either with the v/c ratio exceeding 1.0 (phases 2 and 6) or with GOR and ROR both exceeding 80% (all other phases).

Prior to the diversion, there were only 20 split failures in total at the intersection on 7/22/2015. There were several on the coordinated phases, and the others occurred on phases 3 and 5. Phase 3, the westbound left turn, had historically been a problematic movement, with peaking of demand causing repeated split failures as the ones illustrated in Figure 5.3. However, under free operation for most of the day, there was usually enough time during each cycle to serve the demand.

On 8/12/2015, after the I-65 diversion, the increased volume on northbound phase 2 caused the number of split failures on that movement to be increased by sixfold compared with 7/22/2015. Surprisingly, however, despite the increase in cycle length, there were relatively few increases in split failures on the other movements. Again, the fully actuated control seemed to terminate phases after their queues had dispersed most of the time.

On 8/19/2015, after the implementation of the detour coordination plan, the split failure chart is markedly different. The detour plan used a 150-second cycle and

aggressively sought a large northbound green band to effectively shunt as much diverted freeway traffic through the corridor with minimal stops. Unsurprisingly, the number of split failures on phase 2 fell to zero. Phase 3, on the other hand, sees a dramatic increase in the number of failures. Under coordination, this phase is often forced off before all of the demand can be served.

The 9/9/2015 data represent a brief period between the reopening of northbound I-65 and while the detour timing plan was still in effect on US 231. Although the number of split failures has decreased, phase 3 still experiences a large number. This shows that with the northbound volumes lessened, there is now an amount of spare green to assist phase 3, but it is not really enough to reduce the number of split failures to a similar amount as the previous operation.

Finally, on 9/16/2015, after retiming the corridor for normal conditions, the split failure profile adjusts once again. Cycle lengths were decreased for all times of day. There are now a few split failures on phases 2 and 6 again, similar to the situation on 7/22/2015, and phase 3 has seen the number of split failure cut by half. The timing plan is still relatively fresh and there are likely opportunities to refine it, but the overall number of split failures is less than the diversion route plan.

5.5.5 Estimated Average Delay

The last performance measure to be examined in the single-intersection example will be the estimated average delay, by phase. The distribution of delays across the intersection speaks to the equitability of the performance as well as its efficiency. The five sample dates used in the cycle length and split failure examples facilitate four before/after comparisons that each tell an interesting story. These are presented in the next four figures (Figure 5.51, Figure 5.52, Figure 5.53, and Figure 5.54).

Each of these plots shows a cumulative frequency diagram (CFD) of the cycle-by-cycle average delays from 6:00 to 22:00 for two analysis periods covering the “before” and “after” periods. The average delays are found using the input–output method (Figure 5.20) for phases 2 and 6, and the random-arrival delay method (Figure 5.22) for the other phases.

- *Comparison 1: Addition of the detour traffic under free operation* (Figure 5.51). During normal operations prior to the detour (7/22/2015), delays were relatively low for most of the movements at the intersection. Median delays were 50 seconds or less for the noncoordinated phases, whereas phases 2 and 6 had median delays of about 20 seconds. This reflects the short cycle length that prevailed during most of the operation. The introduction of the detour traffic (8/12/2015) substantially increased the delay on every phase at the intersection, which correlates with the increase in the cycle length (Figure 5.49a).
- *Comparison 2: Implementation of a detour timing plan* (Figure 5.52). The detour coordination plan strongly favored northbound progression, and included very long green times for phases 2 and 6. Unsurprisingly, the delay for phases 2 and 6 is decreased substantially, while it increases a great deal for all the other phases. The median of the

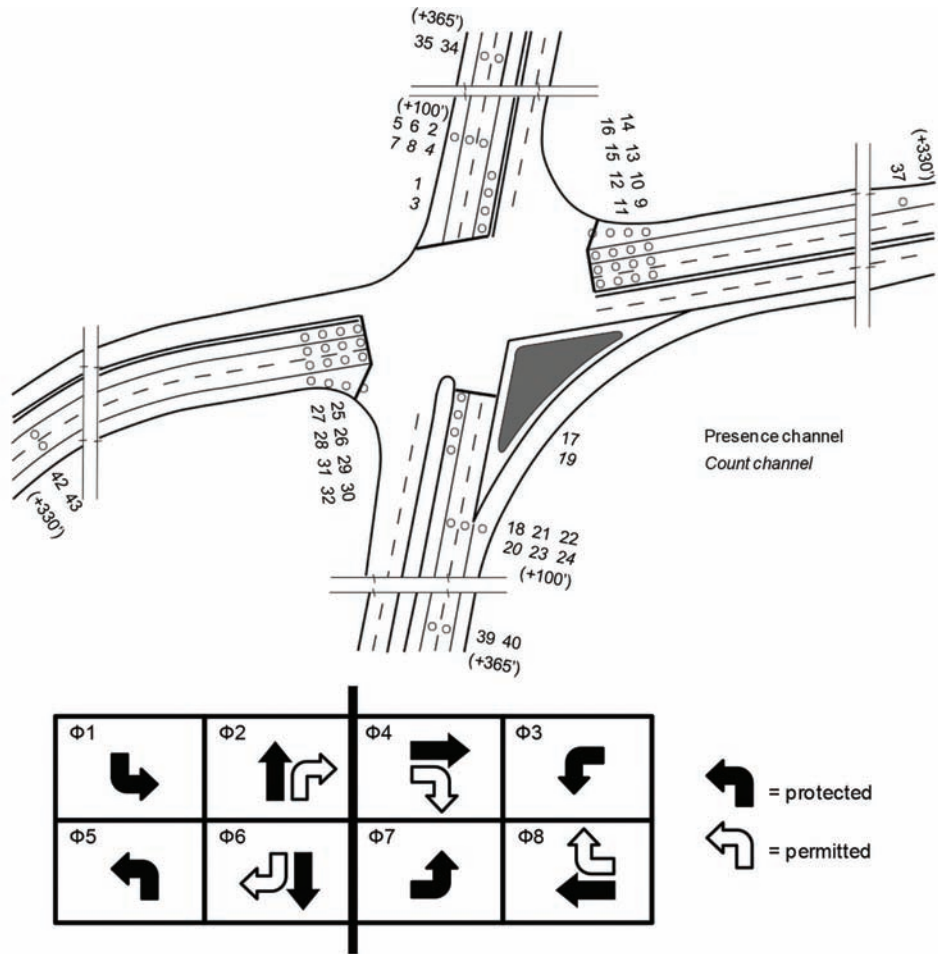


Figure 5.46 US 231 and River Road, West Lafayette, Indiana [75].

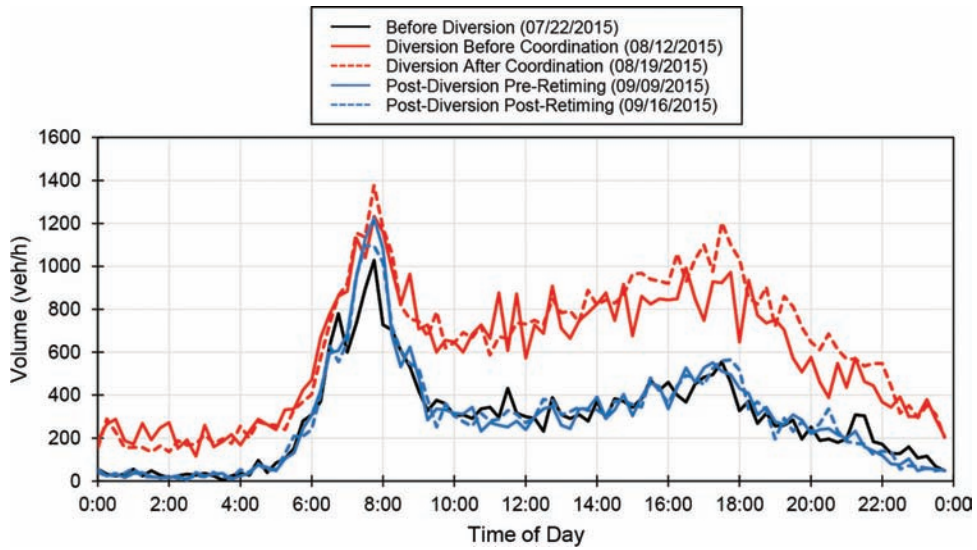


Figure 5.47 Northbound traffic volumes at US 231 and River Road.

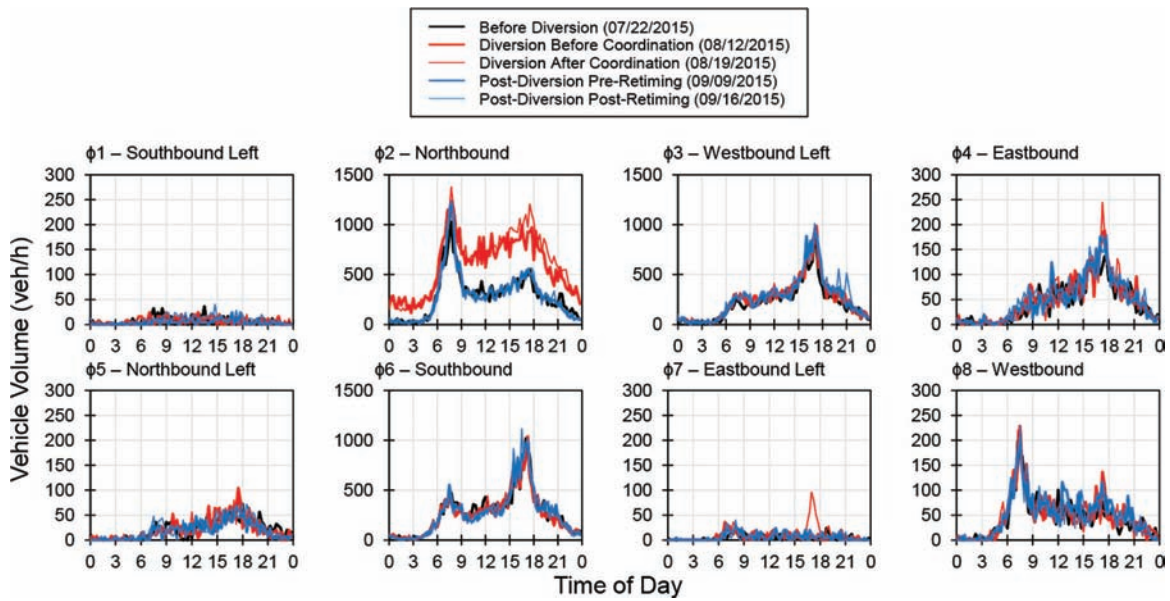


Figure 5.48 Volumes on 8 phases at US 231 and River Road.

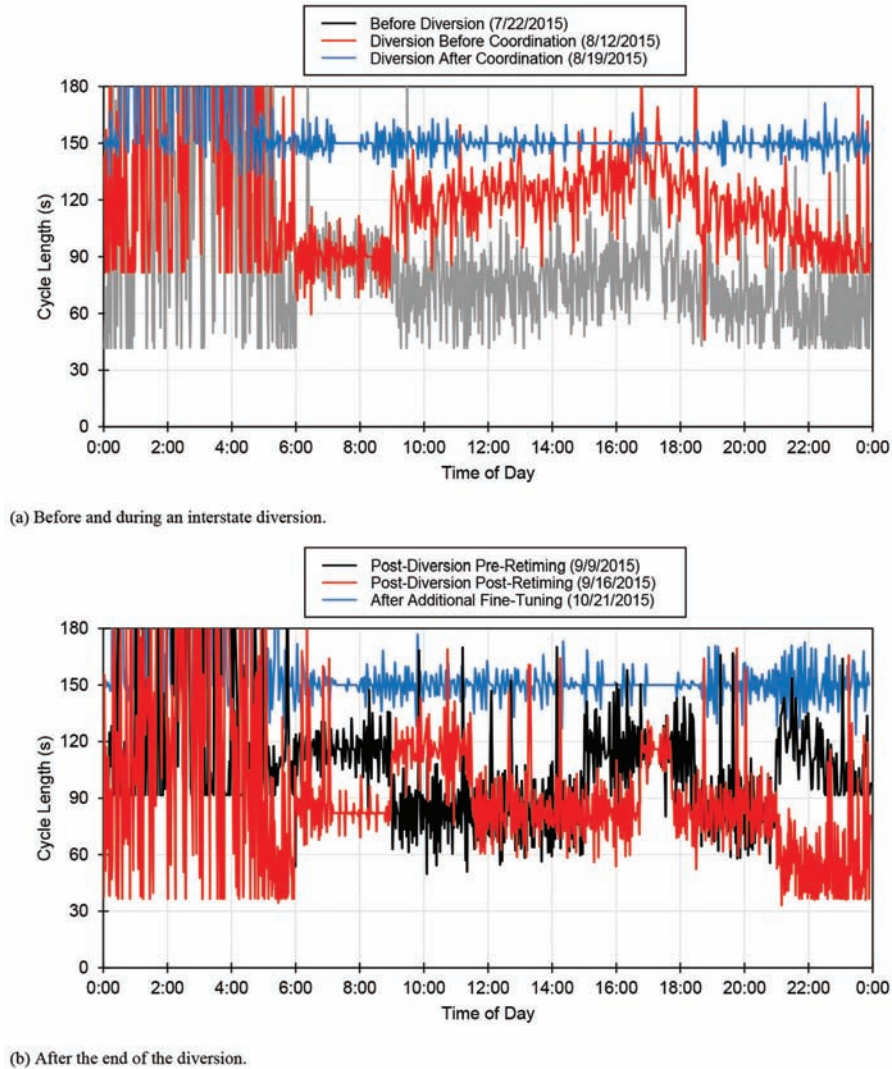


Figure 5.49 Cycle lengths at US 231 and River Road on various dates in 2015.

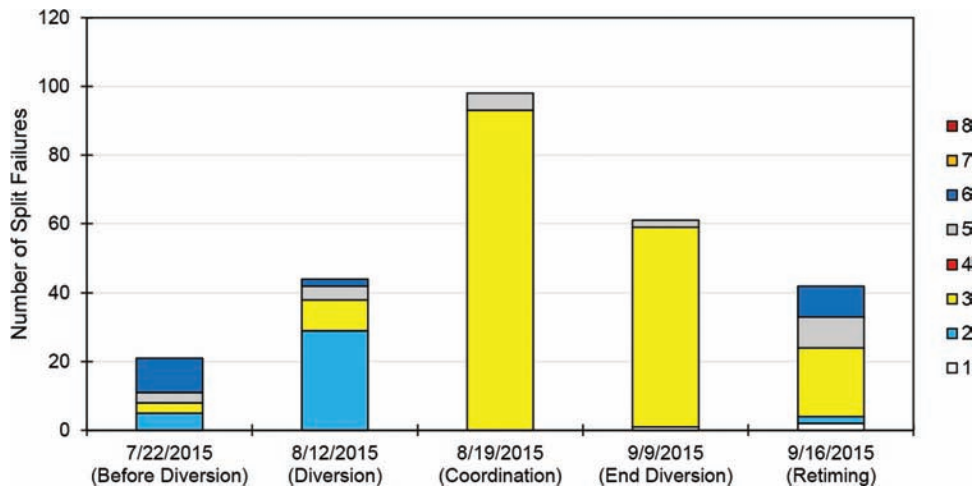


Figure 5.50 Split failures per phase at US 231 and River Road on different dates.

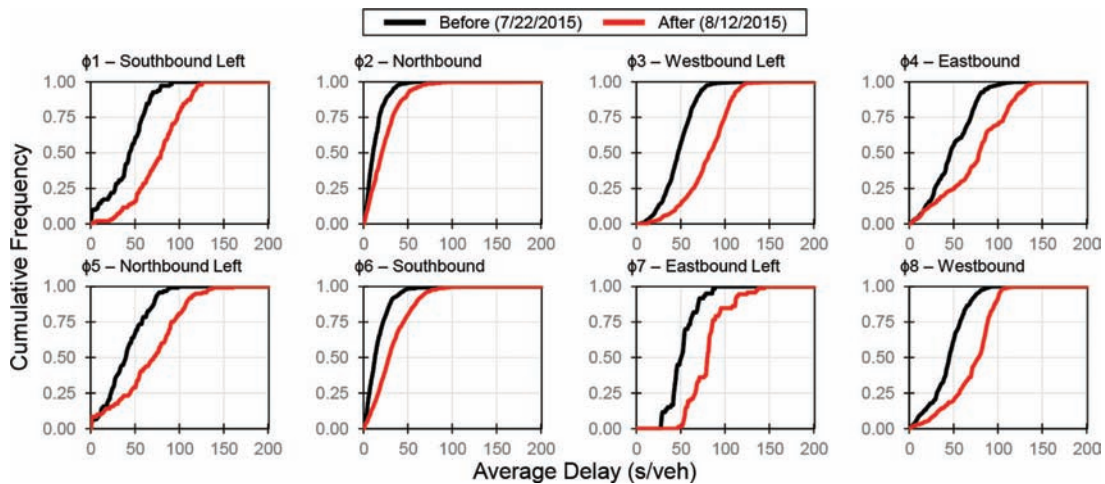


Figure 5.51 CFDs of cycle-by-cycle average delay by phase at US 231 and River Road, before I-65 diversion (7/22/2015) and during I-65 diversion running free (8/12/2015).

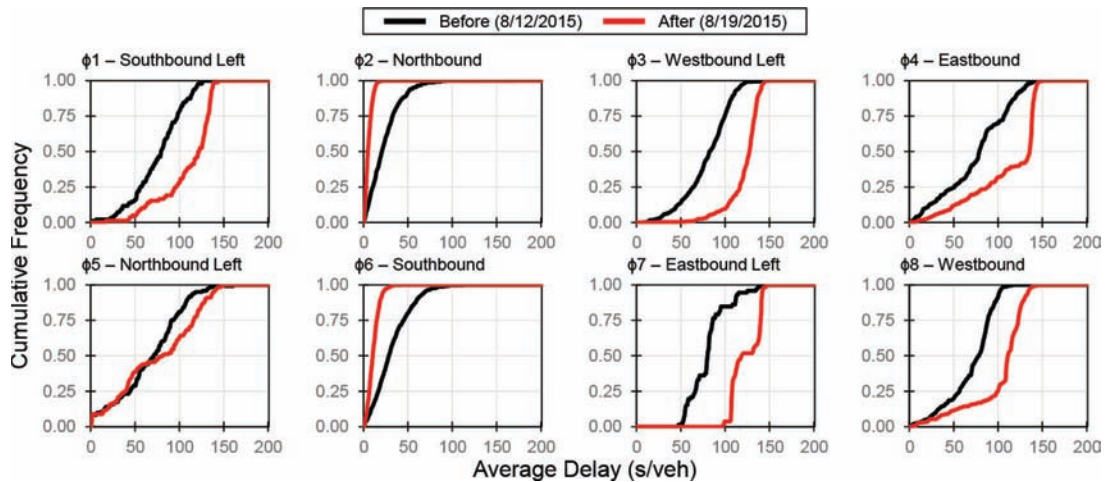


Figure 5.52 CFDs of cycle-by-cycle average delay by phase at US 231 and River Road, during I-65 diversion running free (8/12/2015) and during I-65 diversion running the detour coordination plan (8/19/2015).

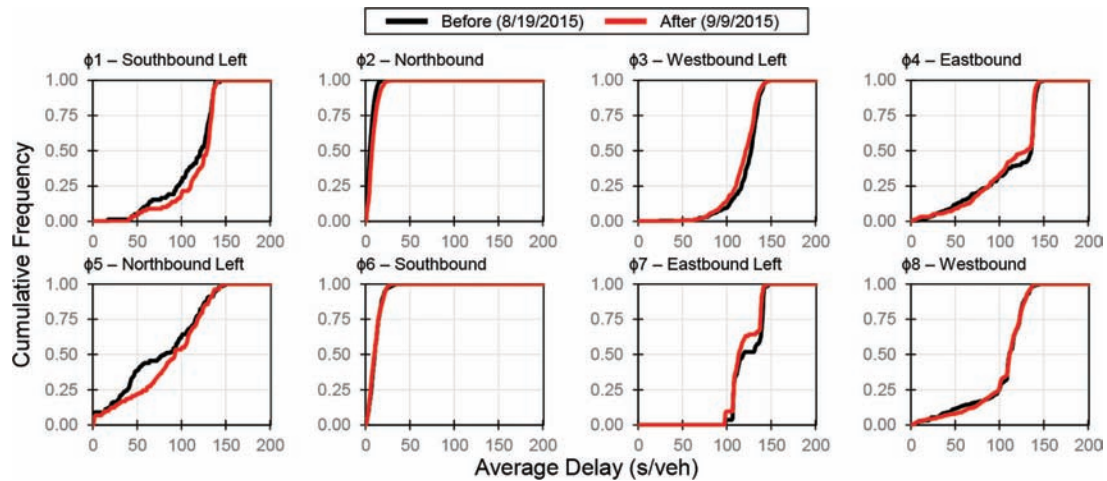


Figure 5.53 CFDs of cycle-by-cycle average delay by phase at US 231 and River Road, during I-65 diversion running the detour coordination plan (8/19/2015) and after the end of the diversion still running the detour coordination plan (9/9/2015).

average delay distribution is above 100 seconds per vehicle (s/veh) for all the noncoordinated phases except for phase 5.

- *Comparison 3: End of the detour* (Figure 5.53). Although the detour ended and the amount of traffic reverted to normal levels, the delay distributions for most of the phases were largely unchanged. The long cycle length of 150 seconds maintained average delays above 100 s/veh for nearly all the noncoordinated phases.
- *Comparison 4: Introduction of a new timing plan* (Figure 5.54). Retiming the corridor for normal traffic conditions saw cycle lengths reduce considerably. Although the delay on phases 2 and 6 increases slightly, it is still very low, with median values less than 25 s/veh. The median values of average delay for the other movements are reduced by about half. These are attributable to reductions in cycle length from 150 to 82–116 seconds.

5.5.6 Conclusion

The previous graphics examined a variety of operational details at a single intersection in great detail.

Such information would be valuable for any engineer operating under duress in similar circumstances. In fact, the ability to quickly validate timing plans after their implementation was put to good use during the I-65 detour. It was especially helpful for ensuring that the coordination was operating as intended, as Chapter 6 explores in more detail. As for the allocation of green times by split adjustments, the detour operation was constrained by the need to hold the northbound movement open. However, the post-detour timing plans were iteratively fine-tuned in the weeks after the reopening of I-65 using feedback from the data.

5.6 Example Network and Corridor Analysis

5.6.1 Overview

Up to this point, this discussion has focused on operations at a single intersection, examining relatively detailed aspects of the performance, such as how a

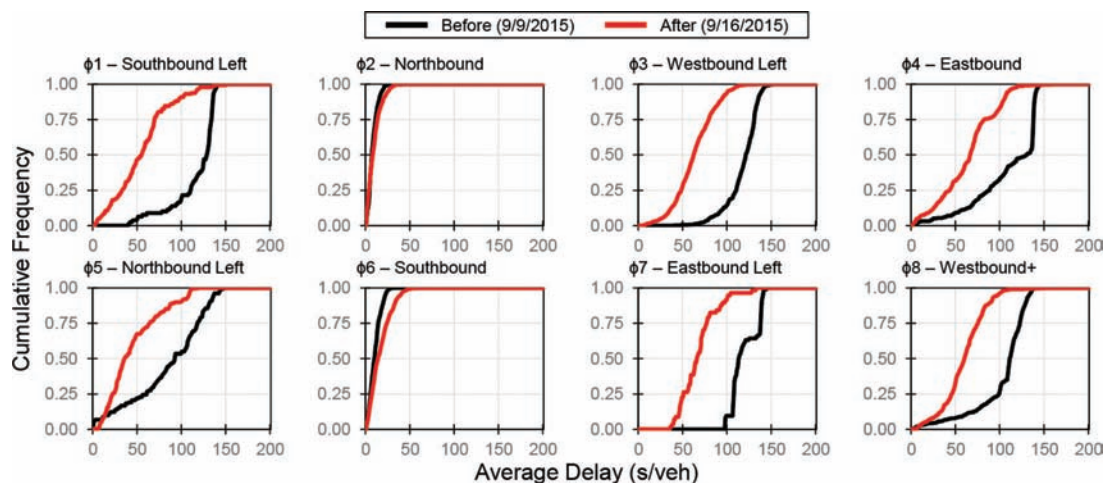


Figure 5.54 CFDs of cycle-by-cycle average delay by phase at US 231 and River Road, after the end of the diversion still running the detour coordination plan (9/9/2015) and after the implementation of a new coordination plan for normal traffic (9/16/2015).

certain movement operated during a certain day. While useful, it is not feasible for an analyst to hunt through the thousands of possible views to discover problems. It is essential to scale these metrics to an agency-wide perspective to effectively monitor the entire system. This final section moves from the intersection level analysis presented earlier and demonstrates the aggregation to corridor and network levels. Although any aspect of intersection performance could be given similar treatment, this discussion will focus on the occurrence of split failures.

A split failure was defined earlier as an occurrence where there is not enough green time during a signal cycle to serve the amount of demand present. Methods for identifying split failures were presented earlier for different detector configurations. Earlier, in the discussion of split failure severity, it was shown that different phases can have a very different signature with respect to consecutive split failures (Figure 5.19). To reiterate, the occasional split failure that is corrected in the next cycle is probably not an operational concern at the network level. To reduce the amount of noise in the data, it is helpful to use a threshold of three consecutive failures in a row [48]. This identifies a movement where

the failures are sustained over long periods of time, with the capacity never quite serving all the demand.

5.6.2 INDOT Network

For this evaluation, a study area was defined that included eight signalized arterial corridors in central Indiana operated by INDOT. Figure 5.55 shows a map of the locations of these corridors. All of the routes are coordinated for at least part of the day, with US 231 having seen numerous changes as the previous case study discussed. The others saw far fewer changes during the analysis period. This analysis focuses on split failure occurrences of these corridors during October and November of 2015.

The concept of the “traffic ticker” was recently developed as a way to visualize system information using probe vehicle data, as applied to freeway systems [85]. Figure 5.56 presents an example view from this dashboard, which shows the performance of the interstate highway system in Indiana. The chart, a stacked-area graph, shows the total number of interstate miles operating at speeds less than 45 mph, with different color series referring to different districts in the state.

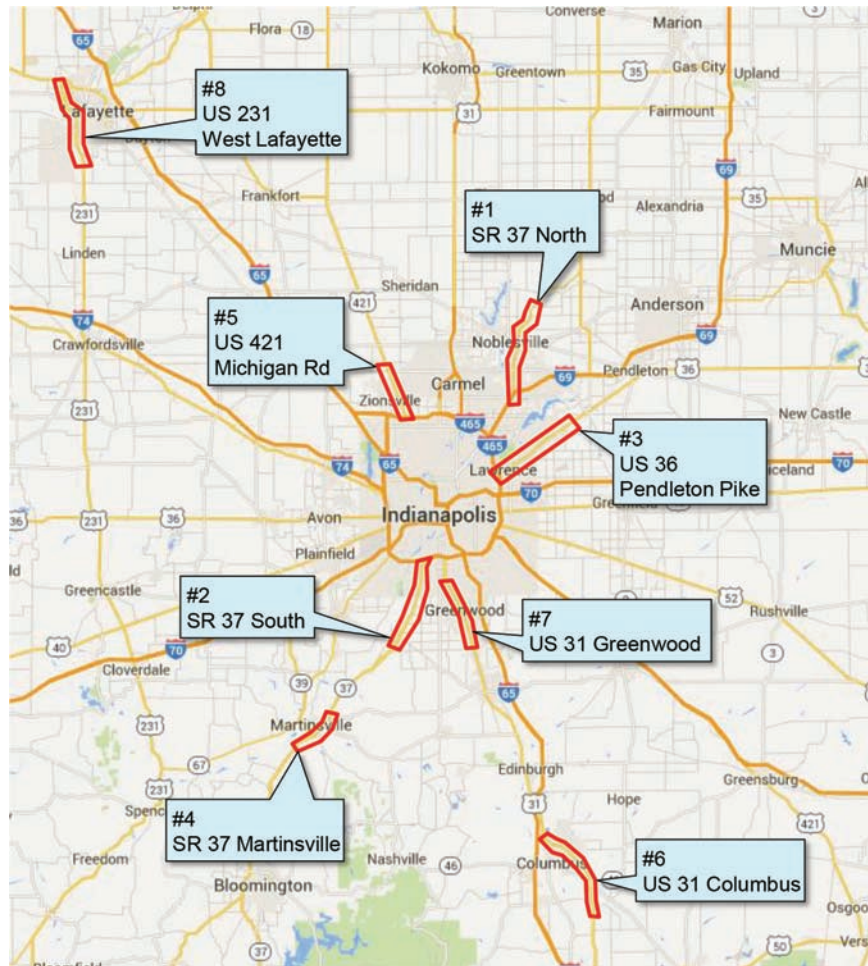


Figure 5.55 INDOT network map of online corridors for split failure analysis.

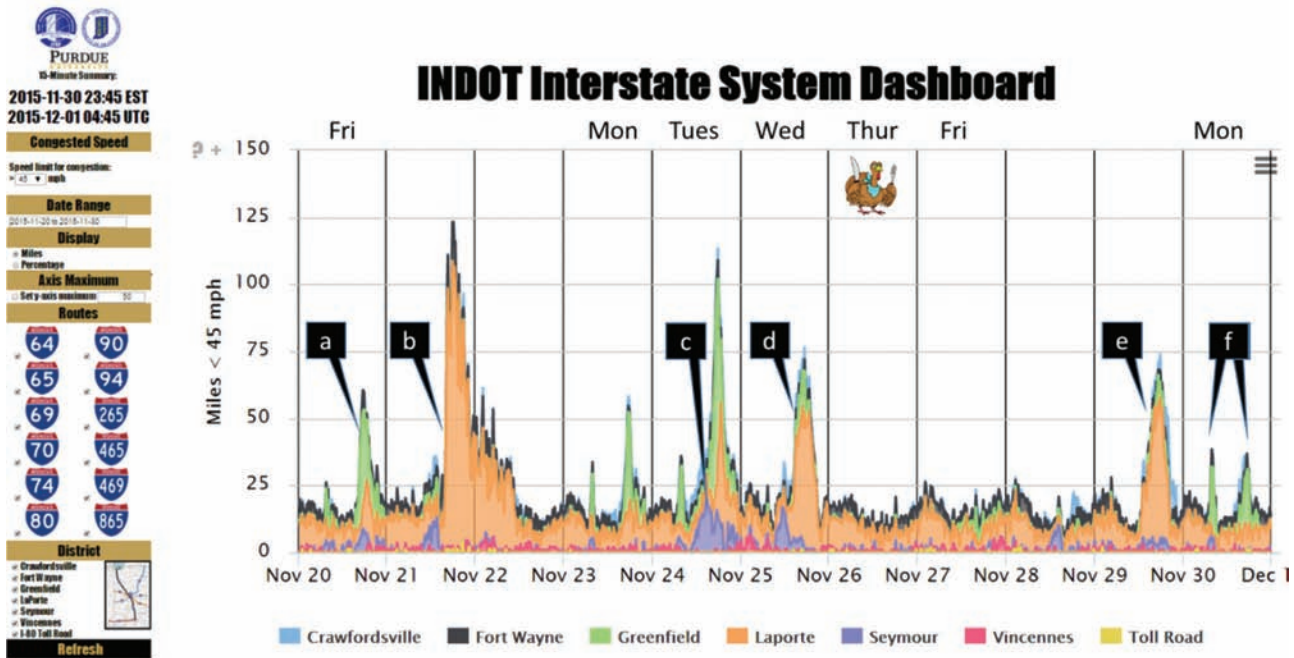


Figure 5.56 Example split failure ticker view showing data for 11/20/2015–12/1/2015 [85]. The graphic shows the sum of all interstate system miles operating under 45 mph per 15-minute bin over the dates shown. The data are colored by INDOT district.

The same concept can be applied to signal performance measures as well, and this is the idea behind the “split failure ticker.” Rather than charting interstate miles below a speed threshold, this chart plots the total number of split failures per category over user-defined dimensions of time. The categories and analysis dimensions are similar in concept to the notion of a cross tabulation, or of spreadsheet pivot tables. At the system level, the split failures can be categorized into different districts, regions, corridors, etc.. It is also possible to drill down to individual corridors, intersections, phases, lanes, or detectors as needed.

Figure 5.57 shows the total number of split failures per corridor by day from 10/1/2015 to 12/1/2015. Each bar in this chart shows the count for one day, and each individual color represents a different corridor subtotal. The horizontal lines demarcate weekends and weekdays. The horizontal axis can be expanded or contracted to look at longer or shorter timespans, from 15 minutes up to multiple years, if enough data are available. The total number of split failures represented in this chart are 153,264 across the eight corridors and two months. The overall trend in this chart seems to be a gradual decrease in the number of split failures per day as we approach the end of the year. The week of Thanksgiving (11/22/2015–11/28/2015) is rather interesting, with Thursday (11/26/2015) having a very small number of split failures, owing to the low amount of traffic, with the following day (“Black Friday”, 11/27/2015) having a much larger number.

Figure 5.58 shows an alternate arrangement of the data. Here, the count of split failures has been flattened across all of the dates, and instead the total count of split failures is shown by corridor and intersection

horizontally. The split failures are categorized by movement type using the color bins. This enables rapid visual identification of the worst performing intersections and movements. The two largest spikes show that Corridor #7, Intersection #1 and Intersection #2 both have an exceptionally high number of split failures on the westbound left turn, and Corridor #6, Intersection #4 has a high number on the eastbound left turn.

5.6.3 Corridor Analysis: US 31 Greenwood

The system-level views show that Corridor #7, US 31 Greenwood, exhibits the highest number of split failures in total and includes two intersections with a much higher amount of split failure than the others. Figure 5.59 shows a map of the corridor and the location of the ten intersections. This corridor is situated on the south side of the greater Indianapolis area. It serves as a major commuter route and serves many retail destinations, such as the Greenwood Park Mall on the north end of the system.

Figure 5.60 shows a split failure ticker that totals up daily counts of split failures per intersection during October and November of 2015. As before, the vertical lines demarcate weekends and weekdays. The graph reveals that Intersection #1 has the largest number of split failures, although these seem to disappear after 11/11/2015. Intersection #3 also has a high number, especially on Fridays and some Saturdays. This agrees with what was seen earlier in the system-level views. Referring back to Figure 5.59, these two prominent intersections are in the region of the system with heavy retail activity.

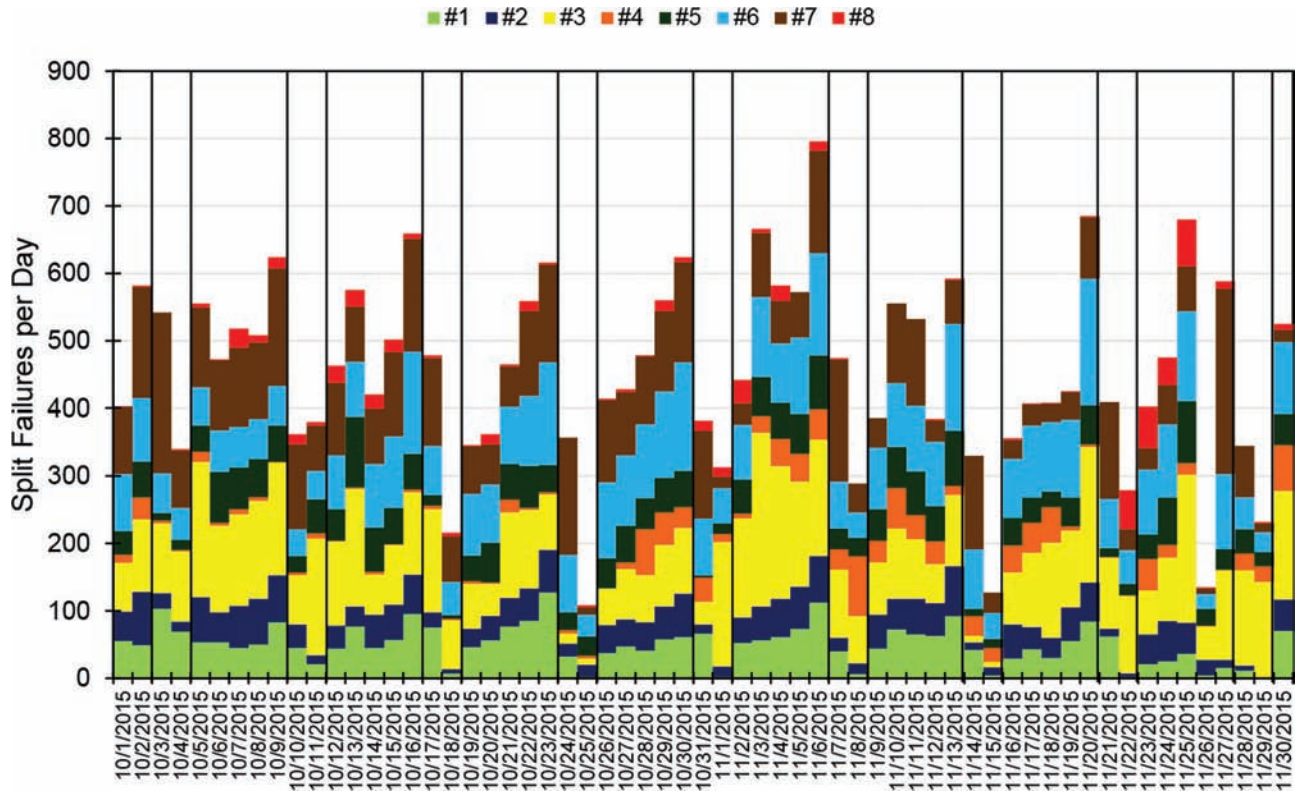


Figure 5.57 System-wide count of split failures per corridor by day from 10/1/2015 to 12/1/2015.

The same data are presented in a slightly different manner by Figure 5.61. Here, rather than plotting by date, the counts of split failures are shown by time of

day for all dates in October and November 2015. They are segmented by intersection as before. The graph shows that the split failures at Intersection #1 are

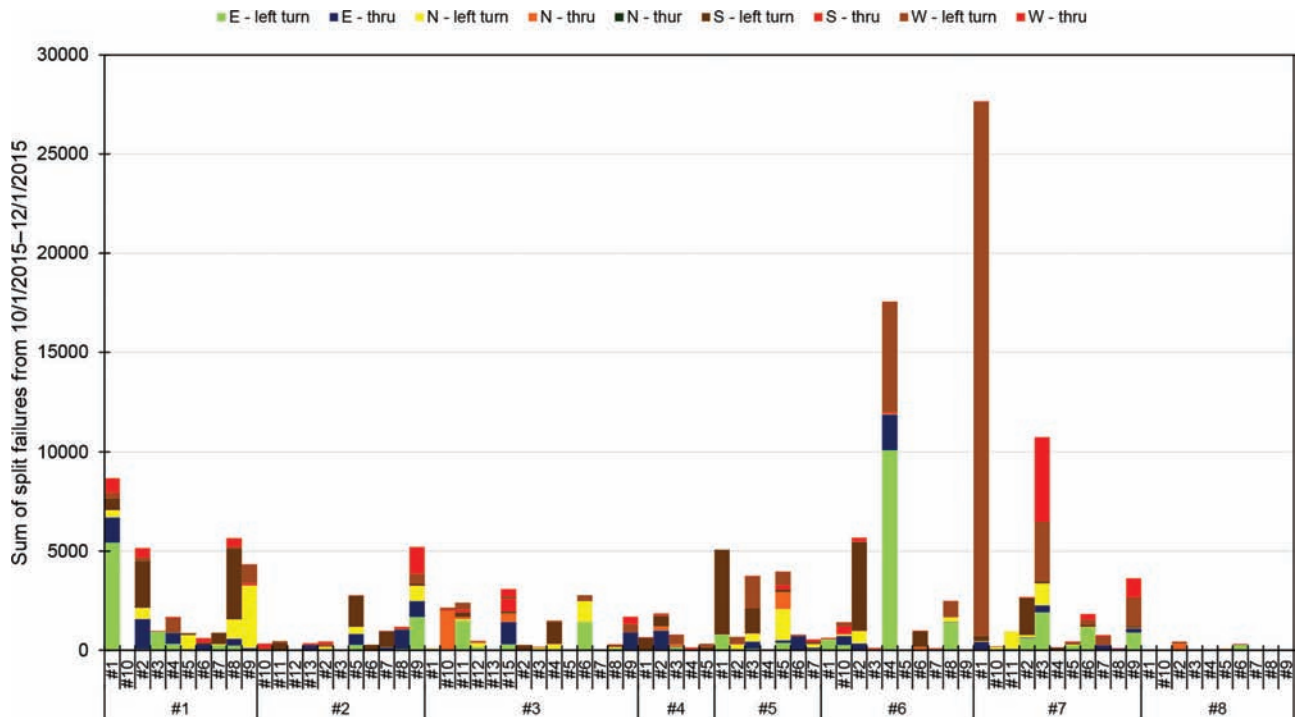


Figure 5.58 System-wide count of split failures by intersection and movement from 10/1/2015 to 12/1/2015.



Figure 5.59 Map of US 31 in Greenwood (Corridor #7).

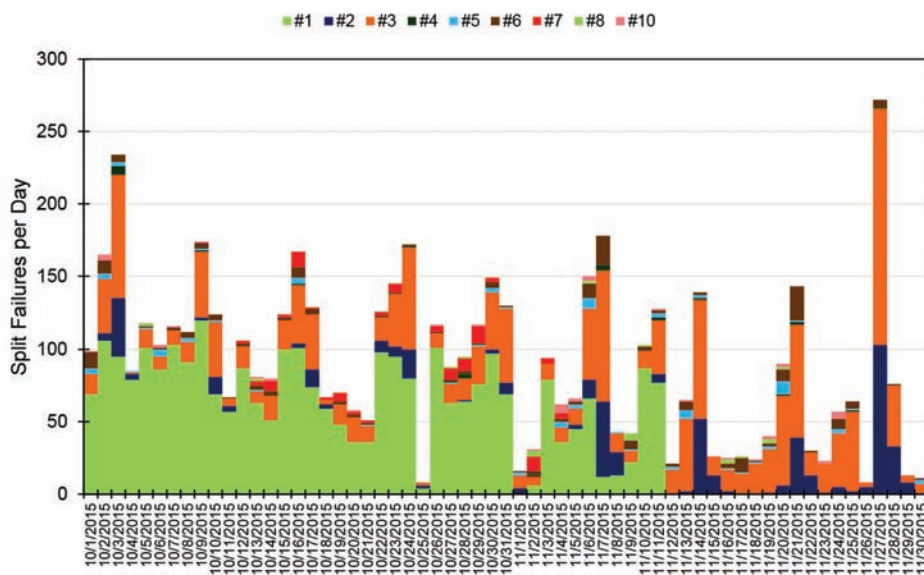


Figure 5.60 Count of split failures per intersection by day on US 31 Greenwood, 10/1/2015–12/1/2015.

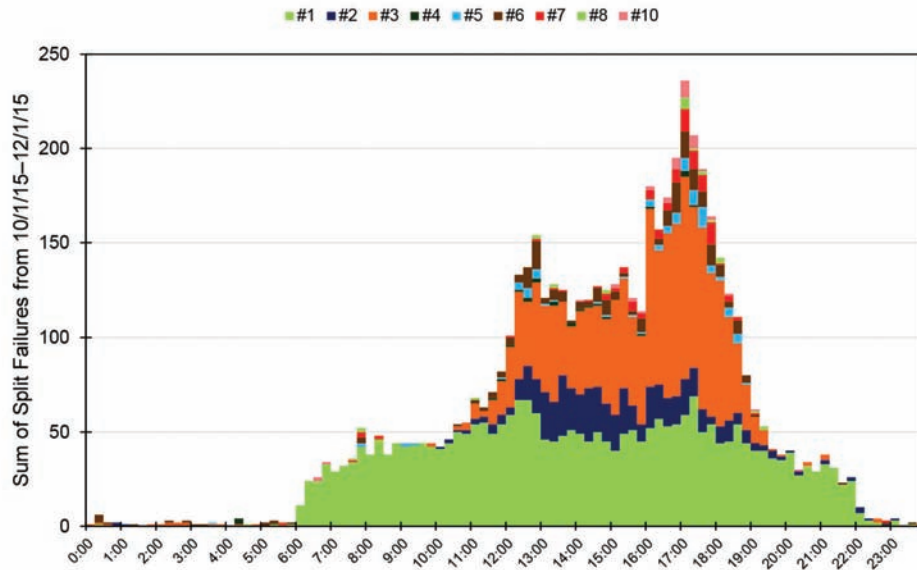


Figure 5.61 Split failures per intersection by time of day on US 31 Greenwood, 10/1/2015–12/1/2015.

distributed through the entire day from 6:00 to 22:00, while those at Intersection #3 are typically in the midday and evening. There were 2,821 split failures in the study period at Intersection #1 and 1718 at Intersection #3.

The ticker view can drill down to the intersection level to identify the individual movements that are having the split failures. Intersection #1 and #3 are candidates for further investigation. Figure 5.62 presents a view of split failures per movement by day at Intersection #1. The graph clearly shows that the westbound left turn is the main source of problems at this intersection, with a few

others contributing a far smaller number of split failures to the intersection total. The count of split failures disappears after 11/12/2015, which suggests that there was a communications outage to the intersection after this date. This underscores the importance of maintaining good communications, as discussed in an earlier chapter. Unfortunately, there is no record of how this intersection performed in the last two weeks of November, but one can easily imagine that split failures continued on the westbound left turn. Figure 5.63 focuses on a particular date, 10/30/2015, which finds that the westbound left turn split failures are distributed

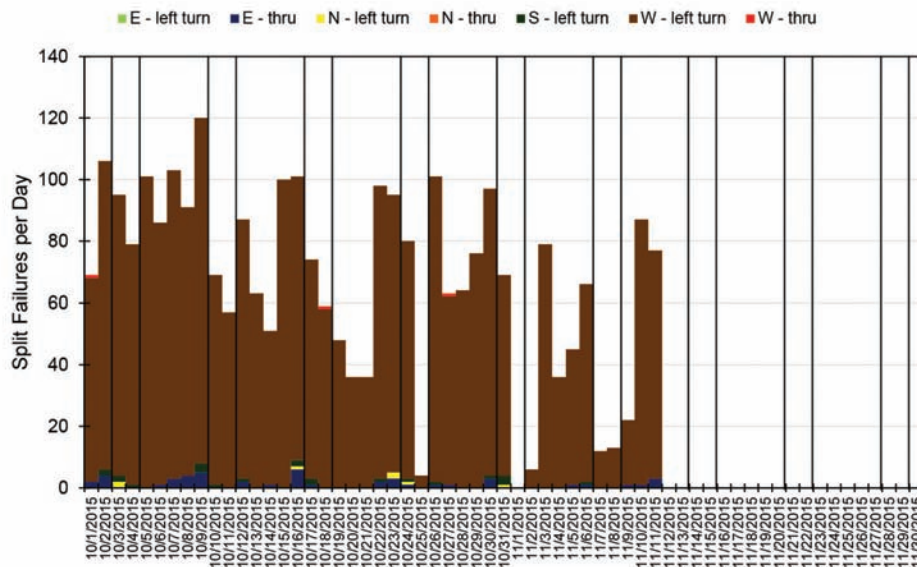


Figure 5.62 Split failures per movement by day at Intersection #1, US 31 and County Line Road, 10/1/2015–12/1/2015.

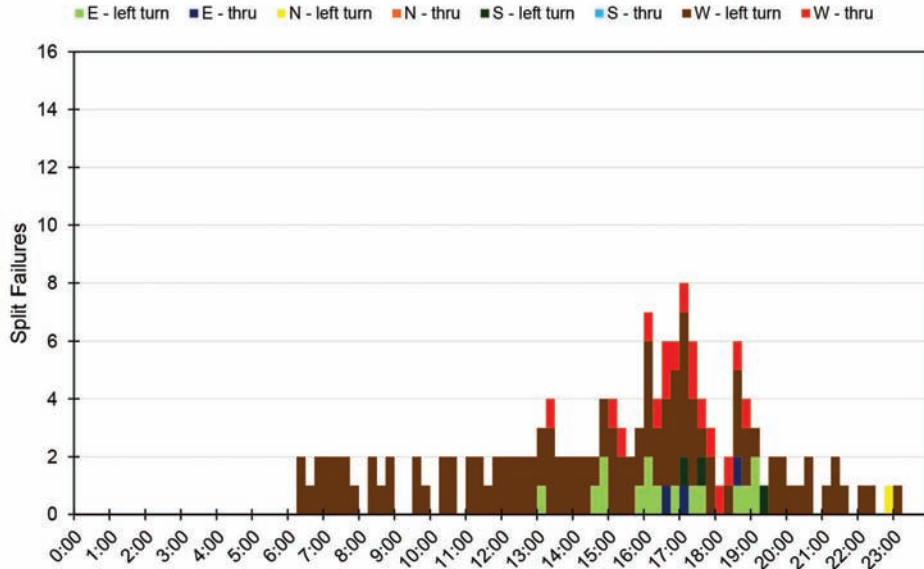


Figure 5.63 Split failures per movement by time of day at Intersection #1, US 31 and County Line Road on 10/30/2015.

throughout the entire day, whereas other movements have failures mostly during the PM peak.

Figure 5.64 shows the number of split failures per movement by day at Intersection #3. This intersection has several movements that have a fair amount of split failures: the westbound thru, westbound left, and east-bound left—in other words, three of the four side street movements. The dynamics are a bit different at this intersection and relate to traffic generation by nearby retail businesses. Fridays and Saturdays typically have considerable congestion, and the “Black Friday” (11/27/2015) following Thanksgiving has an exceptional number of split failures. Figure 5.65 examines 24-hour operation

on that date. Most of the split failures are distributed between 10:00 and 19:00.

5.6.4 Conclusion

The methods shown here demonstrate a way to transform very large quantities of high-resolution signal data into information that engineers can make decisions on. The split failure ticker enables drilling down from the overall system level down to individual intersections, quickly locating problem movements from a large inventory of assets. The times of day when they occur and their distribution over time can also be determined.

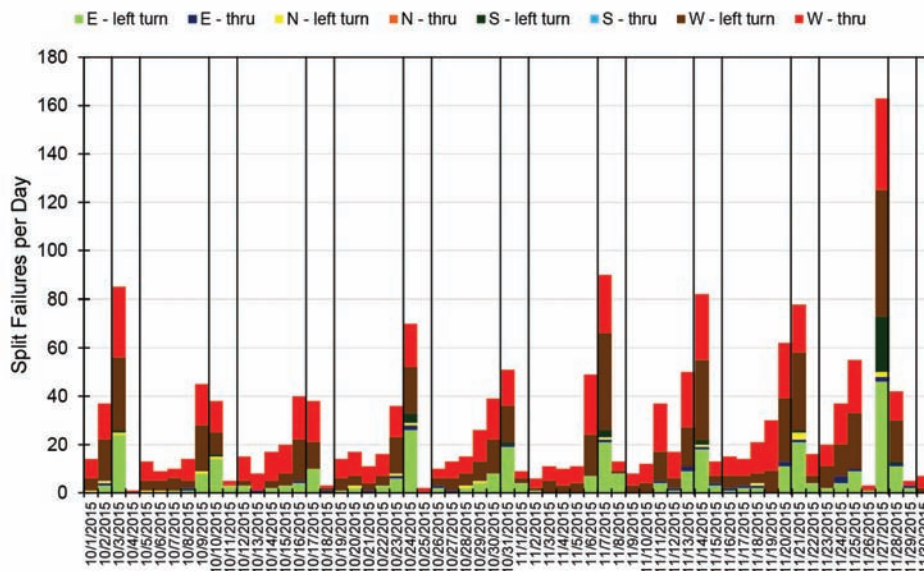


Figure 5.64 Split failures per movement by day at Intersection #3, US 31 and Fry Road, 10/1/2015–12/1/2015.

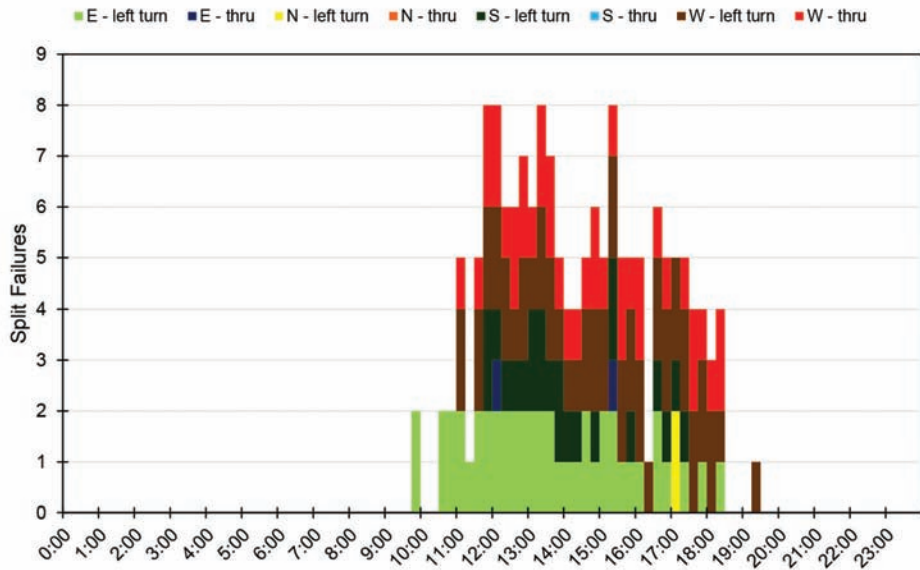


Figure 5.65 Split failures per movement by time of day at Intersection #3, US 31 and Fry Road, 11/27/2015 (Black Friday).

One can imagine scaling the procedure out to an entire agency’s inventory of signals, incorporating hundreds of intersections, with the data assisting engineers in managing the corridors they are responsible for as well as supporting executive reports at the top level that show the overall system performance.

5.7 Chapter Conclusion

This chapter demonstrated several uses of high-resolution data-enabled performance measures to assess the operations of individual intersections. Performance measures to characterize the level of demand and the degree of capacity utilization were presented through applications to several use cases. The use of high-resolution data to support traffic control diagnostics and

assessment of advanced control features was demonstrated. Finally, the value of having such data collected over many locations over a long period of time was demonstrated using longitudinal views of the data as in the “split failure ticker,” which uses a pivot table style approach to arranging data for analysis needs. Such an approach would lend itself well to other performance measures beyond counts of split failures.

Using a systematic method for identifying operational deficiencies across the system will enable engineers to better allocate resources for retiming signals. Another use case is the preparation of executive reports, to help demonstrate accountability and to document the overall state of the system. Figure 5.66 shows an example of an agency-wide summary that is part of the UDOT Signal Performance Measures website. This presents several

Executive Summary						
12/2/2015 to 12/2/2015						
Statewide Summary (intersections with upstream detection)						
Arrival on Red		Delay		Volume	Intersections	
Percent	Platoon Ratio	Daily Average Per Approach (hrs)	Average Per Veh (sec)	Daily Average Per Approach	Total	Number Of Approaches
28.9 %	1.11	0.02	7.15	12,561	498	1,014

Region Summary							
Region	Arrival on Red		Delay		Volume	Intersections	
Name	Percent	Platoon Ratio	Daily Average Per Approach (hrs)	Average Per Veh (sec)	Daily Average Per Approach	Total	Number Of Approaches
1	27.7 %	1.10	0.07	6.32	13,200	172	335
10	NaN	NaN	NaN	NaN	NaN	0	0
2	31.1 %	1.15	0.07	8.30	12,315	198	427
3	27.4 %	1.06	0.10	6.36	12,920	112	222
4	21.0 %	1.14	0.34	5.70	6,427	16	30

Figure 5.66 Example of an agency-wide summary report, from the UDOT Performance Measures website [49].

performance measures for the entire state network and a breakdown of these per region. The user can then drill down to individual corridors and intersections if desired. One goal of detailed analysis will be to prepare such a summary report according to the needs of the agency.

One future step of research in the area of local control performance measures will be to begin suggesting changes to signal control based on the data. One can observe that some movements have many split failures, whereas others have relatively few; this represents a potential opportunity for rebalancing the distribution of green times [48]. A system that can integrate performance measures into the signal timing design process would be a powerful addition to the suite of existing tools. Whereas detector-based actuation and the analogous processes in other control methods tailor the signal response to demands at a microscopic level, a tool to *design* the control policies that govern those processes would do much to integrate these local control process into overall system management.

6. EVALUATION OF SYSTEM CONTROL

6.1 Overview

High-quality traffic signal progression is one of the key objectives of traffic signal system operations, especially for arterial corridors. This concerns the cooperative scheduling of green times at neighboring intersections to establish a pattern of traffic progression through multiple signalized approaches. The previous chapter examined aspects of signal performance from the perspective of local control efficiency. This chapter presents applications of high-resolution data and travel time data for managing arterial progression, including performance measures to assess the quality of progression and data-driven techniques to optimize signal offsets.

6.2 Basic Traffic Progression Concepts

The premise of signal coordination is simple: a group of vehicles passing through an intersection at some time will require a certain amount of time to arrive at the next intersection, and ideally the signal will be providing a green indication for the appropriate movement at that time.

A sample of real-world vehicle trajectories is presented in Figure 6.1 in a time-space diagram format. These trajectories were constructed using software that tracked them across multiple video cameras along an arterial [86] and are representative of actual driver behavior in a signalized system. The process of queue formation at the red signals is evident, as well as some slight dispersion of platoons as they traverse the links. The trajectory and signal phase data together contain all of the most relevant information needed to evaluate the quality of progression. However, such data are almost impossible to obtain on a widespread basis. Rather, the signal phases are known, and arrival times can be sampled using detection.

The process of designing coordinated timing plans is based on modeling, and the time-space diagram is the basic design tool. Figure 6.2 shows some variations on this visualization in various signal timing design software tools. Figure 6.2a shows a time-space diagram with green bands projected forward from each stop bar when the local signal is green. Figure 6.2b shows another view where the space between intersections is colored by a probabilistic distribution of when vehicles are expected to arrive in the cycle. As the vehicles move along the roadway, the distributions appear to disperse, modeling how platoons tend to spread out over time and distance from their formation. Finally, Figure 6.2c presents another view, with traffic flows shown by trajectory lines. Every aspect of these views is based on an underlying traffic model: the speed of the vehicles

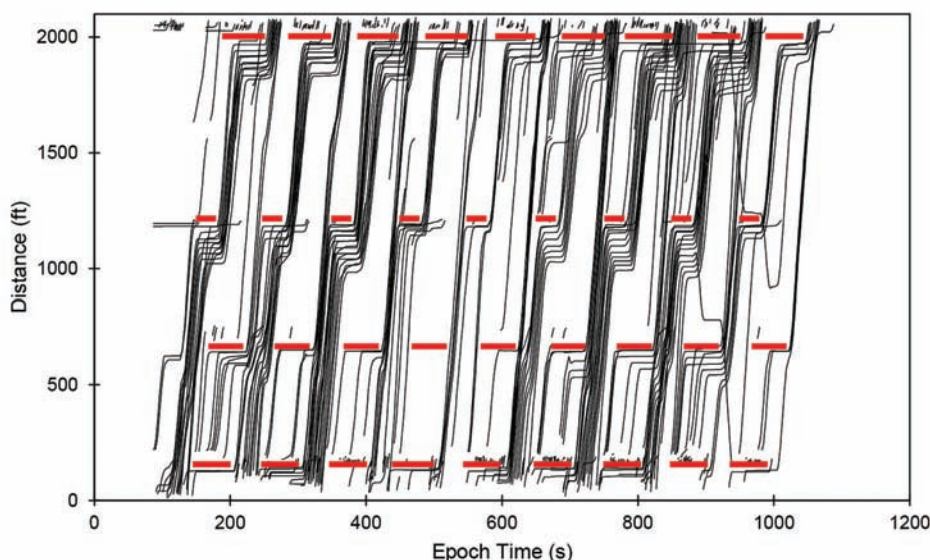
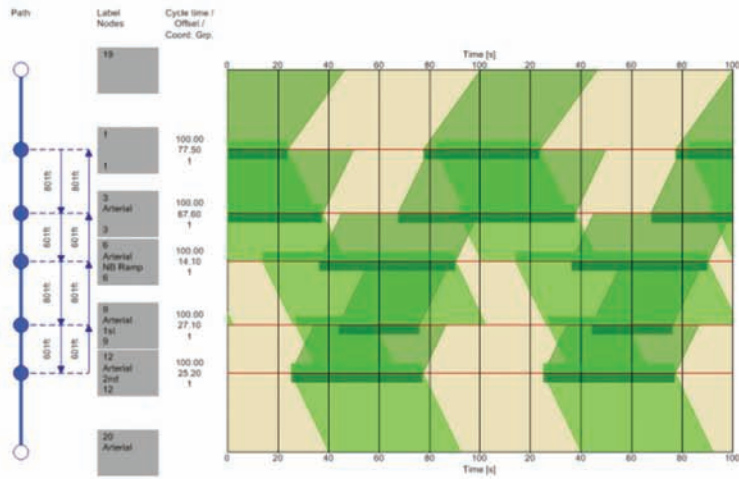
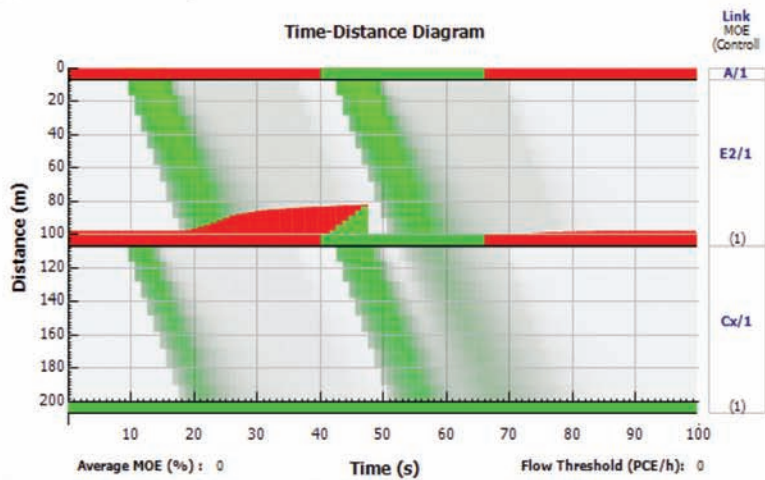


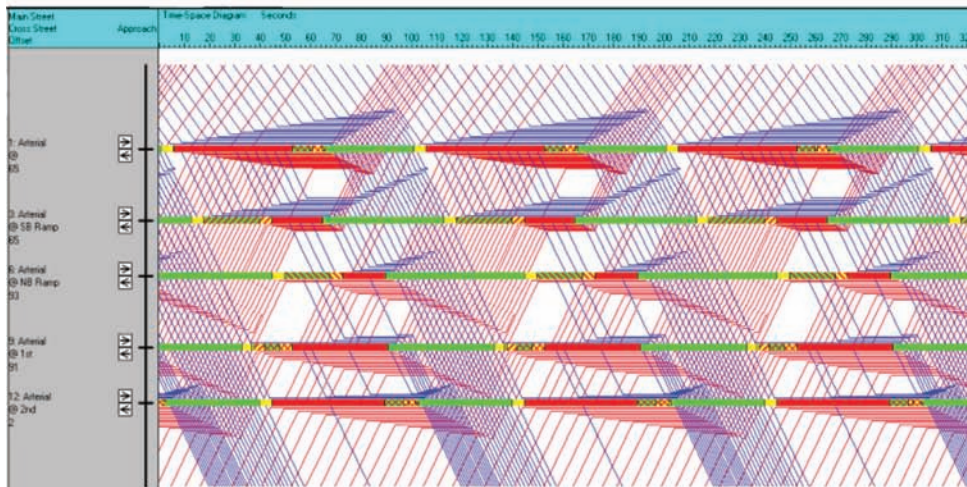
Figure 6.1 Example arterial trajectory data from the NGSIM Peachtree Street dataset [86].



(a) Visualization of green bands in Vistro.



(b) Representation of macroscopic flows in Transyt 15.



(c) Representation of flows in Synchro.

Figure 6.2 Example of time-space diagram views in design software.

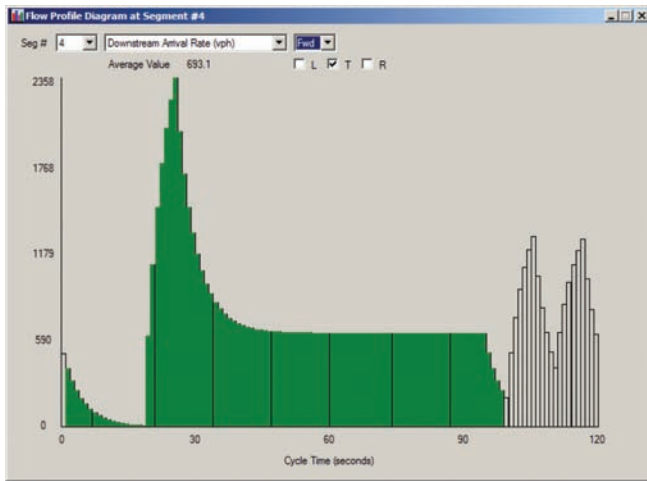


Figure 6.3 A cyclic flow profile from Transyt-7F.

along the roadway, the effective green times resulting from actuation, the amount of platoon dispersal, and so forth. Therefore, for the signal timing to be effective, the model assumptions must be a close fit to reality.

Figure 6.3 shows a different modeling visualization called a cyclic flow profile. The distribution shows the number of vehicles expected to arrive at a certain time in the cycle. The bars colored green are the vehicles

expected to arrive while the signal is green, and the uncolored bars are those expected to arrive while the signal is red. This graph represents a probabilistic view of the dynamics occurring at a signalized approach, and illustrates the two fundamental pieces of information needed for measurement. The arrival profile, in this case modeled using a platoon dispersion model, could instead be directly measured using detection. The probability of green, in this case a simple binary “green” or “not green” condition, could be measured using the actual phase states. The same basic data can also facilitate other visualization tools, as this chapter will show.

6.3 Evaluation of the Quality of Progression

Table 6.1 lists alternative performance measures that evaluate the quality of progression using data from actual signal operations, including, but not limited to, high-resolution event data.

Using phase data alone, it is possible to view the actual bands in a time-space diagram format. Some central systems provide the ability to do this using status data streaming from signal controllers. This can be helpful for validating whether the system is operating as expected, but is rather limited in the scope that it can depict. It becomes difficult to see dynamics beyond about 10 cycles on a screen, so the view is best suited for spot checking. Floating car data are similar in

TABLE 6.1 Performance measures for evaluating the quality of progression.

Performance Measure	Compatible Data/ Detection Types	Criteria for Good Progression	Notes
Time-Space Diagram	Phase States	Visual analysis	
Floating Car Travel Times	GPS data	Traces indicate minimal slowing or stopping.	Very labor-intensive, and usually limited to a very small observation window.
AVI Travel Times	AVI Sensors	Indicated by lower travel times.	Sensitive to sensor location.
Probe Data Segment Speeds	Third-Party Probe Data	Indicated by higher speeds.	Sensitive to segment definitions. Not accurate for high signal density or lower volume roads.
Percent on Green (POG)	Setback Detectors and Phase States	Indicated by high values.	Sensitive to queuing.
Platoon Ratio	Setback Detectors and Phase States	Indicated by values greater than 1.0	Sensitive to queuing.
Arrival Type	Setback Detectors and Phase States	Indicated by values greater than 4.0	Sensitive to queuing.
Estimated Delay	Setback Detectors and Phase States	Indicated by low values.	Sensitive to queuing.
Flow Profile	Setback Detectors and Phase States	Visual analysis, and correlation with other performance measures	Requires a background cycle length. Sensitive to queuing.
Purdue Coordination Diagram (PCD)	Setback Detectors and Phase States	Visual analysis, and correlation with other performance measures	Sensitive to queuing.
Virtual Trajectory	Setback Detectors and Phase States	Indicated by high estimated travel time.	
Trajectory Data	Connected Vehicle	Traces indicate minimal slowing or stopping.	Requires high penetration rate before useful detailed data can be obtained.

that they only cover a snapshot of the signal operation during a small amount of time.

Externally automated collection of vehicle data can be used to measure travel times and speeds along arterial roadways. Automatic vehicle identification (AVI) data, such as Bluetooth MAC address matching, is one option. Third-party probe data from commercial providers currently consist of speed measurements on predefined roadway segments. These data have been used widely for evaluating freeway systems, but have been shown to be useable for at least some arterials, particularly those with a low density of traffic signals and a high traffic volume [87].

High-resolution data can support all of the performance measures that require “Setback Detectors and Phase States,” and these will be the main focus of this chapter. These include numerical performance measures that are calculable for individual cycles (percent on green, platoon ratio, arrival type, and estimated delay) and visualization tools that show the cyclic relationship of individual events (flow profile and coordination diagram).

The last two items in the Table 6.1 are methods for building views of vehicle trajectories as in Figure 6.1. The virtual probe vehicle trajectory concept [12] relies on developing an estimate of queue lengths at all of the intersections in the system. This, combined with the phase states, is used to draw likely trajectories for different start times. These can estimate the likely travel times through the system. With connected vehicle data, it may become possible to eventually begin sampling actual vehicle trajectories. Some of the raw data used by third party data providers will likely become available for that purpose in the near future.

This report will focus primarily on uses of high-resolution data, with independent outcome assessment using travel times measured either by AVI methods or probe data segment speeds. There are two important considerations relevant to the high-resolution data.

- It is important to be aware of potential queuing over the detector. If queues at the intersection extend past the advance detector, then the arrival data measured during that time will not be valid. Rather than measuring vehicle arrivals, the detector will actually be measuring vehicle movement during queue discharge. Monitoring the detector occupancy time can be helpful in identifying when queuing takes place. In the various corridors in Indiana where these data techniques were tested, the detector locations were set at the appropriate location for dilemma zone protection (e.g., 405 ft from the stop line in a 55 mph zone). With this configuration, queuing did not present a problem, except for a few critical intersections during the peak periods.
- It is not possible to accurately predict changes to offsets if the controller clocks at the neighboring intersections are not properly synchronized. Therefore, it is important that controller clock times are kept up-to-date. Usually, this is the case in coordinated systems. However, in our experience, we have found that it is important to verify that synchronization is indeed successful.

At some locations, stop bar detection may be available instead of advance detection. Although stop bar detection can be very useful in making determinations about capacity utilization, unfortunately it is not possible to measure vehicle arrivals except for the very first vehicle entering an unoccupied detection zone. Therefore, simply because an intersection has detection available on an approach does not necessarily mean the detection can be used for progression analysis. The detectors must be located sufficiently upstream of the stop line to measure vehicle arrivals. If a stop bar detector is mistakenly used, most of the vehicle “arrivals” will all appear as though they are occurring during green, whereas what the detector actually measures is the departure flow.

6.3.1 Qualitative Visualization

Figure 6.4 illustrates how the detector and phase states can be translated into two different types of visualizations. To begin, Figure 6.4a shows a plot of vehicle trajectories in a time-space diagram. The stop bar is located at 3000 ft, and the red state of the signal is indicated using red bars. End of green (EOG) and beginning of green (BOG) are the relevant phase events; these are projected backward from the stop bar using the approach speed. The detector is located at 2600 ft in this case, and the blue line traced horizontally across the chart represents the measurement location. The dots appearing on this line are at the intersections with vehicle trajectories; these are represented in the high-resolution data as detection times at the setback detector.

Combining the detection times with the phase times, it is possible to develop a “coordination diagram” that concisely illustrates the situation for this particular cycle. This is shown at the top of the graph. Here, the vehicle arrivals are projected forward to the phase states at the time when they arrive at the intersection, and are plotted at the estimated time in cycle when they arrive. In this case, most of the vehicles arrive as a platoon during green. That group of relatively straight-line platoons that experience little delay are represented by the cluster of dots within the green region of the diagram.

The advantage of this view is that it visualizes several cycles at once when multiple cycles are situated side-by-side, as illustrated by Figure 6.4b. This graphic is known the “Purdue Coordination Diagram” (PCD) [19]. This chart plots time of day on the horizontal axis and time in cycle along the vertical axis. Each vertical strip in the data shows a single cycle. BOG is represented by the green line and EOG by the red line. The EOG line also shows the cycle length. The shaded region between BOG and EOG shows when the signal is green, and each dot represents a single detection event, projected forward from the detector location. In this example, the vehicle arrivals are coincident with green during coordination, which lasts from 6:00 to 22:00. During the other portions of the day, the signal runs free and there is no cyclic arrival pattern.

Data from the time period spanning 12:00–15:00 are extracted and processed into a cyclic flow profile, which

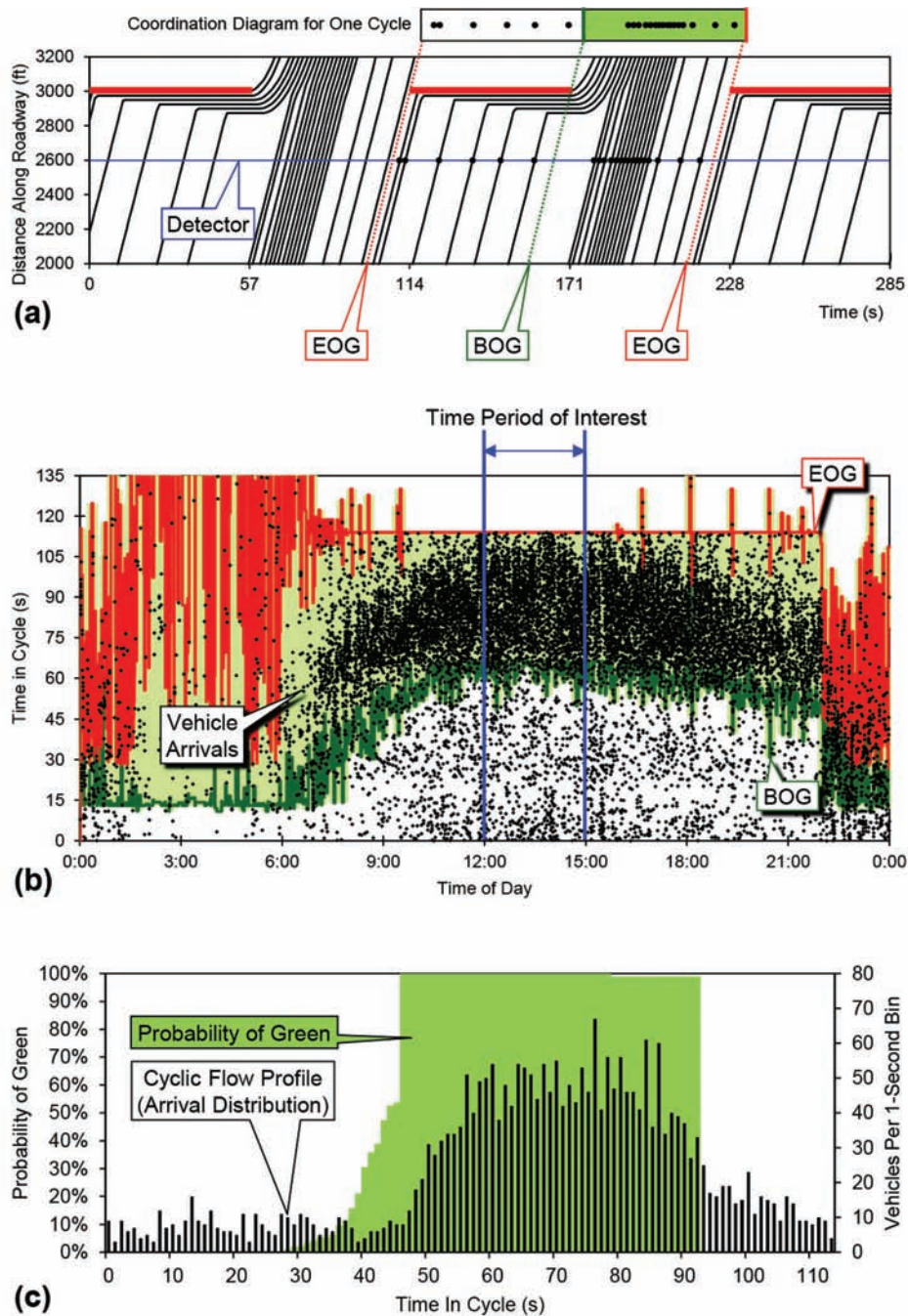


Figure 6.4 Three graphical representations of vehicle arrival and phase status data based on conventional (complete) detection [89]: (a) a time–space diagram, and a coordination diagram for one cycle; (b) a Purdue Coordination Diagram (PCD) covering 24 hours; (c) a cyclic flow profile showing arrival and probability of green distributions.

is shown in Figure 6.4c. This plot consists of two elements. The probability of green shows the percentage of cycles in which the signal was green during that particular time in cycle. The arrival distribution (the part of the graph called the “cyclic flow profile” in its original conception [88]) shows the total number of vehicles arriving during that time in the cycle. The bulk of the distribution represents vehicle platoons. Good

progression is typified by coincidence of these platoons with the portion of the cycle where the probability of green is high.

This chapter will focus more on use of PCDs than on flow profiles, but both are useful tools for visualizing the quality of progression. The advantage of the PCD is that multiple times of day can be visualized, and microscopic events can be scrutinized. The advantage

of the flow profile is that it more succinctly describes the arrival and service patterns.

6.3.2 Quantitative Evaluation

Both the PCD and the flow profile support the computation of quantitative performance measures that can summarize the overall performance in the graphic. This is useful for showing total system performance in a quick view, without having to dig into the details. Several such numbers can be reported:

- **Arrivals on Green (AOG)/Percent on Green (POG).** The total number of vehicles arriving on green is a useful metric for quantifying the quality of progression, especially when expressed as a percentage. Increases in POG are useful for quantifying the effect of a signal timing change. Figure 6.5 shows a view of PCDs as implemented in the INDOT signal metrics website. Six different approaches are shown for three intersections, and the AOG and POG are indicated for the entire 24-hour period shown.
- **Arrivals on Red (AOR)/Percent on Red (POR).** This is the opposite of AOG/POG, using a number where more means worse, similar to split failure analysis. This can be useful when trying to identify problems across an broad inventory of locations.
- **Platoon Ratio.** This metric divides the POG by the green-to-cycle ratio. This accounts for the fact that the greater the percentage of time that the signal is green, the more likely it is for vehicles to arrive on green. The metric

therefore tends to reward high POG that is achieved with smaller greens and scales down the results of high POG that is achieved with extremely long arterial green times. A platoon ratio of 1.0 represents random arrivals; higher numbers represent good progression, while lower numbers represent poor progression.

- **Arrival Type.** This is a Highway Capacity Manual metric that is based on the Platoon Ratio concept, and scales the value to a 1–6 scale [15,90,91]. A category of “1” represents very poor progression and “6” perfect progression. A category of “3” represents random arrivals. Values in the range 4–5 are typical for most well-coordinated systems.
- **Estimated Delay.** Using the input–output method described in Chapter 5, it is possible to compute a delay estimate using arrival profile data.

Figure 6.6 shows an implementation of the PCD graphic on the UDOT Signal Performance Measures website. The PCD has been combined with a black line that shows the hourly volumes. Each TOD pattern includes several summary statistics, including the percent arrivals on green (“AoG”), the percent of cycle used for green on the phase shown (“GT”), and the platoon ratio (“PR”).

Figure 6.7 shows an example corridor summary from the UDOT Signal Performance Measures website, in which several of these quantitative metrics have been summarized for all of the signalized approaches within each corridor. These are coupled with vehicle volumes and the number of signalized approaches, which give

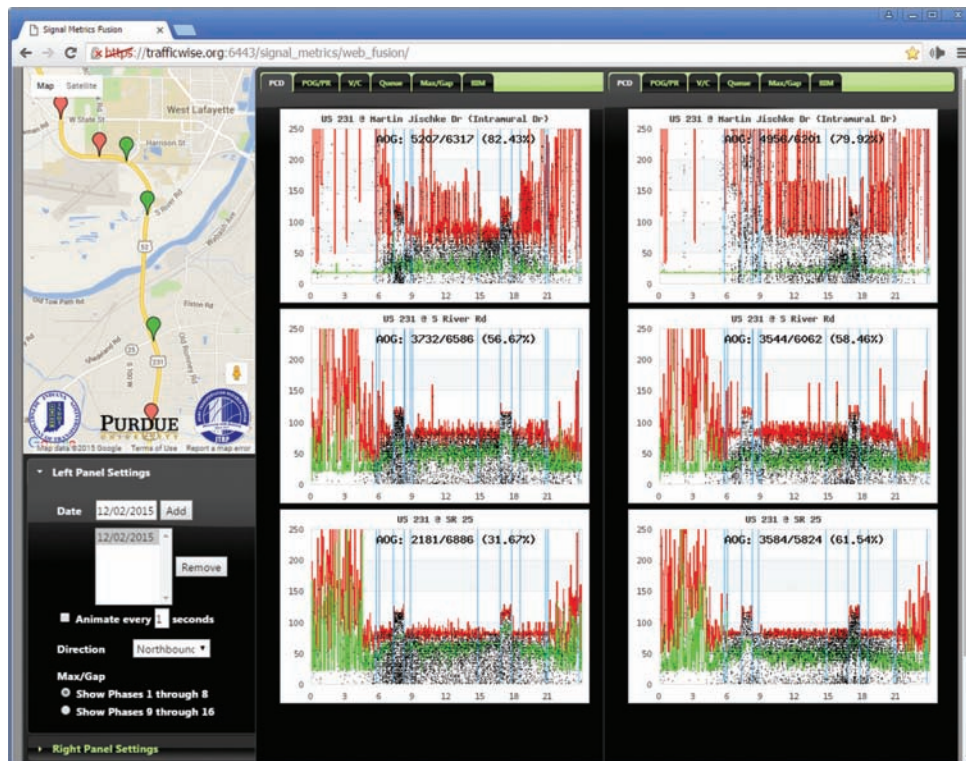
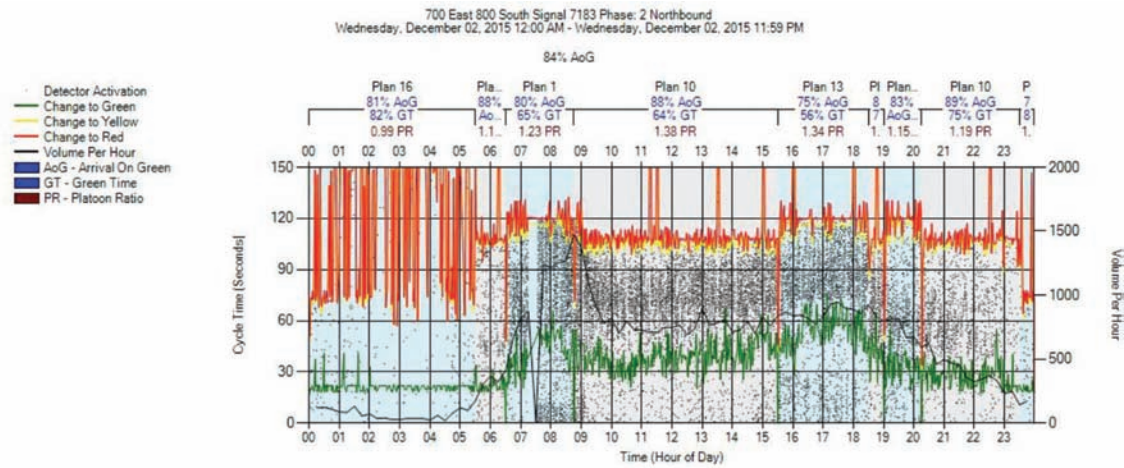
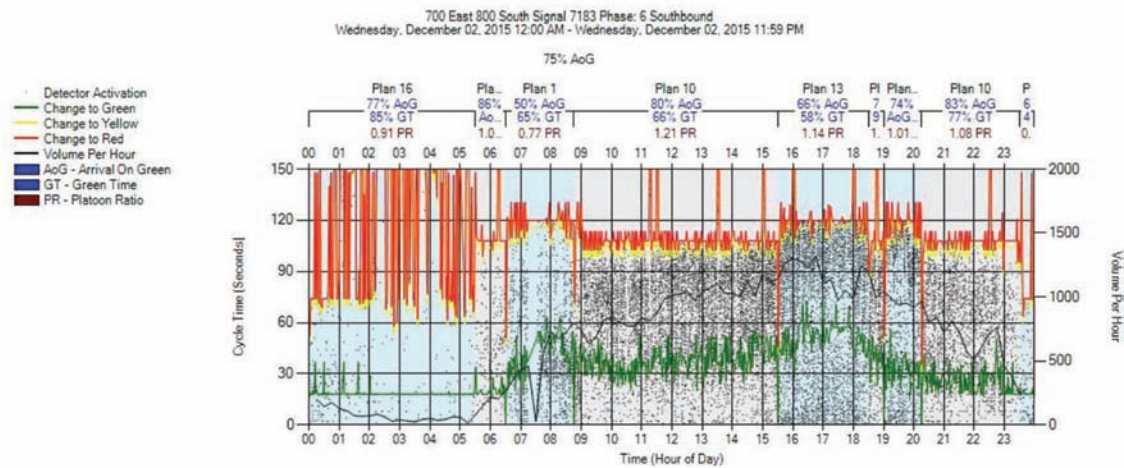


Figure 6.5 Example PCDs from INDOT signal performance measures website. Data are shown for three intersections along US 231 in West Lafayette, Indiana for 12/2/2015.



(a) Northbound.



(b) Southbound.

Figure 6.6 PCDs from Utah’s Signal Performance Metrics website [76] for the intersection of 700 East and 800 South, 12/2/2015.

some context to the numbers. This report shows the performance for a particular time period, the AM peak on 12/2/2015. This chapter will show some additional methods for tabulating similar types of information for longitudinal analysis.

6.4 Operational Diagnostics with PCDs

This section demonstrates the use of PCDs to examine operational characteristics of coordinated intersections, both for evaluating progression and also to see whether advanced control achieves the desired effect.

6.4.1 Field Data Examples

A good starting point is an example of exceptionally good progression. Figure 6.8 shows a PCD for a location where, for nearly all of the day, almost all of the arrivals are coincident with the green band. This type of performance is quite rare because, at most intersections,

there will be side-street entries and left-turning vehicles adding to the demand on the street, in addition to coordinated platoons coming from the other intersection. The PCD in this example is taken from one a diamond interchange, where the upstream intersection is the other end of the interchange. There are two platoons that can be seen in the point cloud, from the upstream left turn and the upstream through movement, and both of these are accommodated by the local green. In fact, the green seems to be somewhat longer than it needs to be, especially in the morning: the last 10–20 seconds of green have very few vehicles. The signal may be continuing to serve traffic in the other direction, but this suggests a potential opportunity to consider the split for the phase, especially if a conflicting movement has a heavy demand.

Figure 6.9 shows a PCD for a location that receives traffic from a freeway off-ramp. The arrivals are randomly distributed throughout the cycle, and there is no apparent platoon formation pattern, except for the afternoon hours. What happens here is that queues have extended to the advance detector. This situation has a particular signature

Corridor Summary

	Corridor Name	Arrival on Red		Delay		Volume Daily Average Per Approach	Intersection Number Of Approaches	
		Percent	Platoon Ratio	Daily Average Per Approach (hrs)	Average Per Veh (sec)			
Region 1	US-89 SB	10.06%	1.13	6	1.73	36,048	3	
	US-89 NB	15.41%	1.03	6	2.14	19,119	2	
	Riverdale NB/EB	24.58%	1.20	2	4.99	17,468	10	
	Riverdale (both)	26.66%	1.11	1	5.42	14,715	19	
	Riverdale SB/WB	30.13%	1.01	2	6.14	11,657	9	
	SR-126 SB	27.88%	1.08	2	5.20	11,079	10	
SR-126 NB	29.21%	1.10	2	5.78	10,521	10		
Region 2	5400 S EB between Bangerte & 2200 W - including Bangerte and 2200 W.	28.63%	1.10	16	9.63	30,349	5	
	Foothill Blvd NB/WB	19.24%	1.14	4	5.35	26,546	11	
	5400 S EB	31.25%	1.18	9	9.83	25,247	8	
	SR-36 NB	24.86%	1.00	6	3.51	22,800	4	
	Foothill (both)	20.76%	1.10	2	5.41	22,529	22	
	10400 S EB	38.05%	1.08	10	11.67	22,239	7	
	Foothill Blvd SB/EB	22.93%	1.05	3	5.49	18,511	11	
	Bangerte N/W	31.27%	1.91	4	11.62	17,397	15	
	BangerteHwy(both)	34.61%	1.78	2	11.91	15,548	30	
	SR-36 SB	18.13%	1.15	2	3.17	15,063	6	
	Wasatch Blvd SB	20.29%	1.30	2	2.81	14,682	5	
	Bangerte S/E	38.85%	1.63	3	12.29	13,698	15	
	5400 S WB	34.56%	1.05	5	8.68	13,082	7	
	10400 S WB	42.81%	0.98	7	14.10	12,632	7	
	WB on 5400 S from Bangerte to 2200 W (includes both Bangerte and 2200 W)	33.55%	1.06	7	10.95	11,346	5	
	Wasatch Blvd NB	26.26%	1.28	2	4.09	10,966	5	
	11400 S WB	20.22%	1.09	2	4.23	8,358	6	
	Region 3	Redwood Rd NB	13.58%	1.07	8	2.96	36,839	4
		University Ave. SB	15.31%	1.10	2	2.37	34,129	10
		University Ave (Both)	15.39%	1.09	1	2.55	26,201	20
Pioneer Crossing EB		14.1%	1.31	2	1.98	26,044	6	
SR-92 WB		30.87%	1.16	10	7.25	25,259	5	
Redwood Rd SB		13.72%	1.00	3	1.73	22,602	4	
SR-92 EB		22.84%	1.16	4	3.79	18,643	5	
University Ave. NB		15.55%	1.08	1	2.88	18,273	10	
SR-6 EB		32.79%	1.14	12	8.30	15,396	3	
SR-6 WB		24.19%	1.18	7	4.77	14,946	3	
Pioneer Crossing WB		20.04%	1.27	2	3.59	12,173	6	
Geneva Rd SB		29.85%	1.03	2	4.78	10,398	8	
Geneva Rd NB		31.89%	1.12	2	5.62	10,119	7	
State St NB		22.57%	1.12	1	3.67	9,755	7	
State St SB		27.12%	1.03	1	4.46	8,432	7	
Region 4		Main St (Richfield) NB	13.35%	1.01	1	1.27	5,766	3
	Main St (Both)	13.44%	0.99	0	1.23	5,448	5	
	Main St (Richfield) SB	13.58%	0.97	1	1.16	4,971	2	
	Bluff St (both directions)	20.51%	1.23	2	2.96	2,838	1	

Figure 6.7 Example executive report summary from the UDOT Signal Performance Measures website [49]. The report shows corridor performance for 12/2/2015 from 6:00 to 9:00.

to it: the only time when vehicle movement over the detector occurs is after the queue departure shockwave extends past the detectors. Thus, the first 20 seconds of

green have few detections. This is followed by saturation flow through the rest of green and well past the end of green as the queuing area fills up again.

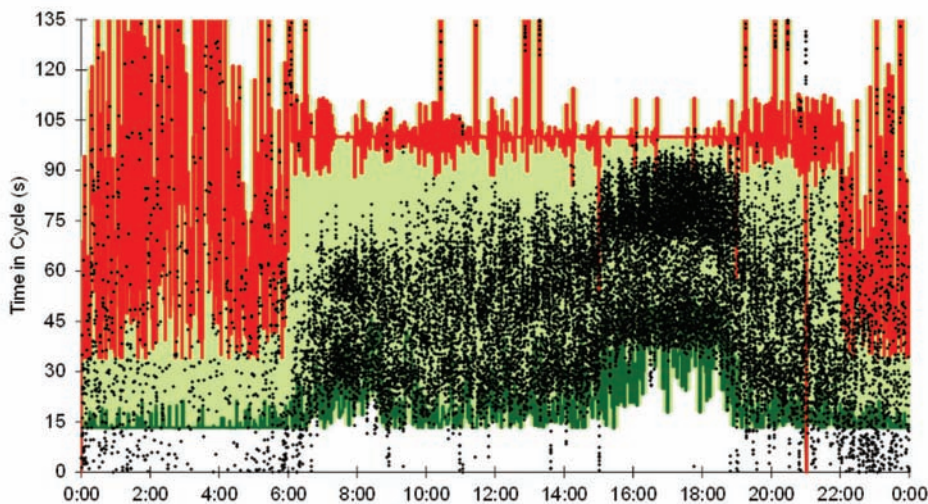


Figure 6.8 Example of exceptional progression: Eastbound, Pendleton Pike at I-465 Northbound Ramp, 7/1/2015.

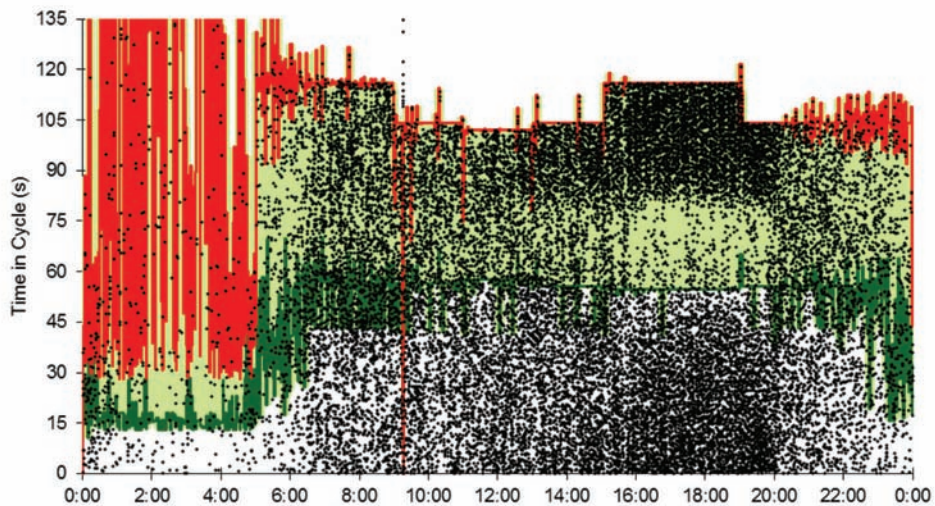


Figure 6.9 Random arrivals and queuing at Northbound, SR 37 and 126th Street, 7/1/2015.

The impacts of preemption can also be evaluated with PCDs. Figure 6.10 shows a location where railroad preemption is used. Preemption will sometimes truncate the coordinated phase to run a track clearance phase on a conflicting movement; this is followed by limited service in which the signal can cycle between the coordinated movements and left turns. Finally, the preemption exit phases are the two coordinated movements. The 24-hour view includes some cycles where the cycle length is much shorter than the usual 100-second cycle length, as well as other cycles that are much longer. There also appears to be an anomaly around 14:00 where, for several cycles in succession, the green times for the coordinated phase are much shorter than usual.

Figure 6.11 shows a PCD with an interesting repeating pattern during certain times of day. The

24-hour view (Figure 6.11a) exhibits what appears to be nearly vertical striping in the arrivals during the 9:00–11:00, 13:00–15:00, and 19:00–22:00 periods. Figure 6.11b shows a zoomed-in view of the 13:00–15:00 time period, which shows that platoons are arriving at the intersection; during each successive cycle, the platoon seems to arrive at an earlier and earlier time in the cycle. There are a few cycles where, by chance, the platoon arrives on green, and others where it arrives during red. This pattern is typical of a cycle length mismatch. In this particular case, the upstream intersection had a malfunction, and technicians replaced the controller with a spare that did not have the most recent timing plan programmed. The result was that during some times of the day, the cycle lengths had changed, leading to the mismatches.

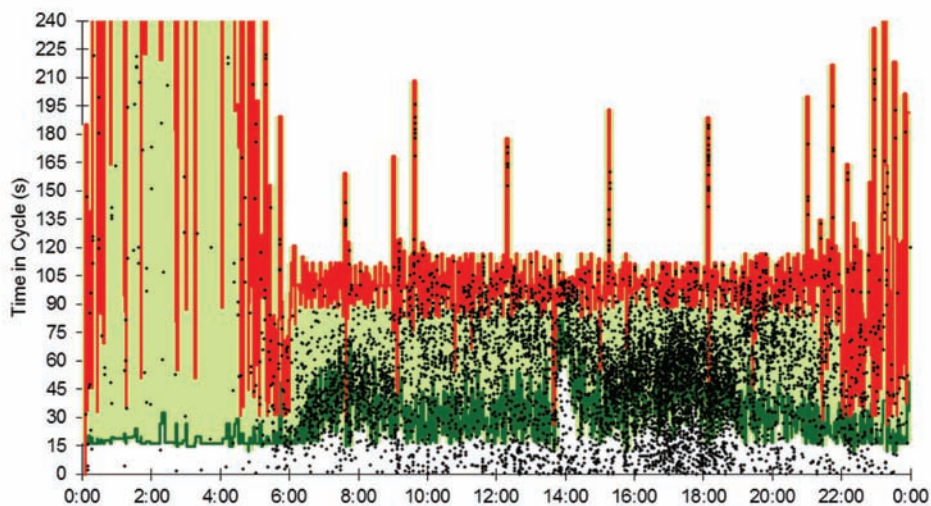
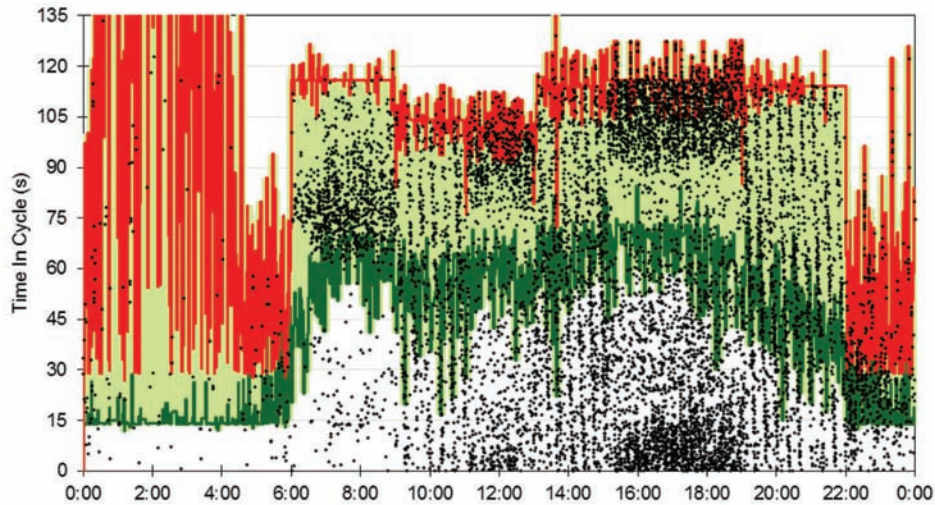
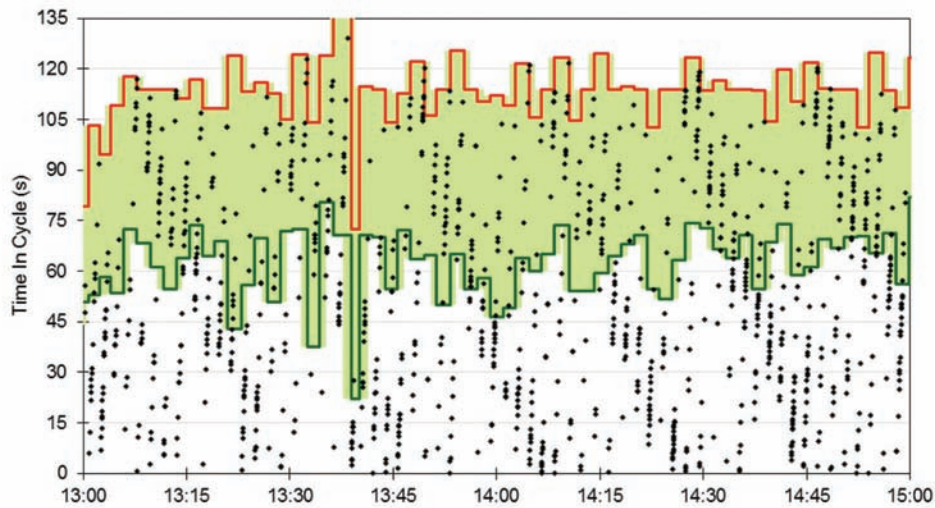


Figure 6.10 Shorter and longer cycles due to preemption, and transition back to normal operations: Eastbound, Pendleton Pike and Carroll Road, 7/1/2015.



(a) 24-hour view.



(b) Zoomed-in view of 13:00–15:00 time period.

Figure 6.11 Cycle length mismatch: Northbound, SR 37 and SR 32, 4/13/2009.

6.4.2 Diverging Diamond Interchange Examples

The PCD is also helpful for visualizing signal timing under unconventional intersection and interchange designs. For example, the diverging diamond interchange (DDI) is an interchange configuration that has been seeing increased use recently. In this design, the through movements cross over to the opposite sides of the street, which eliminates the conflict between the left turns at the two interchange intersections. This enables the two interchange intersections to be operated as “two-phase” signals, with some nuances regarding the clearance times.

Figure 6.12 shows a plan of the first DDI in Indiana, at S.R. 1 and I-69 in Fort Wayne [92]. Figure 6.13 shows PCDs for the two exiting movements of the interchange. Figure 6.13a shows the westbound

movement and Figure 6.13b the eastbound movement. These exhibit rather different behavior.

The westbound direction (Figure 6.13a) includes vehicle arrivals in nearly every portion of the cycle, yet two distinct platoons are visible. The current signal timing plan is able to progress the smaller platoon and the leading portion of the larger platoon. In this case, queuing does not pose a problem for this intersection, as many of the arriving vehicles are making the left turn onto the I-69 southbound ramp. Meanwhile, the eastbound direction (Figure 6.13b) appears to have a single platoon, which arrives on green during most of the cycle. There are relatively few vehicles arriving at other times in the cycle.

Figure 6.14 shows results from an evaluation of two alternative DDI signal phasing plans from an interchange in Utah [93]. Figure 6.14a shows the prior plan, which

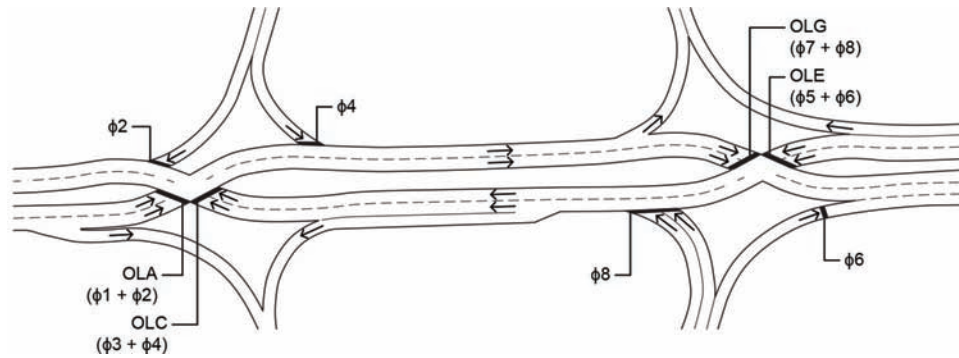


Figure 6.12 The interchange of S.R. 1 and I-69 in Fort Wayne [92].

used two-phase operation at each signal, yet led to a situation where there was more traffic entering the interchange than could exit at the other end. Figure 6.14b shows three-phase operation, which adds a third “holdback” phase into the sequence, which balances out these demands. Figure 6.14c and Figure 6.14d show the ring diagrams that represent two-phase and three-phase operation.

The results of using the two plans are shown by the PCDs. In these, the dots for vehicle arrivals have been colored to reflect their originating phase. The black dots represent the upstream through movement and the gray dots the upstream ramp movement. Figure 6.14e shows the PCD for two-phase operation; because vehicles arrive at all times in the cycle, it is only possible to progress one group of vehicles. Figure 6.14f shows how the signal operated with the three-phase scheme. In this case, it was possible to progress nearly all of the traffic. The quantitative performance measures show that POG was increased from 52.6% to 92.0%.

For this study, a video was made of the three-phase operation using video recorded from a drone (Figure 6.15). This both visually confirmed the operation and validated the identification of ramp and upstream vehicles in the PCDs [93,94].

6.4.3 Advanced and Experimental Control Examples

The same concepts can also be applied to evaluate advanced control algorithms. Figure 6.16 shows a view of a simulation network that was used to test several different control schemes for an arterial system in Morgantown, West Virginia [95]. There are two intersections in this network; we will show example PCDs for the labeled signalized approach. At this particular location, the upstream intersection was half-cycled, sending two platoons to the local intersection in each cycle. Figure 6.17 shows some PCDs from the evaluation.

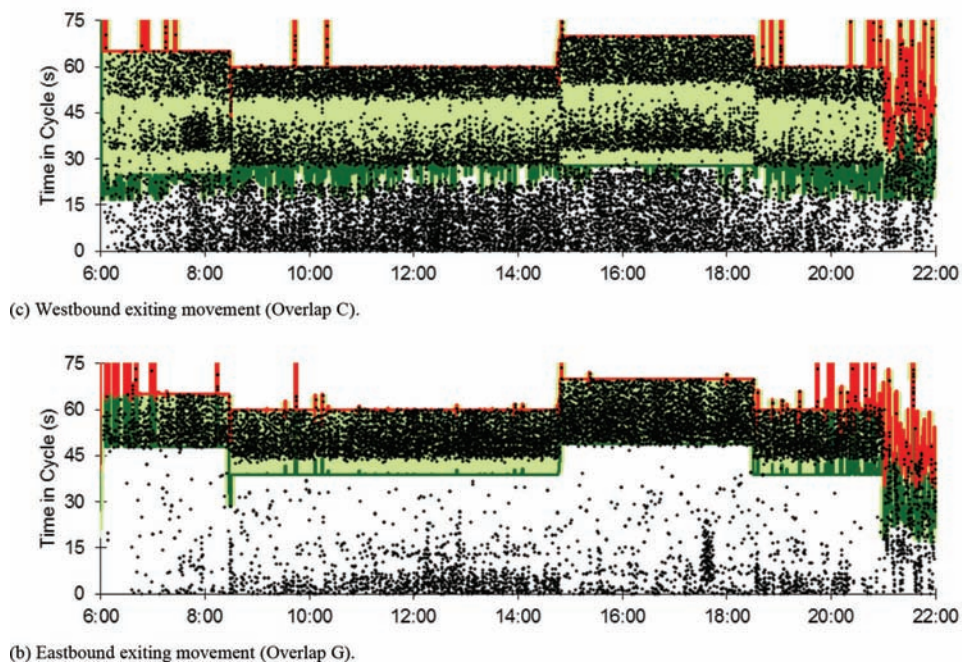


Figure 6.13 Evaluation of diverging diamond interchange signal timing: exiting movements at S.R. 1 and I-69 in Fort Wayne, from 7/1/2015.

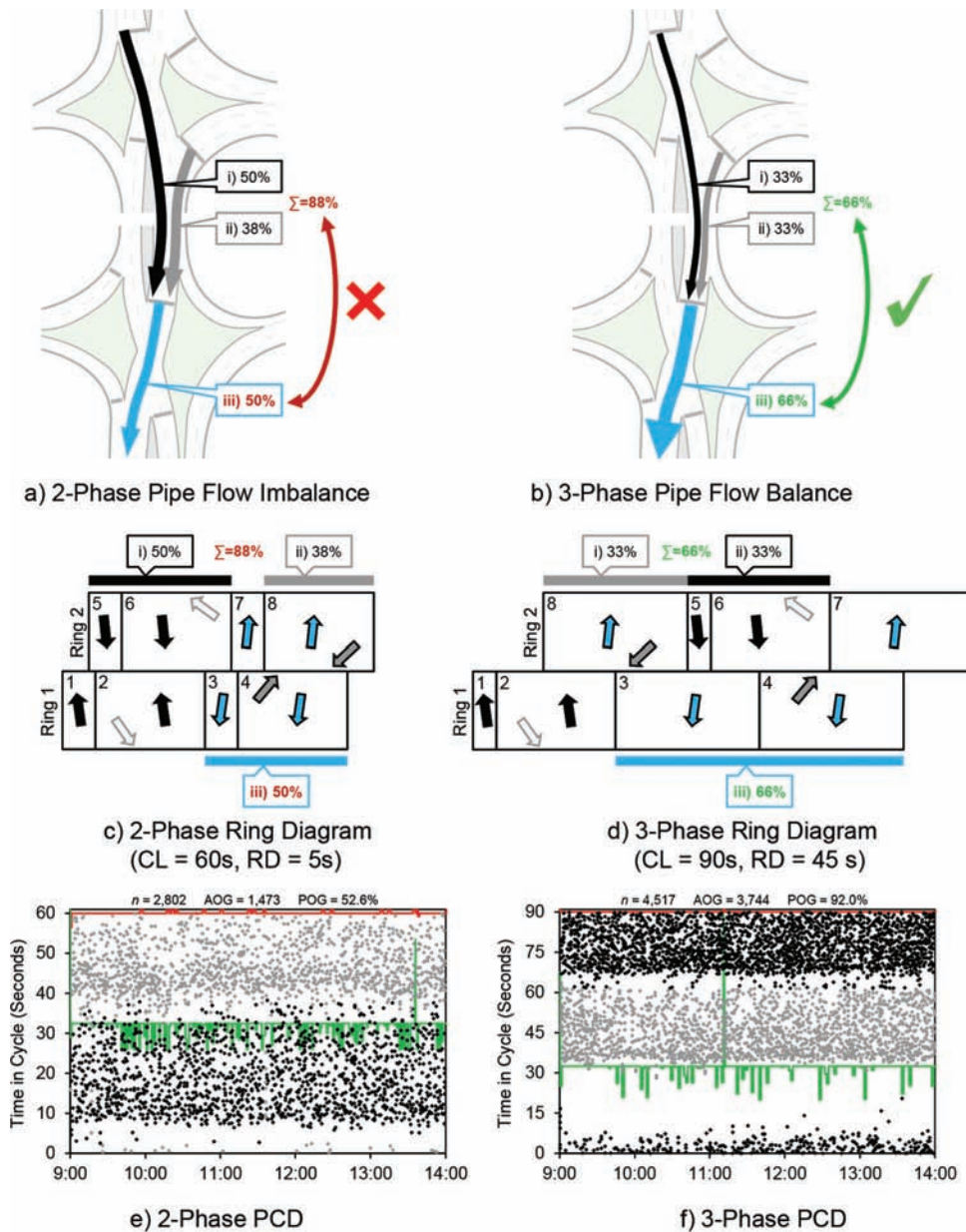


Figure 6.14 “Three-phase” operation of a DDI [93]. Data shown are from SR 201 and Bangerter Highway in Salt Lake

- Figure 6.17a shows the operation with the existing TOD plan. Two platoons arrive each cycle, but the local signal is only able to progress one of them. Every other platoon must stop. There are three separate TOD plans and this situation seems to be true of each of them.
- Figure 6.17b shows the operation after adaptive control has been enabled. The adaptive control system is capable of adjusting the offset and splits, and attempts to make as many vehicles arrive on green as possible. A visual comparison with Figure 6.17a shows that for most times of day, the progression has been improved for at least part of the TOD pattern. One can also see the adjustments as they occur, with the platoon arrivals shifting in response to adaptive adjustment of the offsets, and the duration of green expanding as the split is increased.

The PCD can also provide some insights on new, experimental forms of traffic control before they are used in the field, revealing operational details that totals and averages of delay do not convey, and for which vehicle trajectory diagrams can show in detail, but only for relatively small intervals of time. Figure 6.18 shows three PCDs from ad hoc tests of some experimental ideas, zoomed in on a 2-hour period.

- Figure 6.18a shows operation under a conventional TOD plan. These are the same data as shown earlier in Figure 6.17a, only here we have focused on a 2-hour period.
- Figure 6.18b shows an experimental scheme where the signal is preempted at regular intervals to account for

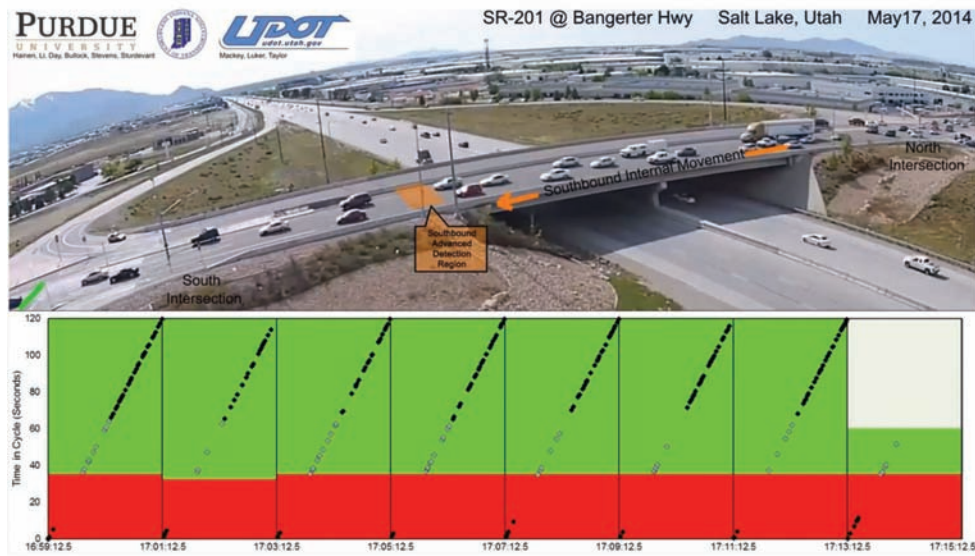


Figure 6.15 Visual assessment using video captured from a drone at SR 201 and Bangarter Highway in Salt Lake City, Utah [93,94]. The video is available at <http://dx.doi.org/10.4231/R7C24TC4>.

platoon arrivals from the upstream intersection. We can see that the effective cycle lengths are very short. Although most of the vehicles arrive in green, the amount of the cycle available for the side street is greatly truncated, which an analysis of the capacity metrics would further reveal. But regarding the progression quality alone, the scheme seems to have some promise, perhaps if the time between preempts were to be increased.

- Figure 6.18c shows another experimental scheme that uses transit priority to place requests for the mainline phase instead of preemption. In this case, the signal is allowed to use more time to serve the side street. The scheme appears to have some success. Most cycles have arrivals that are well aligned with green, but some others are slightly off the mark, and there is at least one cycle where no arrivals appear at all during green.

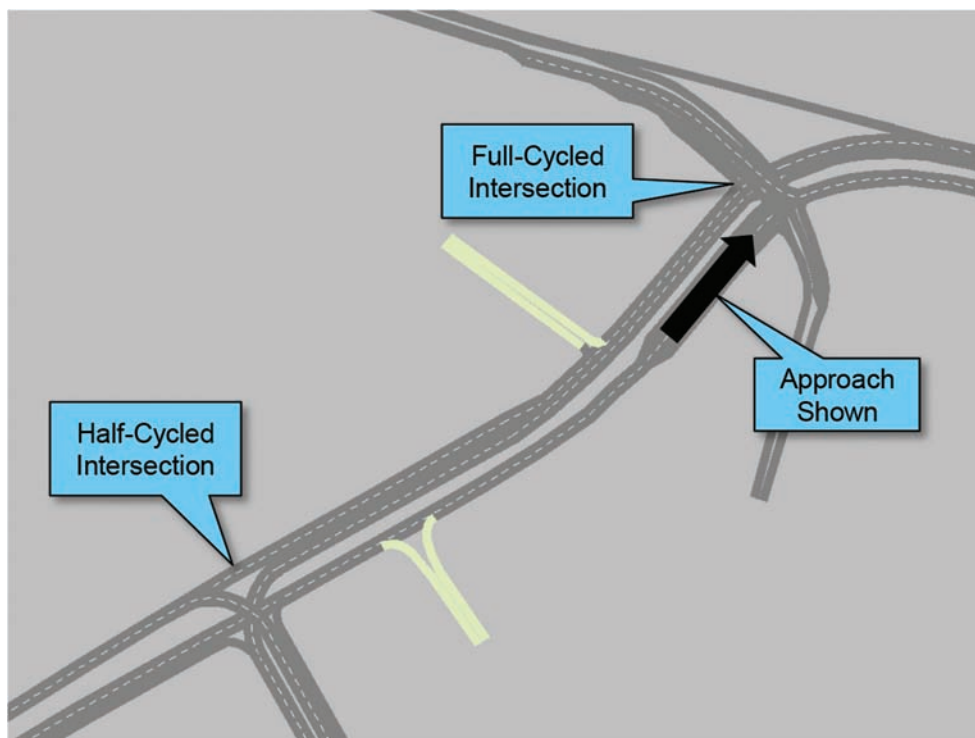


Figure 6.16 Portion of the simulation test network for testing advanced control [95].

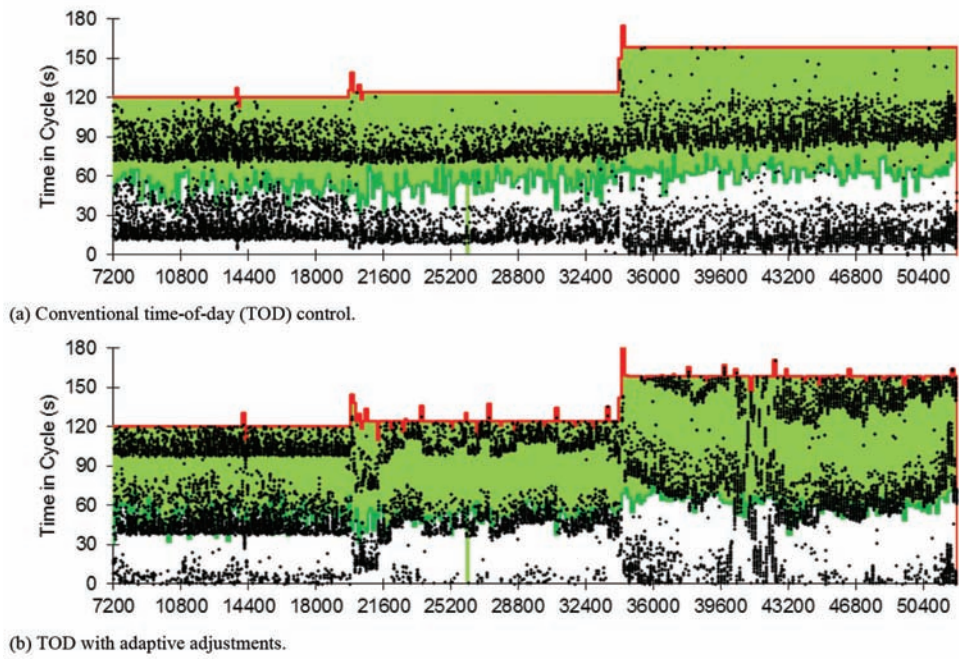


Figure 6.17 Advanced control diagnostics using PCDs from simulation data [95].

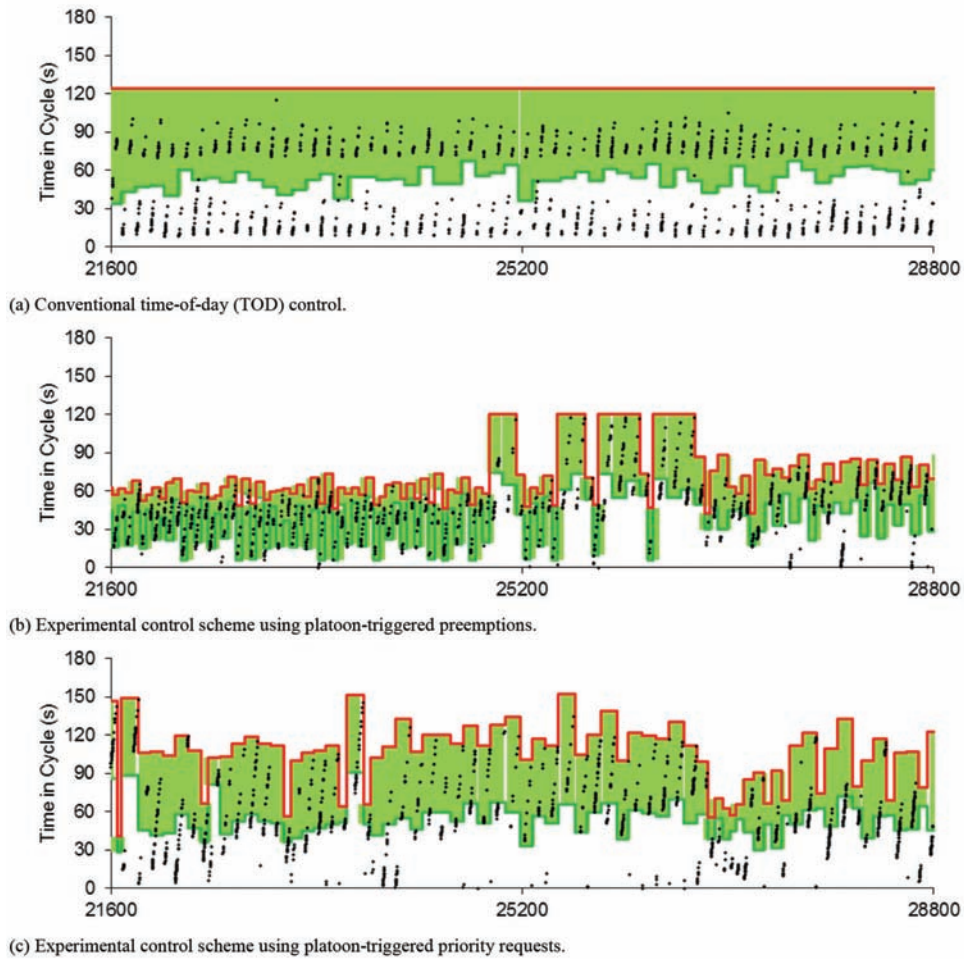
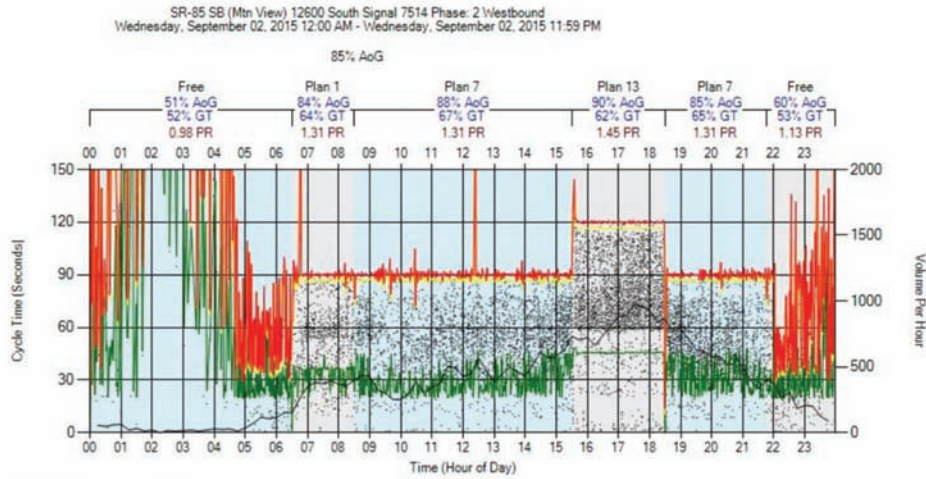


Figure 6.18 Experimental control diagnostics using PCDs from simulation data.

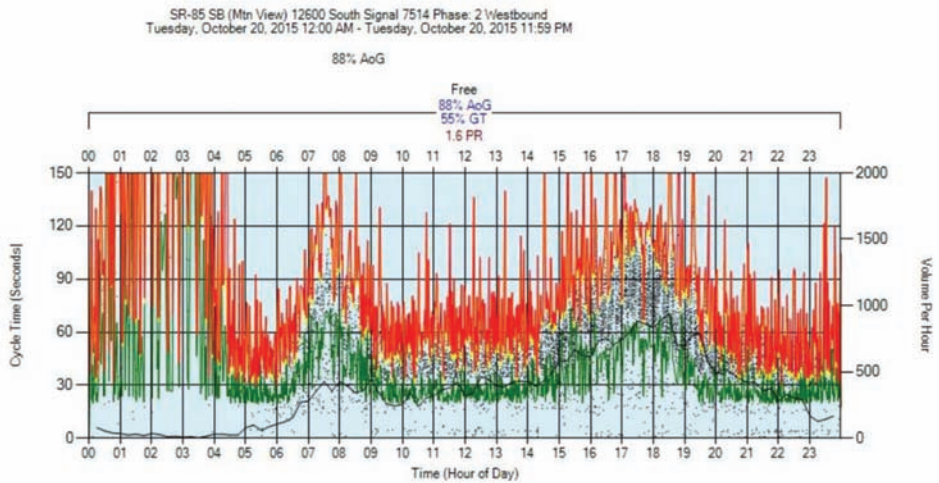
One can imagine that considerable value would be added to more comprehensively developed control schemes by the introduction of this sort of graphical view to explain how they work. Moreover, the cost of introducing the necessary data collection objects in a simulation environment is trivial.

Figure 6.19 shows a real-world example of a more recent control scheme developed by Matt Luker and colleagues at Utah DOT. This method uses the peer-to-peer feature that has become available in a few different signal controllers in recent years. This is a communication protocol that enables multiple controllers to share data with one another in a local network. The control scheme illustrated here is able to allow traffic patterns at a “master” controller to influence operations at neighboring “slave” controllers. The purpose is to enable cycle-free coordination within small clusters of intersections.

- Figure 6.19a shows an example of conventional TOD coordination for the westbound movement at SR 85 SB at 12600 South. The coordinated timing plan operates at 90 or 120 seconds throughout the day, and progression appears to be relatively good overall. The vehicle arrivals are coincident with green and the POG values are above 80%. As is typical, arrivals are random during the noncoordinated overnight hours.
- Figure 6.19b shows the same movement using the peer-to-peer control scheme. Now, rather than cycle length being fixed at 90 or 120 seconds, it naturally expands and contracts in response to traffic demands as they occur. Throughout the entire day, the vehicles appear to be well aligned with green, and the overall POG for the entire 24-hour period is 88%. Although not shown directly in this graph, the substantial reduction of cycle lengths during the midday period would very likely correspond to delay reductions for conflicting movements.



(a) Conventional TOD control.



(b) Peer-to-peer control.

Figure 6.19 Experimental control diagnostics using PCDs: real-world example. Data were obtained from the UDOT Signal Performance Measures website [76].

6.5 Optimizing Signal Timing with High-Resolution Data

After the quality of progression has been measured along a corridor, it is possible to use these data to improve the signal timing. With regard to coordination, the most relevant set of parameters that affect the quality of progression are the signal *offsets* that determine the relative timing of neighboring intersections. To optimize the offsets, it is first required to develop a working model that can predict the signal state given a set of adjustment to the offsets, and an optimization algorithm that can calculate optimal offsets from the full range of possible offsets.

6.5.1 Predicting Changes and Optimality of Solutions

Figure 6.20 shows a time-space diagram for three intersections, and illustrates a model of the likely conditions when the offset of Intersection 2 is adjusted by the value ΔO_2 .

- At the two downstream intersections (Intersection 1 and Intersection 3), the vehicles that originated from Intersection 2 will arrive later. Their arrival times are adjusted by the same value as the adjustment ΔO_2 .
- At the local intersection (Intersection 2), the red times (and therefore also the probability of green) will move later by the same adjustment, ΔO_2 . Equivalently, one may also adjust the vehicle arrivals so that they occur earlier, by the time $-\Delta O_2$.

Each approach (link) should be modeled independently, taking into account the local and upstream adjustments to the vehicle arrivals and/or the green distribution.

- Example 1: The local adjustment is +25 and the upstream adjustment is +50. The cycle length is 100 seconds.

- Move the local arrivals by +50 seconds.
- Move the local green by +25 seconds, or move the local arrivals by -25 seconds (in addition to the +50): $(-25 + 50 = 25)$.

- Example 2: The local adjustment is +75 and the upstream adjustment is +25. The cycle length is 100 seconds.

- Move the local arrivals by +25 seconds.
- Move the local green by +75 seconds, or move the local arrivals by -75 seconds (in addition to the +25): $(-75 + 25 = -50 \text{ mod } 100 = 50)$.

These transformations can be applied to both the PCDs and the flow profiles, and this model has been demonstrated to work well [19,96]. For example, Figure 6.21 shows a comparison of PCDs for an offset adjustment that was predicted by the above described model. Figure 6.21a shows the prior conditions, with most of the arrivals occurring in red. Figure 6.21b shows the predicted arrival patterns after an adjustment of +41 seconds to the arrivals, resulting from combined local and upstream adjustments. Finally, Figure 6.21c shows the actual field conditions after the new offsets for the corridor were deployed, which closely matches the initial prediction.

Figure 6.22 shows a more comprehensive comparison for an entire corridor, SR 37 North in Indianapolis [96]. The most dramatic changes were at SR 37 and 141st Street, especially the southbound approach. Here, the vehicle arrivals were changed from mostly arriving in red to mostly arriving in green, and the predicted profile strongly resembles the actual profile from after the new offsets were actually deployed. This is also true of the other approaches.

The prediction methodology also enables us to evaluate the quality of a set of offset adjustments prior

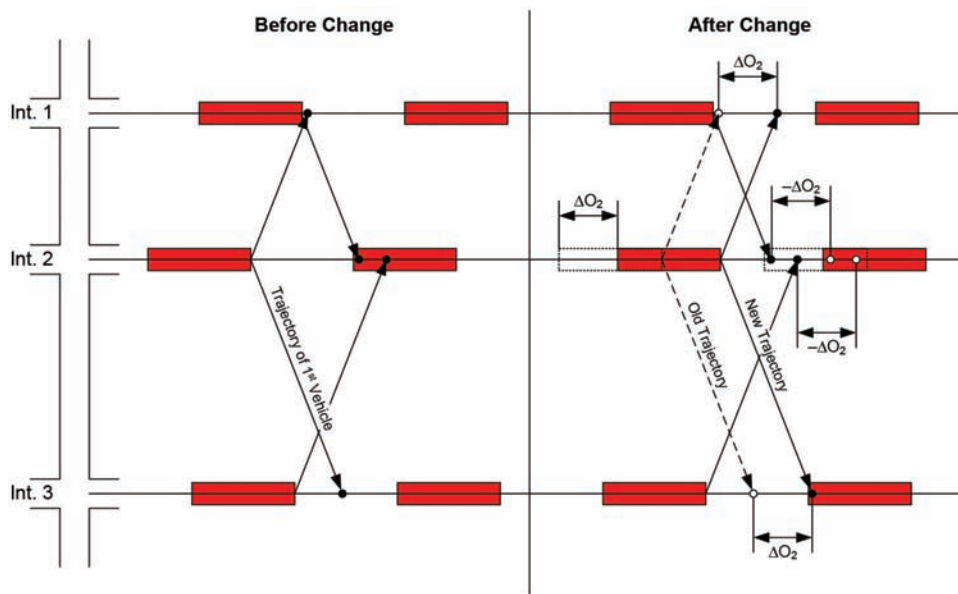


Figure 6.20 Predicting new arrival conditions after an offset adjustment [19].

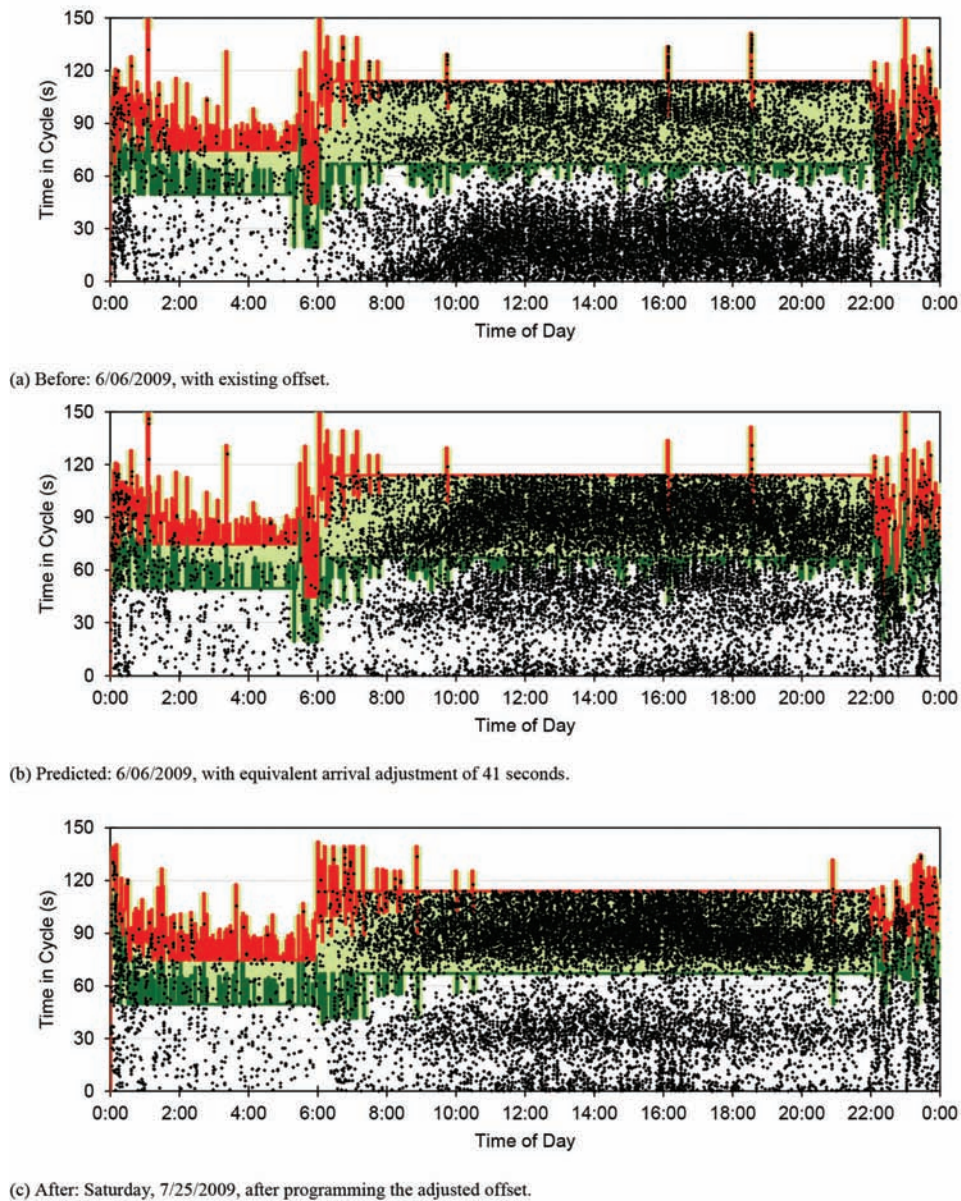


Figure 6.21 Comparison of predicted and actual vehicle arrivals from the pilot corridor study [19].

to their deployment in the field. Figure 6.23 explains the concept. The figure contains a plot of what appears to be a sinusoidal wave. This is an offset-performance chart; every point along this graph shows the predicted performance of a signalized approach under a trial adjustment. In this case, the performance measure is the percent on green (POG), which has an existing value of about 48% (where the adjustment = 0), a maximum value of about 82%, and a minimum value of about 15%. All of these are possibilities under different equivalent adjustments to the local offsets (combining the local and upstream adjustment).

To the left of the offset-performance chart is a diagram comprising a bar, a red dot and a green dot,

and a line connecting the two. The upper and lower bounds of the bar show the possible range of POG values that can possibly be achieved on this signalized approach. The red dot indicates the existing POG, and the green dot indicates the predicted POG with the new adjustments. The line indicating the two shows the predicted increase in POG. In this case, the new POG will be much higher than the existing one, although it does not quite reach the maximum possible value.

Figure 6.24 shows this graph for eight signalized approaches in the SR 37 North system, which is presented for the same dataset as the flow profiles in Figure 6.22. Increases in POG are indicated by green bars, while decreases in POG are indicated by red bars.

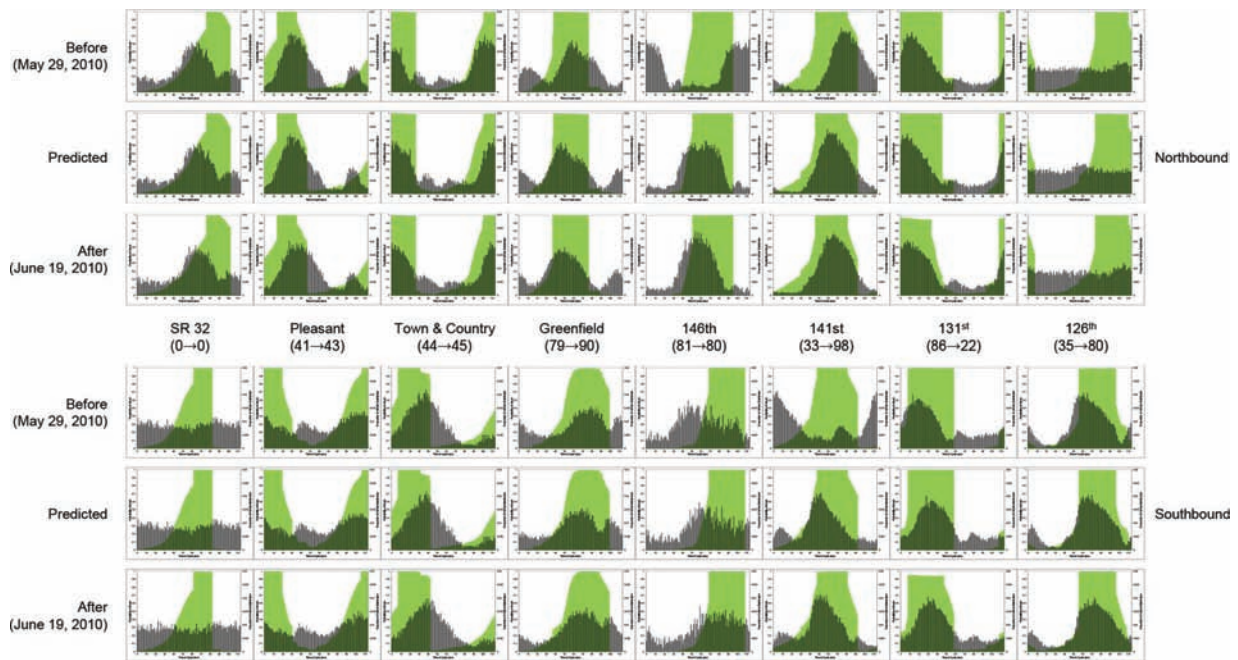


Figure 6.22 Comparison of flow profiles: before offset adjustment, predicted offset adjustment, and after offset adjustment [96].

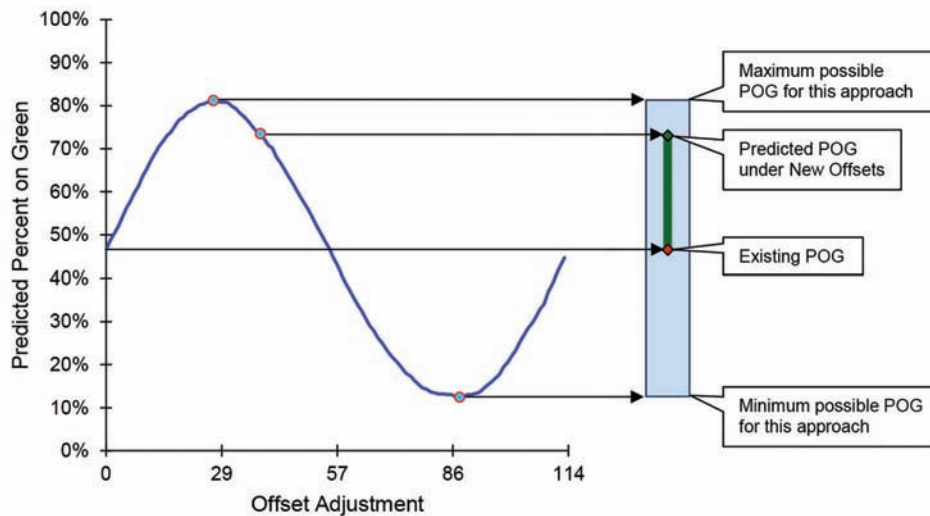


Figure 6.23 Explanation of the progression potential graph.

The chart shows that approaches 5sb, 6sb, and 7sb have very dramatic increases in POG. Approach 6sb appears to achieve the maximum possible value. There are more marginal improvements at 4sb, 4nb, and 5sb. Two approaches (3nb and 7sb) are predicted to have slight decreases in POG. The rest of the system is largely unchanged.

The two entry approaches (1sb and 8nb) that have random arrivals have a very narrow range of possible values (i.e., the bar is relatively flat), because the offset does not really have any impact on the arrival characteristics at those intersections, with no upstream signal. The approaches at 2nb and 3sb both have relatively low POG. These could likely be improved, but

probably only by decreasing the POG at either 2sb or 3nb. One can easily imagine alternative options being visualized by this method to make such tradeoff decisions.

6.5.2 Algorithms for Offset Optimization

Now that we have a modeling system that can reasonably predict the impacts of offset adjustments, we can now develop a method of systematically varying the offsets in the system to find the optimal condition. This is not as easy a problem as it may seem at first glance, since the number of offset combinations is equal to $C(n - 1)$, where C is the cycle length and n is the number of intersections.

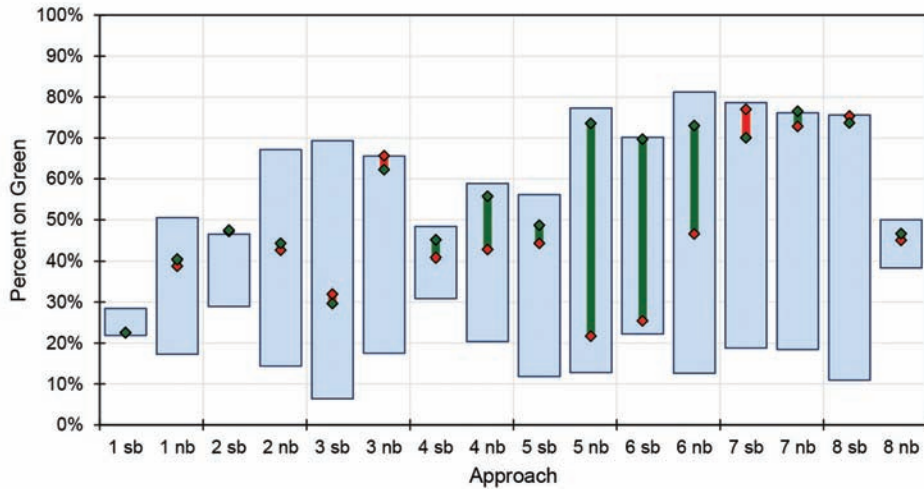


Figure 6.24 Progression potential graph for the 2010 offset optimization study.

In a previous study [97], five algorithms for optimizing the offsets were examined. These included:

- A quasi-exhaustive search that scanned the range of possible solutions by considering only five possible offsets at each intersection within the cycle length, reducing the number of combinations from $C(n-1)$ to $5(n-1)$.
- Monte Carlo selection, in which n random offsets were selected over hundreds of different iterations, thereby randomly searching through the $C(n-1)$ solutions.
- A genetic algorithm, in which the offsets are coded as a binary “chromosome” with an associated “fitness” that is proportional to the objective function. The initial population of solutions is randomly generated, and successive populations (iterations) are generated by roulette wheel selection where the odds of selection are proportional to the fitness, and the best solution is preserved from one generation to another. The iterative process continues until a stopping parameter (number of generations without improvement) is reached.
- Hill climbing, in which a set of possible offset adjustments are considered at each intersection, and the best one is kept. This procedure is repeated iteratively until a stable solution is reached [88].
- Link pivot, an arterial-focused implementation of the combination method developed in the 1960s in Great Britain [98]. This procedure begins at one end of the system and exhaustively searches for the best adjustment on the next link. When moving to the next link, the offset adjustment for that step are also made to all of the intersections previously optimized. This enables the global optimal solution to be identified.

While each of the five algorithms resulted in substantial improvements, slightly better solutions were found using link pivot, as illustrated by Figure 6.25. The amount of computational resources required by each algorithm is shown in Figure 6.26, which shows the number of computations needed to produce the optimal offsets. Link pivot and hill climbing required similar levels of computational resources and were considerably more efficient than the other algorithms considered.

When the underlying data type is flow profiles, and with some economization of the algorithm code (e.g., only refreshing states on a subset of links on each internal step during the program runtime), the computations involve relatively small quantities of numbers, and could rather easily be implemented in a variety of platforms. In particular, it is possible to set up the individual link calculations as parallel processes, which speeds the calculation further. The fast calculation of optimal offsets would make it a good candidate for embedding in another process to optimize other settings, such as cycle lengths and splits.

Figure 6.27 and Figure 6.28 show flow charts that describe the function of hill climbing and link pivot, respectively, in more detail. The two algorithms can be used for most signal systems, depending on the network geometry:

- A geometry that contains closed loops (e.g., one could drive around in circles in the network without making

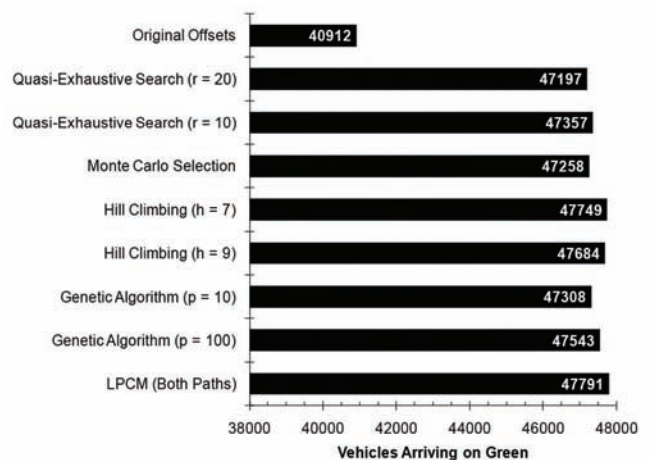


Figure 6.25 Performance comparison of several alternative offset optimization algorithms [97].

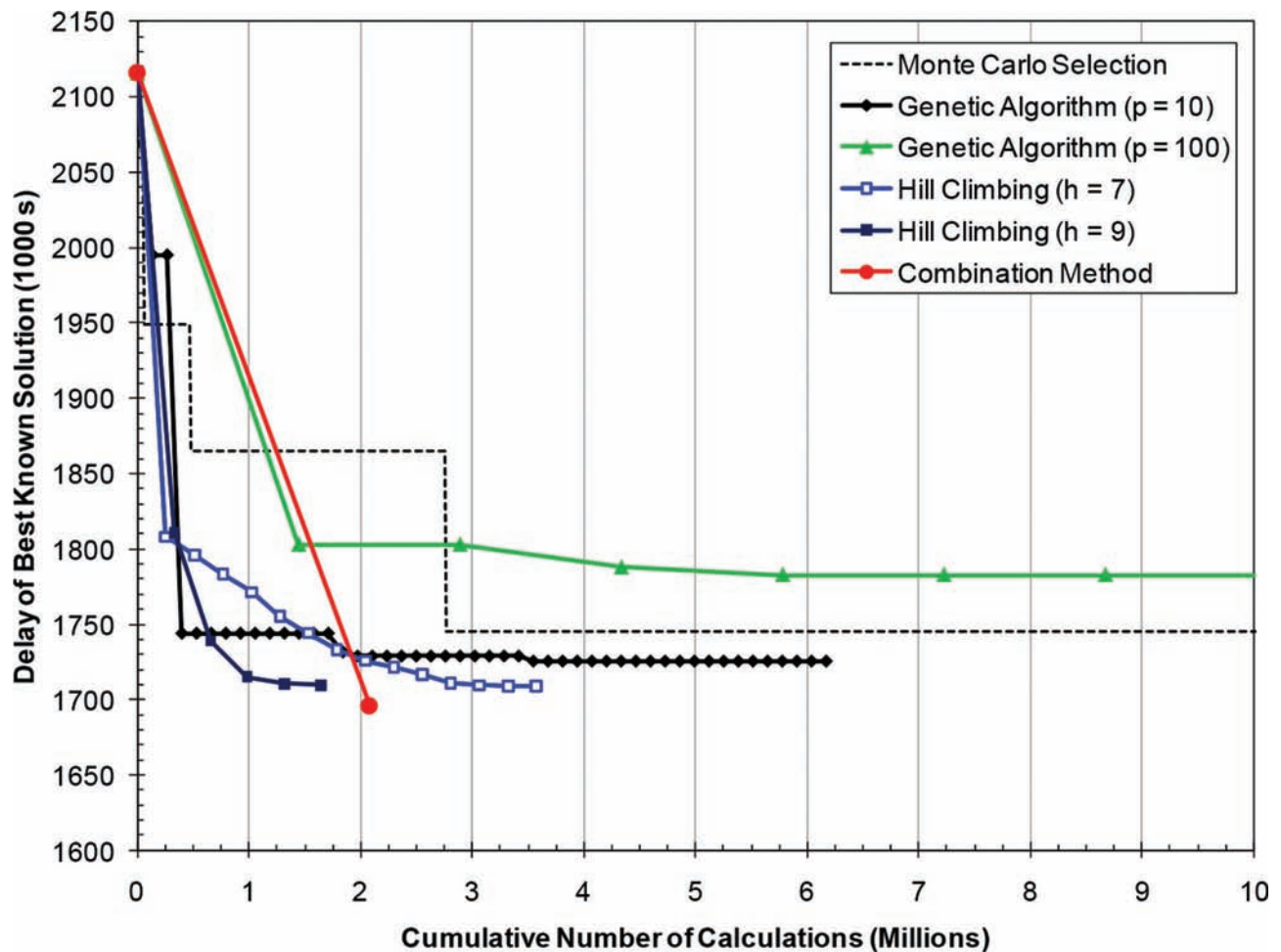


Figure 6.26 Computational performance of several alternative offset optimization algorithms [97].

U-turns) can be optimized using hill climbing. An example of such a network is a grid system.

- A geometry that does not contain closed loops can be optimized using link pivot. Examples include an arterial or a crossing arterial. For a crossing arterial, the offsets on the secondary roadway can be made to match up with the primary roadway using a common adjustment on either group of offsets, to match up offsets at the shared intersection where the crossing occurs.

Hill climbing (Figure 6.27) is an iterative algorithm. On each iteration, it walks through all of the intersections in the system. At each intersection, it considers an array of “hill climb increments,” or potential adjustments to the local offset. Each adjustment is modeled. The hill climb increment that provides the best performance (lowest delay, highest AOG) is saved, and the algorithm moves to the next intersection. This process repeats until no further improvements are obtained after scanning over all the intersections.

Link pivot (Figure 6.28) is a constrained exhaustive search that reduces the search space by considering the network relationship of the system. First, the intersections are listed in the order to be considered. For an

arterial, this begins at one end and ends at the other. At each intersection in the list, all of the possible adjustments to the offset are modeled (“ δ ” in Figure 6.28). The best performing adjustment is then saved. However, starting with the second intersection, not only are the trial adjustments applied locally, they are also applied to all of the previously optimized intersections—which preserves the link flow relationships that we just finished optimizing. In this manner, the algorithm moves from one end of the system to the end, “pivoting” on each link during the process until the entire system is optimized. The last offset can be given any arbitrary adjustment, which can reconcile coordination with other systems (e.g., a crossing arterial).

The process of finding the individual δ values can also be parallelized. That is, because each individual scan of δ values affects two links (that are not involved in any other search), each of the iterations through the j loop can be done in a parallel process, then the final offsets computed from properly summing up the δ values at the end.

Link pivot has been implemented as a web-based dashboard in both the INDOT and UDOT systems, as

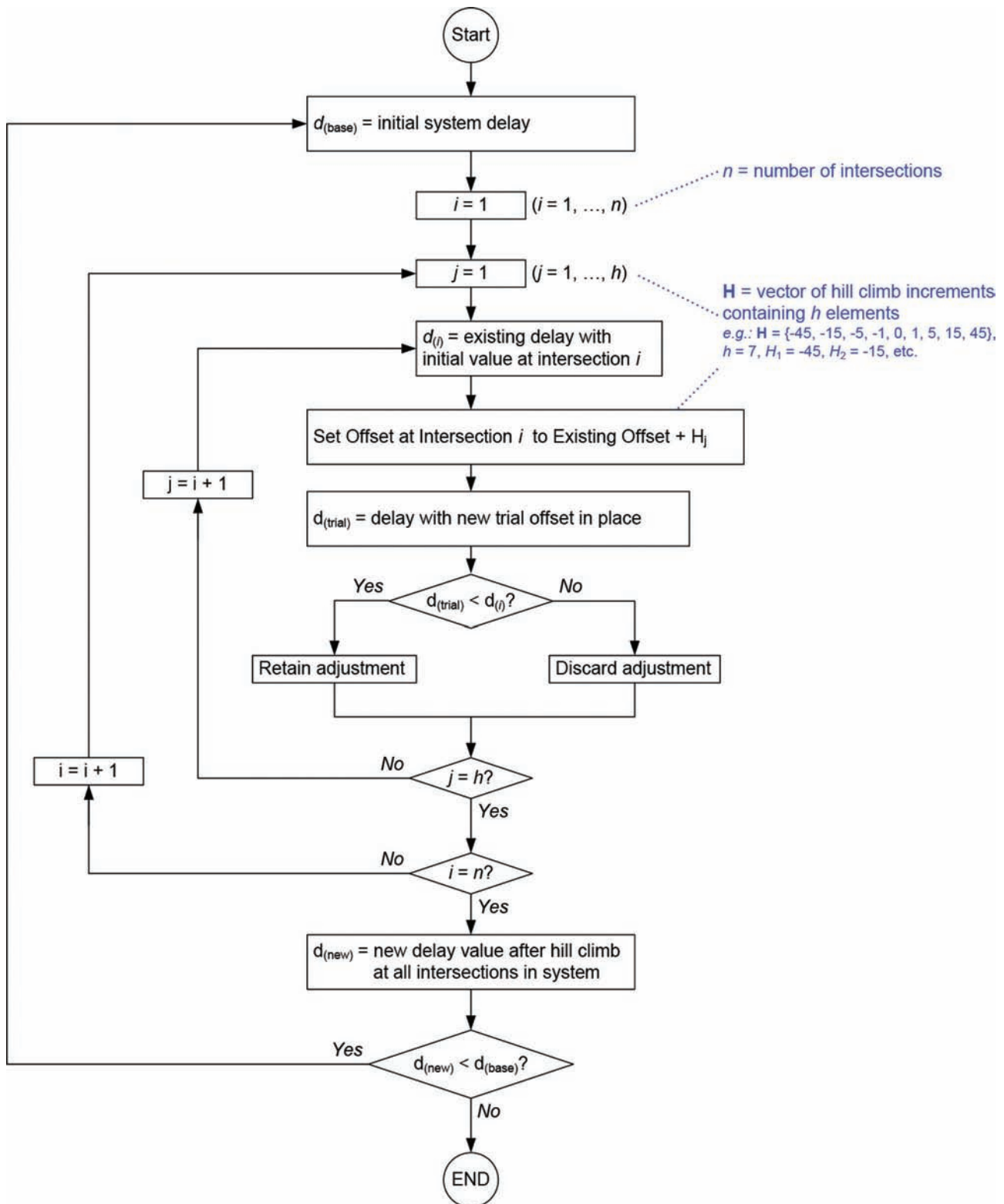


Figure 6.27 Hill-climbing algorithm.

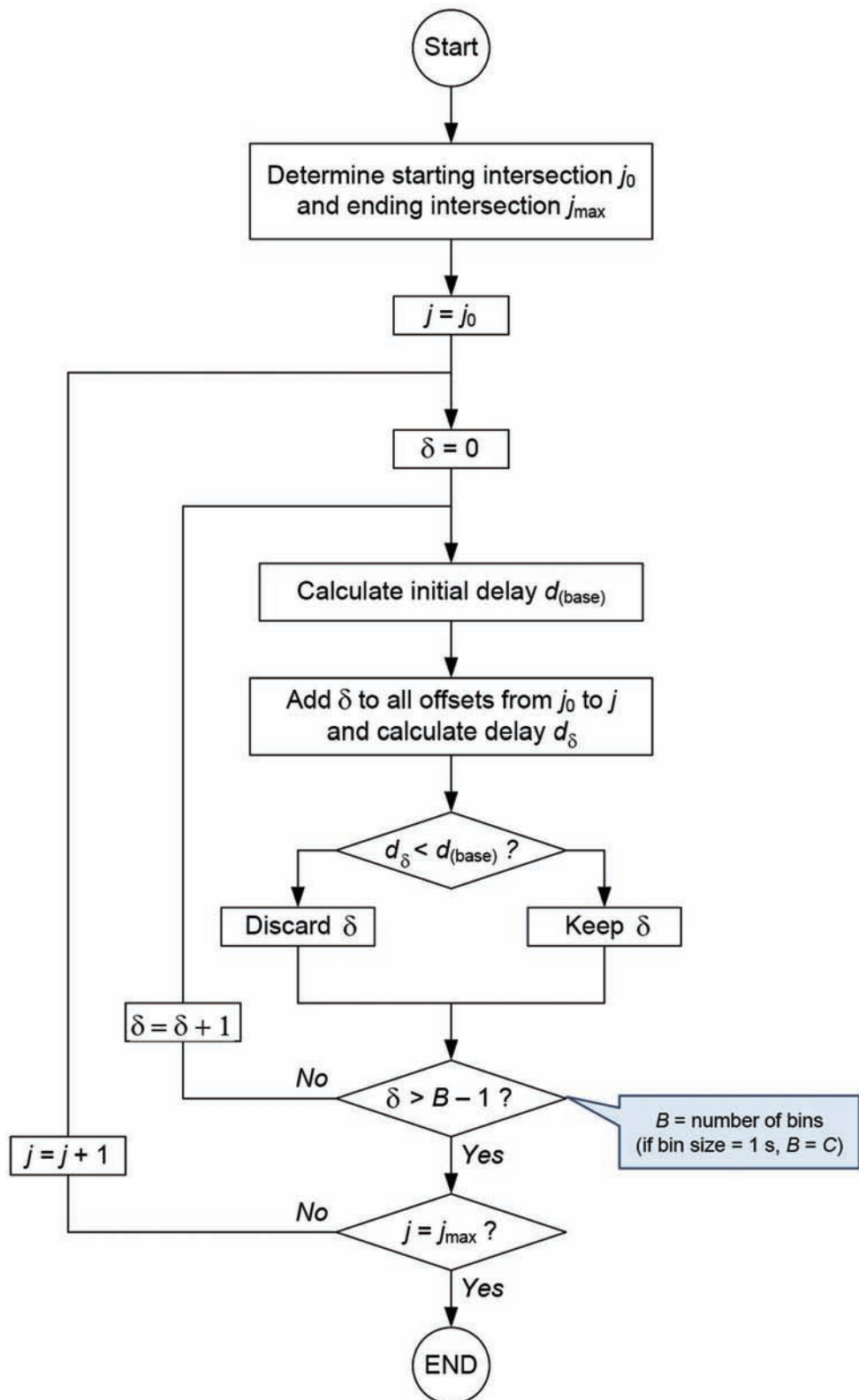


Figure 6.28 Link pivot algorithm.

illustrated by the screen shots in Figure 6.29 and Figure 6.30. At present, these dashboards display the suggested offset adjustments, which the user must reconcile with the existing offsets and then program in the new offset. However, with the addition of a utility to perform SNMP “gets” and “sets” of the relevant offsets, it would be possible to automate the process of reprogramming the field controllers.

6.5.3 Example Field Implementation

Link pivot has been used extensively in the field in the past several years. This subsection shows some results of its use in Indiana. The first corridor where it was tested was SR 37 North. A map of this corridor is provided in Figure 6.31. The corridor consists of nine intersections in total. The map shows the locations of Bluetooth devices that were used to measure travel times during the 2010, 2013, and 2015 offset-optimization studies.

- An initial study of offset improvements was conducted on the north end of the corridor (Intersections 1–4) in 2009 [19].
- Offsets were optimized for the entire corridor, then eight intersections, one year later in 2010 [96].
- A new intersection with 135th Place was opened in 2012.
- Offsets were optimized once more in 2013 and again in 2015 [52].

Figure 6.22, presented earlier, shows an example of the Saturday plan flow profiles before and after the 2010 offset optimization study. A more detailed example from that year is provided in Figure 6.32 and Figure 6.33. These respectively show the coordination on the south end of the system before and after optimization of offsets. At the time of the study, Intersection 7 had not yet been signalized. The “before” PCDs show that progression in both directions at 146th Street is rather poor, with the arrivals completely misaligned with green (Figure 6.32a, Figure 6.32b). Additionally, both approaches at 141st Street appear to have platoons cut off by the end of green (Figure 6.32c, Figure 6.32d). Finally, the southbound approach at 131st Street has platoons arriving a bit early (Figure 6.32e).

After optimizing the offsets, the arrival patterns were improved. 146th Street has much better progression in the southbound direction (Figure 6.33a), while the northbound is improved, although not perfect (Figure 6.33b). Both directions at 141st Street, however, appear to have platoons arriving nearly perfectly within the green band (Figure 6.33c, Figure 6.33d). Southbound at 131st Street is now better aligned as well (Figure 6.33e). Meanwhile, northbound at 131st Street and southbound at 126th Street did not see any worsening (Figure 6.33f, Figure 6.33g). Finally, northbound at 126th Street had random arrivals in both cases and consequently was not affected (Figure 6.32h, Figure 6.33h).

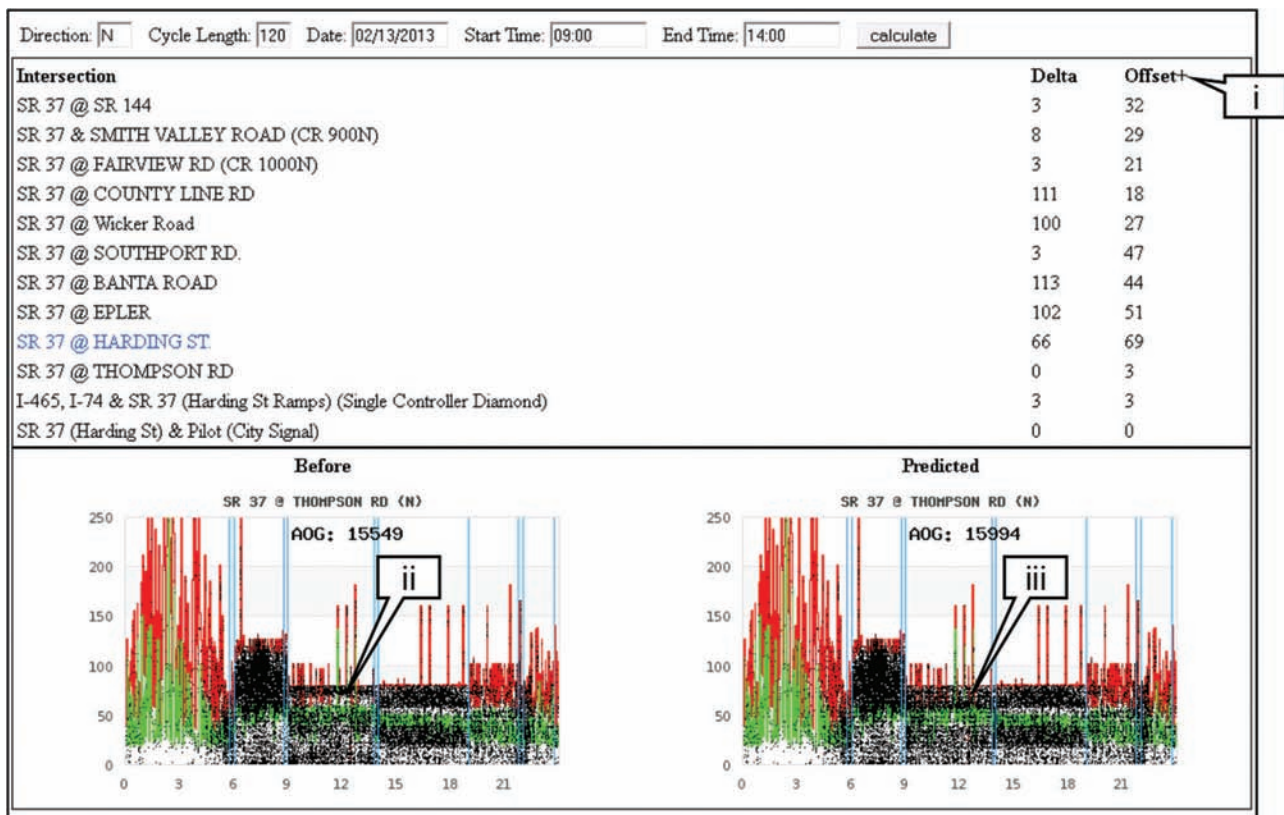


Figure 6.29 Implementation of link pivot on INDOT Signal Performance Measures website.

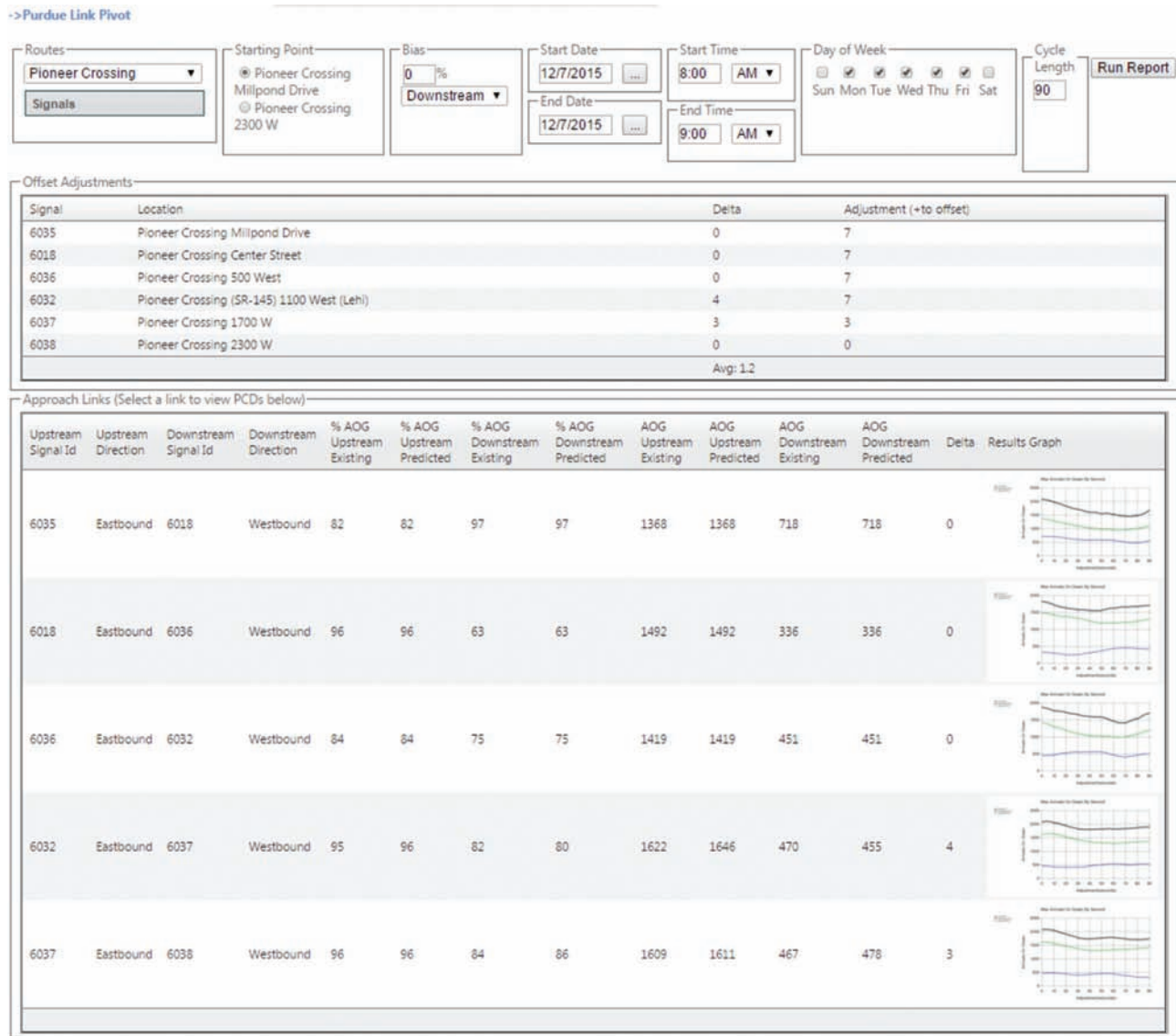


Figure 6.30 Implementation of link pivot on UDOT Signal Performance Measures website [76].

One additional observation from these PCDs that can be made is that southbound at 146th Street appears to have very light traffic (Figure 6.32a, Figure 6.33a). This was due to the fact that one of the two advance detectors was not functioning at the time. This was corrected during the optimization process by weighting the arrivals on green on that approach by a factor of 2, which assumed that both lanes were utilized equally. This ensured that the value of AOG on the approach would not be out of balance with the other direction.

As an aside, a question that is sometimes asked regarding the optimization process is whether it is possible to add some directional weighting to the AOG numbers. Typically, if there is one direction or another that is dominant, the volumes themselves should reflect this, so no weighting would necessarily be needed. However, it is certainly possible to add a multiplier for any individual approach to either correct for known detector inaccuracies

or to cause the algorithm to favor a particular approach or direction along the arterial.

The outcome of the offset retiming was assessed using travel time measurements with the Bluetooth sensors. Three sensors divided the corridor into two sections, as shown in Figure 6.31. Travel times on the north section and south section can be considered separately, as well as travel times through the entire corridor.

Six different coordination patterns were optimized in total: the Saturday pattern, which runs all day from 0600–2200; and five weekday patterns: 0600–0900; 0900–1100; 1100–1300; 1300–1500 and 1900–2200 (i.e., same pattern for two periods and 1500–1900). Figure 6.34 shows the results for the 2010 study for the Saturday timing plan. The six plots show cumulative frequency diagrams for northbound and southbound travel times along the entire corridor (Figure 6.34a, Figure 6.34b), the north end of the corridor (Intersections 1–4, Figure 6.34c,

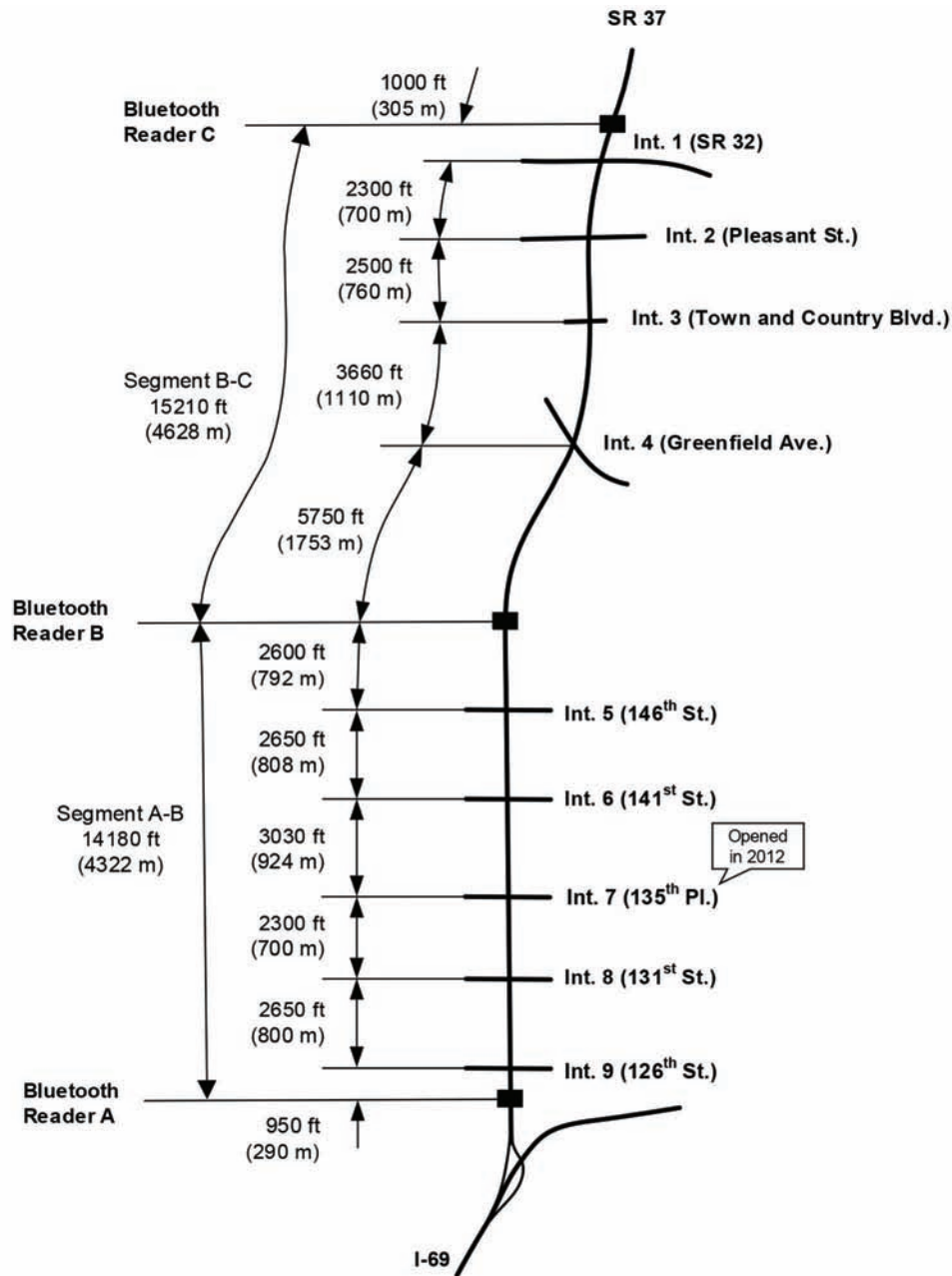


Figure 6.31 Map of SR 37 North [52].

Figure 6.34d), and the south end of the corridor (Intersections 5–9, Figure 6.34e, Figure 6.34f). The five lines show one “before” date (labeled “Base”), and four “after” lines in which four different optimization objective functions were tested:

- I. Maximize arrivals on green
- II. Minimize delay
- III. Minimize delay and stops
- IV. Maximize arrivals on green, with queue clearance time (i.e., the first 10 seconds of green were considered to be red for optimization purposes).

The travel time results show very little difference between the four different objectives; ultimately the fourth resulted in the lowest median travel time for the entire corridor in both directions, but it was only marginally less than the others. The overall corridor saw a travel time reduction of about a minute in both directions (Figure 6.34a and Figure 6.34b). Most of this resulted from reductions on the south end. Figure 6.34e and Figure 6.34f both show a separation of about one minute between the baseline conditions and the four “after” datasets. The north end saw only a

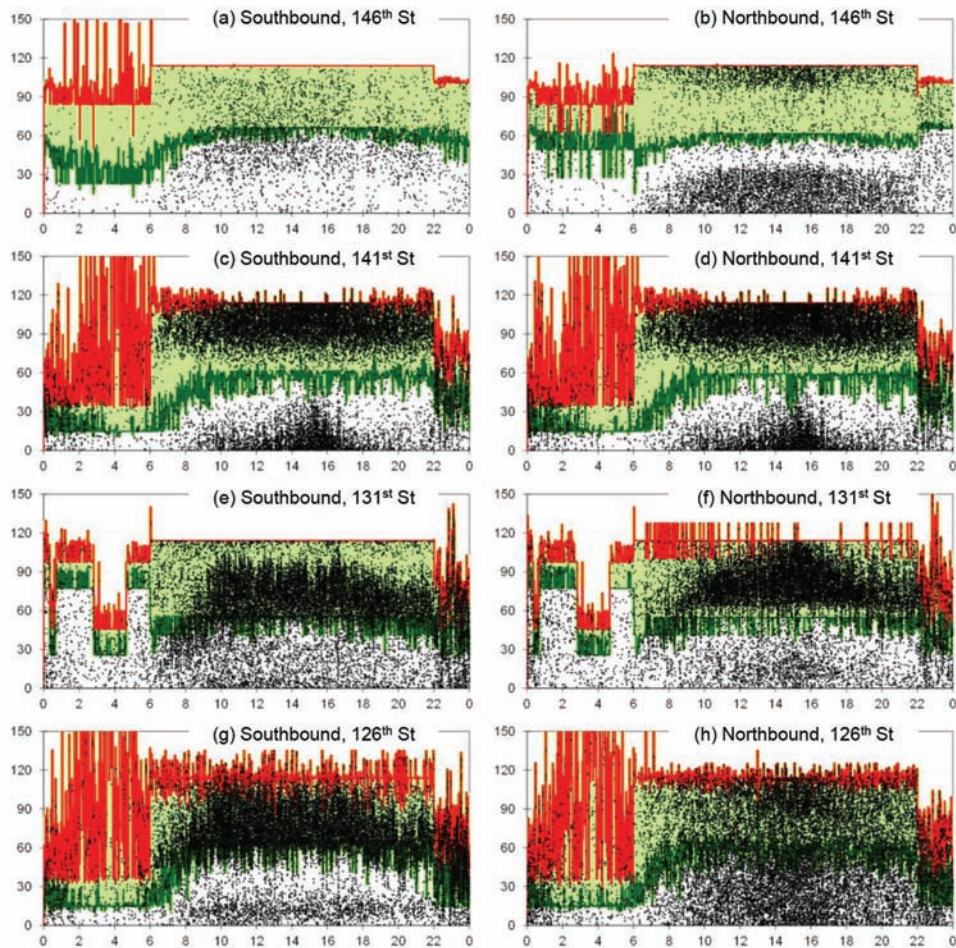


Figure 6.32 Selected PCDs from SR 37 before optimization, Saturday, 5/29/2010.

little improvement because those offsets had been optimized about a year before. The northbound direction saw slightly more reliable travel times (Figure 6.34d), shown by the reshaping of the curve.

Figure 6.35 shows some example results from 2013 and 2015, showing two different outcome types that were typical of the weekday plans:

- *2013, weekday PM peak (15:00–19:00)*. Figure 6.35a and Figure 6.35c show CFDs where travel time was reduced for both the northbound and southbound directions. The northbound direction saw a substantial reduction of the variability of travel times (Figure 6.35a), whereas the southbound direction had mostly an improvement in the median travel time (Figure 6.35b). These are shown as box plots in Figure 6.35e.
- *2015, Saturday plan (sampled from 12:00–15:00)*. Figure 6.35b and Figure 6.35d show CFDs where travel time decreased for the northbound direction (Figure 6.35a), but increased and became less reliable for the southbound direction (Figure 6.35b). Box plots are shown in Figure 6.35f.

There is often a tradeoff between two directions in the quality of progression, with the direction with the more

dominant volume winning over the other. However, the net outcome is typically positive for the corridor as a whole.

Figure 6.36 shows a comparison of the arrival on green (AOG) and arrival on red (AOR) performance of the system before and after each of the retimings in 2010, 2013, and 2015. These are shown separately for Saturdays and weekdays. Figure 6.36a and Figure 6.36b present these in terms of the total arrivals, and Figure 6.36c and Figure 6.36d show these in terms of percentage.

The overall height of the bars in Figure 6.36a and Figure 6.36b represent the total volume along the corridor, summed over all the approaches (Intersection 7 was excluded from all the sums to compare 2010 directly with 2013 and 2015). In the 5-year period, traffic volumes increased by 25% for Saturdays and 16% for weekdays. In spite of the increased volume, however, efforts to optimize offsets along the corridor were able to maintain an increase in AOG beyond the growth of the total volumes. The total of all approach AOGs increased by 46% between 2010 and 2015 for Saturdays, and by 28% for weekdays.

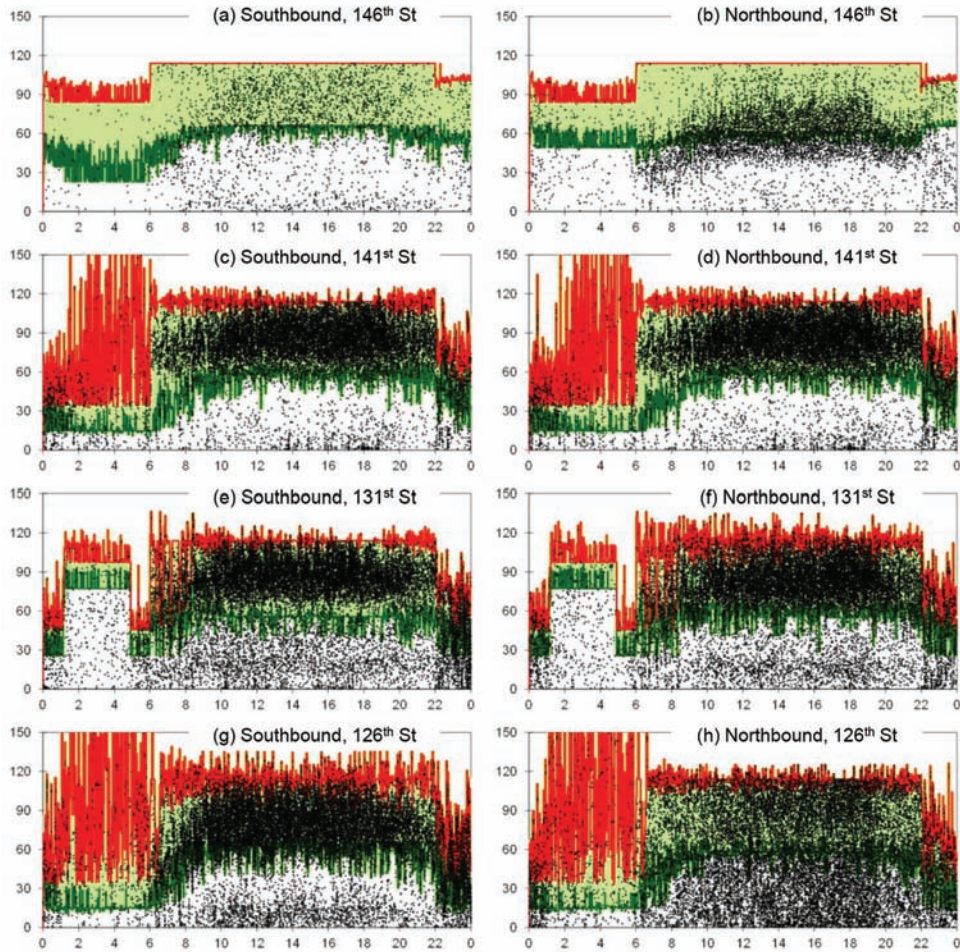


Figure 6.33 Selected PCDs from SR 37 after optimization, Saturday, 7/10/2010.

The percent arrivals in Figure 6.36c and Figure 6.36d show the proportion of the total arrivals belonging to northbound AOG, southbound AOG, and all AOR. For example, on 2015 Saturdays (Figure 6.36c), after optimization, the northbound AOG accounted for 31% of total arrivals, whereas southbound AOG was 39% of the arrivals, a total of 70% (callout *i*), leaving 30% AOR. Each year saw a growth in the total share of AOG when comparing the “before” and “after” scenarios. The charts also show some degradation between years, especially between 2013 and 2015 for weekdays.

Figure 6.37 shows the total duration of green for each of the before/after datasets during each study year, for Saturdays (Figure 6.37a) and weekdays (Figure 6.37b). The height of the bar represents the average of the total arterial green time among the intersections. For example, the first bar in Figure 6.37a shows that there were approximately 13 hours of green provided to the arterial, averaged across all the intersections. What this chart shows is that the increases in AOG shown previously were not obtained by substantially increasing the coordinated phase splits. In other words, despite the growth in volumes, it was not necessary to change the basic

capacity parameters: neither the splits nor cycle lengths were changed to favor larger arterial bands. Although such actions are not likely avoidable if the volumes continue to grow, re-optimizing the offsets on a regular basis was perhaps able to forestall the increasing of cycle length or coordinated phase splits.

To quantify the overall impact and provide an external outcome assessment, the difference in travel times for each before/after pair were converted into user costs. The average and standard deviations of the travel times were used to evaluate both the shift in travel time and the changes in travel time reliability [52]. The following formula was used:

$$\text{COST} = \frac{364}{60} \left(T_{\text{avg}} v_{\text{pc}} o_{\text{pc}} u_{\text{pc}} + k_{\text{pc}} T_{\text{std}} v_{\text{pc}} o_{\text{pc}} u_{\text{pc}} \right. \\ \left. + T_{\text{avg}} v_{\text{hv}} u_{\text{hv}} + k_{\text{hv}} T_{\text{std}} v_{\text{hv}} u_{\text{hv}} \right) \quad (1)$$

where

T_{avg} is average travel time (min), T_{std} is the standard deviation of travel time (min);

v_{pc} and v_{hv} are the total volumes of passenger cars and heavy vehicles;

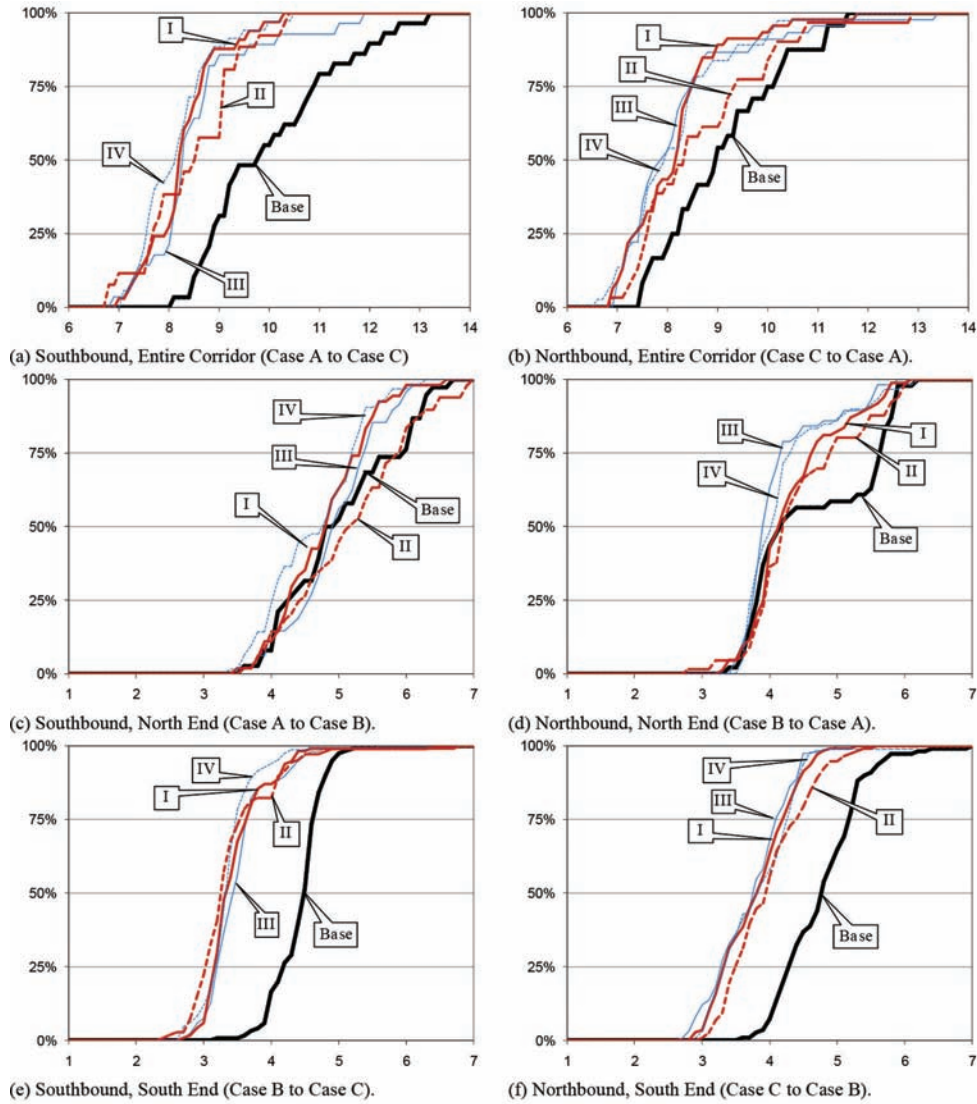


Figure 6.34 Travel times before and after 2010 offset optimization [96].

o_{pc} is the occupancy rate of passenger cars (persons/vehicle);

u_{pc} and u_{hv} are the unit values of time for passenger vehicles and heavy vehicles;

k_{pc} and k_{hv} are the unit values of reliability for passenger vehicles and heavy vehicles; and 60 is a conversion from minutes to hours, and 364 annualizes the results.

This equation yields the user cost in dollars per year. The k factors convert the value of travel time reliability to the same units of the travel time. For example, $k = 1.0$ means that one unit change in the standard deviation of travel time is equal to one unit change in the average. We used the values $k_{pc} = k_{hv} = 1.0$, $o_{pc} = 1.25$ persons/vehicle, $u_{pc} = \$17.67$ per hour, and $u_{hv} = \$94.04$ per hour.

Figure 6.38 shows the total user costs for each year in which the corridor was retimed. The number at the

top of each series shows the percent reduction in user costs due to retiming. The charts show the overall cost increasing due to the volumes increasing. Saturday user costs (Figure 6.38a) increased by 22% and weekday user costs by 38% between 2010 and 2015.

However, although total user cost increased, reductions in user costs were achieved for both directions within each year thanks to the optimization. The largest reduction, -8.5% , was observed for 2010 Saturdays. Subsequent optimization years saw a slightly smaller reduction for Saturdays. For the weekdays, the amount of reduction varied from -3.5% to -7.8% . The average decrease in user costs was about 6% per optimization.

The total user benefits are shown in Table 6.2. The individual results by direction and Saturday/Weekday show some directional tradeoffs that yielded some disbenefits in exchange for greater benefit in the opposite direction. The totals for each year in both

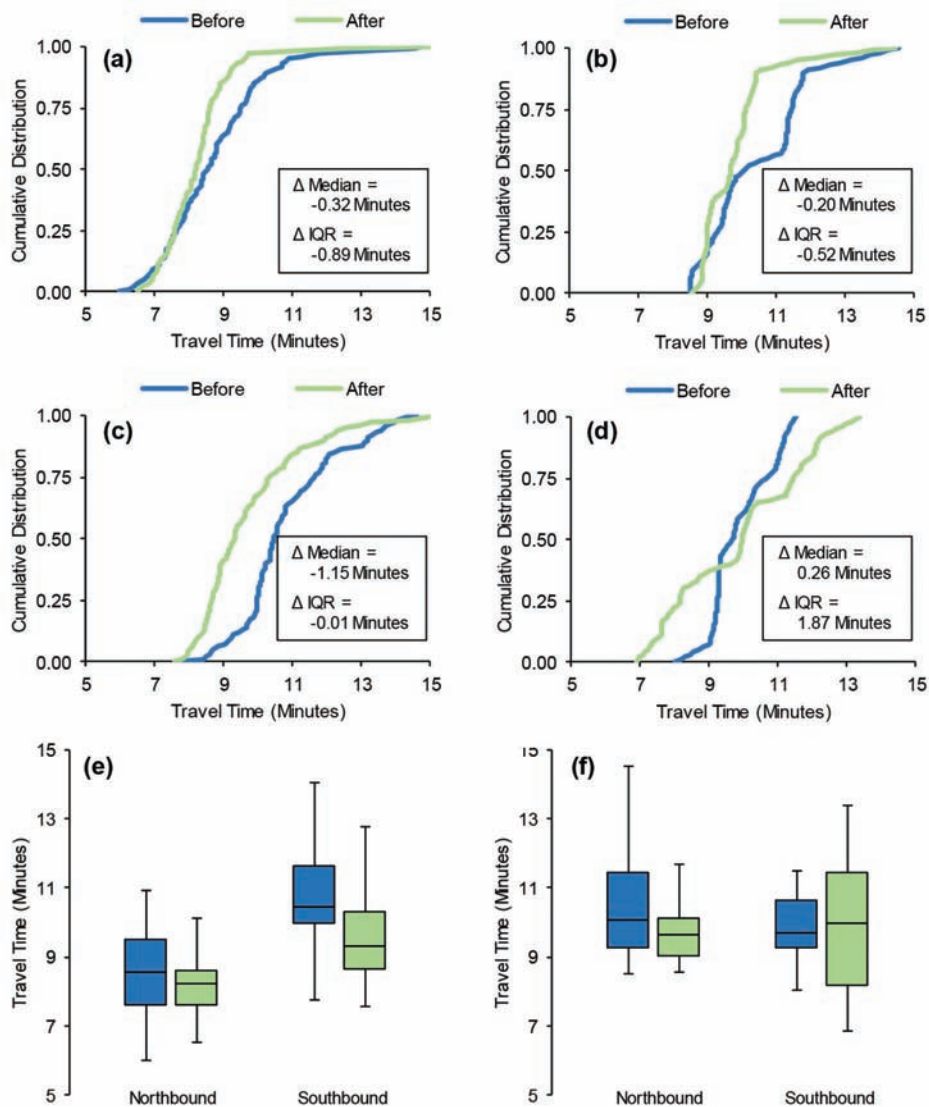


Figure 6.35 Comparison of offset optimization outcomes. (a)/(c)/(e) demonstrate northbound and southbound travel time improvements, and (b)/(d)/(f) show northbound improvements at the expense of southbound vehicles [52]. (a) Northbound before/after travel time CFDs, 2013, TOD plan 15:00–19:00. (b) Northbound before/after travel time CFDs, 2015, TOD plan 12:00–15:00. (c) Southbound before/after travel time CFDs, 2013, TOD plan 15:00–19:00. (d) Southbound before/after travel time CFDs, 2015, TOD plan 12:00–15:00. (e) Northbound and southbound before/after travel time box-whisker plots, 2013, TOD plan 15:00–19:00. (f) Northbound and southbound before/after travel time box-whisker plots, 2015, TOD plan 12:00–15:00.

directions are a net positive in each case, with 2013 having the largest reduction in user costs. The final total is a user benefit of about \$4.2 million. Figure 6.39 shows a bar chart that shows how the benefit breaks down per year.

6.6 System and Corridor Analysis

6.6.1 Overview

This chapter has presented a series of performance measures that examine vehicle progression, and has applied these to various individual use cases. Another use of the data is to track the performance of a system

over time. This section examines the longitudinal analysis of the quality of progression for the system level, with drill down to a corridor.

6.6.2 Network Performance

Similar to split failures, a metric that can show poor progression performance is the number of arrivals on red (AOR). More specifically, anomalous operation should stand out in such a view. Figure 6.40 shows a view of AOR per date for of the systems (see Figure 5.55 for a map showing the system locations). Of these, System #6 was excluded because the intersections in that corridor do not have advance detection.

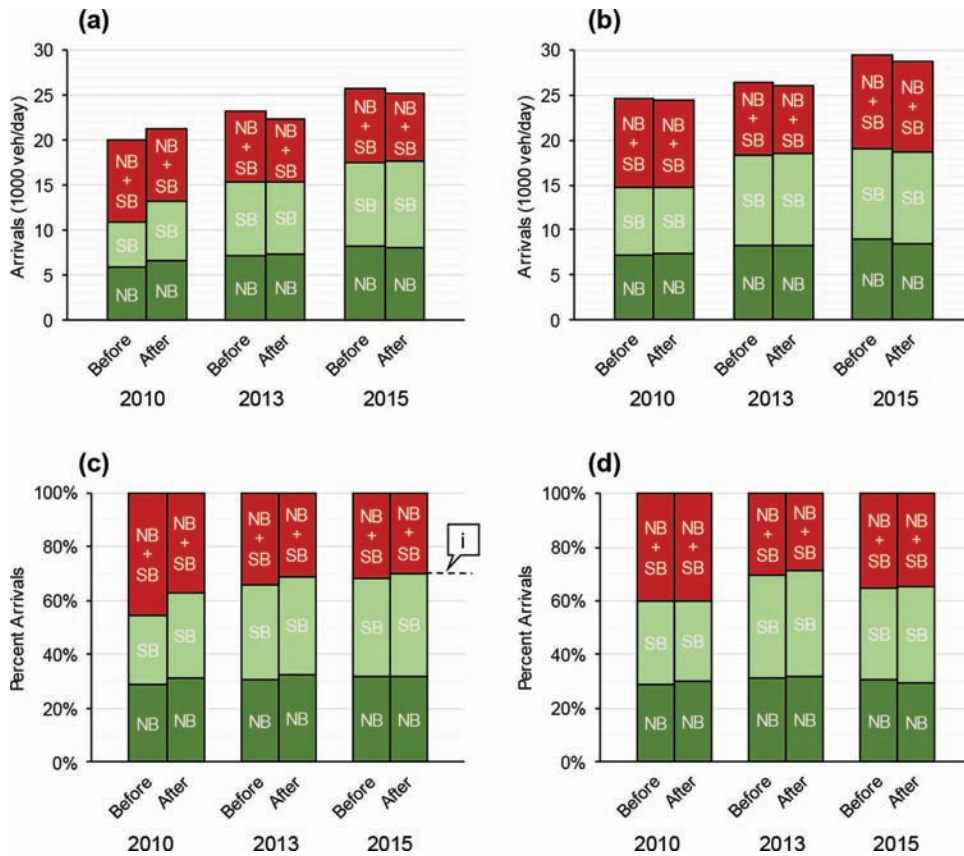


Figure 6.36 Comparison of arrivals on green (AOG) and percent arrivals on green (POG) for weekdays and Saturdays from 06:00–19:00, before/after offset optimization [52]. (a) Average AOG and combined northbound/southbound arrivals on red (AOR), per intersection, Saturdays. (b) Average AOG and combined northbound/southbound arrivals on red (AOR), per intersection, weekdays. (c) Average POG and combined northbound/southbound percent arrivals on red (POR), per intersection, Saturdays. (d) Average POG and combined northbound/southbound percent arrivals on red (POR), per intersection, weekdays.

The vertical lines in the plot separate weekdays and weekends. The overall trend loosely follows the traffic volume, with there being fewer AOR on weekends and the highest value on Fridays. There are no particular anomalies that stand out; however, there is a trend of increased AOR on System #4 at the end of the series, and of a decrease on System #1 that is apparent in this view.

To separate performance from the volumes, AOR can be converted into the percent on red (POR). Although POR no longer lends itself to a stacked bar view as the sum is not meaningful, it is possible to develop a sorted list of approaches by their performance to find locations where progression is poor. Figure 6.41 shows a chart of the 40 highest POR values among all of the signalized approaches across the eight

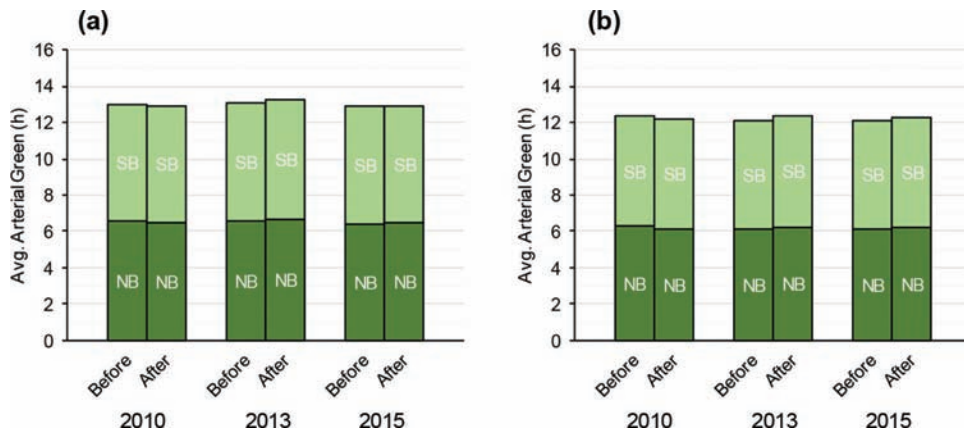


Figure 6.37 Comparison of arterial green time, excluding 135th Street, for weekdays and Saturdays from 06:00 to 19:00, before/after offset optimization [52]. (a) Average Arterial Green (h) for Saturdays. (b) Average Arterial Green (h) for weekdays.

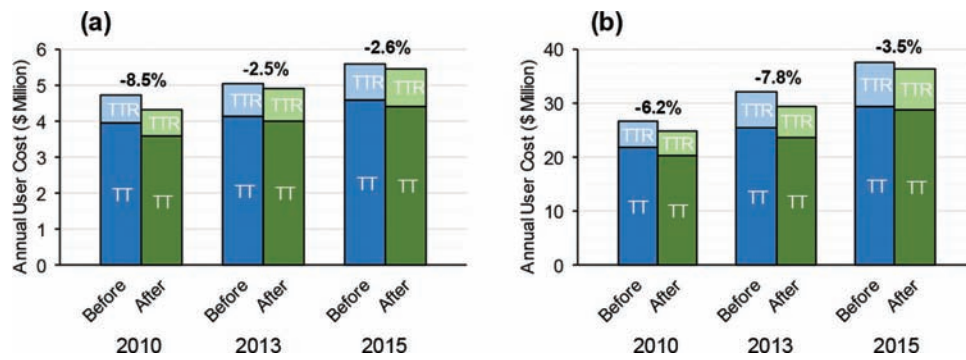


Figure 6.38 Annualized user costs for (a) Saturdays and (b) weekdays, from 06:00 to 19:00, before/after offset optimization [52].

TABLE 6.2

Benefits (user cost reductions) due to offset optimization on SR 37 North, 2010–2015 [52]. All figures in USD. Negative numbers indicate increases in user costs.

	Northbound			Southbound			Both Directions by Year			Grand Total
	2010	2013	2015	2010	2013	2015	2010	2013	2015	
Saturdays	119,458	-13,425	76,220	285,080	139,387	68,429	404,538	125,962	144,649	675,149
Weekdays	1,482,716	1,238,969	-810,762	-948,806	1,273,386	1,331,068	533,910	2,512,355	520,306	3,566,571
Total	1,602,174	1,225,544	-734,542	-663,726	1,412,773	1,399,497	938,448	2,638,317	664,955	4,241,720

corridors included in the study. The value represents the average daily AOR for all the dates in October and November 2015. Notably, there are 23 approaches with AOR above 50%, meaning fewer than half of the vehicles arrive on green.

The next logical step once presented with the list of the worst performing approaches is to begin drilling down to each one to examine what is occurring. In a dashboard environment, one might imagine clicking on each individual bar to go to an “AOR ticker” view, similar to the split failure ticker in Chapter 5, that will show the data for each individual approach. Figure 6.42 shows a ticker view combined with a representative PCD for the top 5 approaches.

1. *System 3, Intersection 2, Southbound.* The ticker (Figure 6.42a) shows a weekly trend in the data, with Sundays having a lower POR than the rest of the week. The PCD (Figure 6.42b) shows that there are certainly opportunities to improve progression, especially during

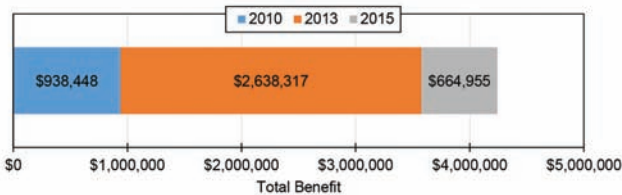


Figure 6.39 Bar chart showing the breakdown of the total user benefit by analysis year [52].

the PM peak. This intersection is located at an interchange with a neighboring local signal that seems not to be well coordinated.

2. *System 8, Intersection 3, Northbound.* The ticker (Figure 6.42c) shows that there are a lot of missing data for this intersection. The PCD indicates that progression is rather poor across much of the day, with none of the green bands really capturing the platoons.
3. *System 1, Intersection 10, Southbound.* The ticker (Figure 6.42d) shows that data here have been missing since mid-November. There does not seem to be much trend in AOR over time. The PCD (Figure 6.42e) shows that arrivals are random at this intersection, and also that the intersection runs free during all of the day except for the 15:00–19:00 hours. This approach is at the entry to a signalized system.
4. *System 2, Intersection 2, Southbound.* The ticker (Figure 6.42g) here shows what appears to be a weekly trend. The PCD (Figure 6.42h) shows what appear to be random arrivals throughout the day. A closer inspection of the intersection finds that this is actually an interstate off-ramp with an advance detector.
5. *System 1, Intersection 9, Southbound.* The ticker (Figure 6.42h) shows that AOR hovers at around 60% most of the time. The PCD (Figure 6.42i) has some interesting properties. During certain times of day, the arrivals appear random, but during the PM peak (15:00–19:00), there is evidence of platoons. This intersection is south of Intersection 10 in the same system, whose southbound PCD is shown in Figure 6.42f. Because Intersection 10 runs free most of the day, it sends platoons south to Intersection 9 during those times of day with no cyclic pattern. During the free portion of the day (15:00–19:00), however, it does send platoons. The high AOR at southbound Intersection 9 suggests there may be some opportunity for improving coordination,

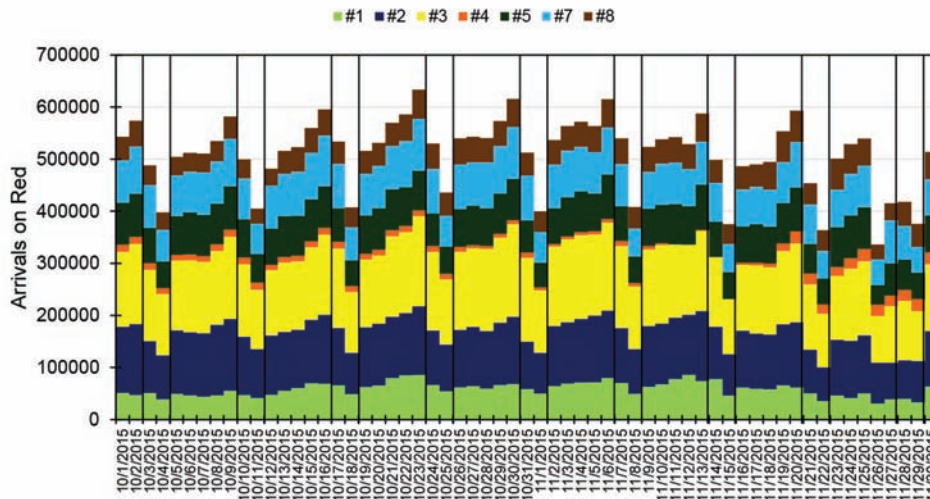


Figure 6.40 Plot of AOR, 10/1/2015–11/30/2015.

perhaps by synchronizing Intersection 10, at least during some other times of day besides the PM Peak.

6.6.3 US 231 Case Study

Overview. This section will present a longitudinal analysis for US 231 between 7/1/2015 and 12/1/2015. As mentioned in Chapter 5, this corridor was affected by the closure of a parallel Interstate route that diverted traffic in the northbound direction onto it [99]. Figure 6.43 shows a map of the detour route. The bridge closure occurred to the east of Lafayette, but because of road work underway on the alternate route through the city, the decision was made to instead route the northbound traffic along a detour further west, as Figure 6.43 shows.

There are 12 operating traffic signals along US 231 on the detour route. Of these, 10 are equipped with high-resolution data. A map of these is shown in Figure 6.44. The other two signals are located on the east–west section of US 231 after it makes a left turn at Intersection #10, and that section is excluded from this analysis. Of the 10 high-resolution data intersections, one (#9) had not yet been completed at the beginning of the detour, and the traffic signals at that location were activated during the detour timeline. Prior to the detour, high-resolution data were only available for Intersections #4 through #10. The other three intersections were brought online during the detour.

The initial closure of I-65 occurred on August 5, 2015, in response to observed structural problems with a bridge over Wildcat Creek in Lafayette. After a brief reopening of I-65, additional bridge movement was observed, and the Interstate closed again on August 7. It was determined that one of the bridge piers would have to be strengthened with micropiles. The closure lasted until September 6.

At first, detour traffic was heavily metered at choke points where the detour route had turns. These were the left turn from US 52 onto SR 28 and the right turn

from SR 28 onto US 231. These resulted in extremely long delays on the first few days of the closure. The efforts of the signal engineers were at first focused entirely on the expedient installation of emergency signals at those locations (as well as a third intersection with SR 47, to improve safety). However, the queuing at these turns initially limited the amount of traffic flowing through the remainder of the detour route.

After the emergency signals had been installed and traffic began flowing into the rest of the system more evenly, engineers deployed a coordinated timing plan on US 231 to handle the diverted traffic. This was a 150-second cycle plan that established very long green bands through the system. After the reopening of the interstate and the end of the detour, this plan ran for about another week until new timing plans for normal traffic were deployed.

In Situ Assessment during the I-65 Detour. The availability of high-resolution data to assess conditions during the I-65 detour was advantageous to engineers trying to assess the situation. The first use of the data was to assess the amount of increase in traffic volumes that affected the detour. This was examined in the last chapter in the case study of US 231 and River Road (Figure 5.47). The volume data were also integrated into a dashboard view to track conditions along the detour route, as shown in Figure 6.45.

The uppermost chart in this figure is a “congestion ticker” showing the number of miles along the roadway operating under 45 mph. Note the very large number of miles in this category at the beginning of the detour, gradually decreasing as improvements are made. In particular, the orange colored section of data representing the signalized portion of US 231 is dramatically reduced following August 15, which shows when the detour timing plan was deployed.

The middle chart shows a heatmap of segment speeds. The red and purple sections show where speeds were less than 25 mph. The beginning of the detour saw

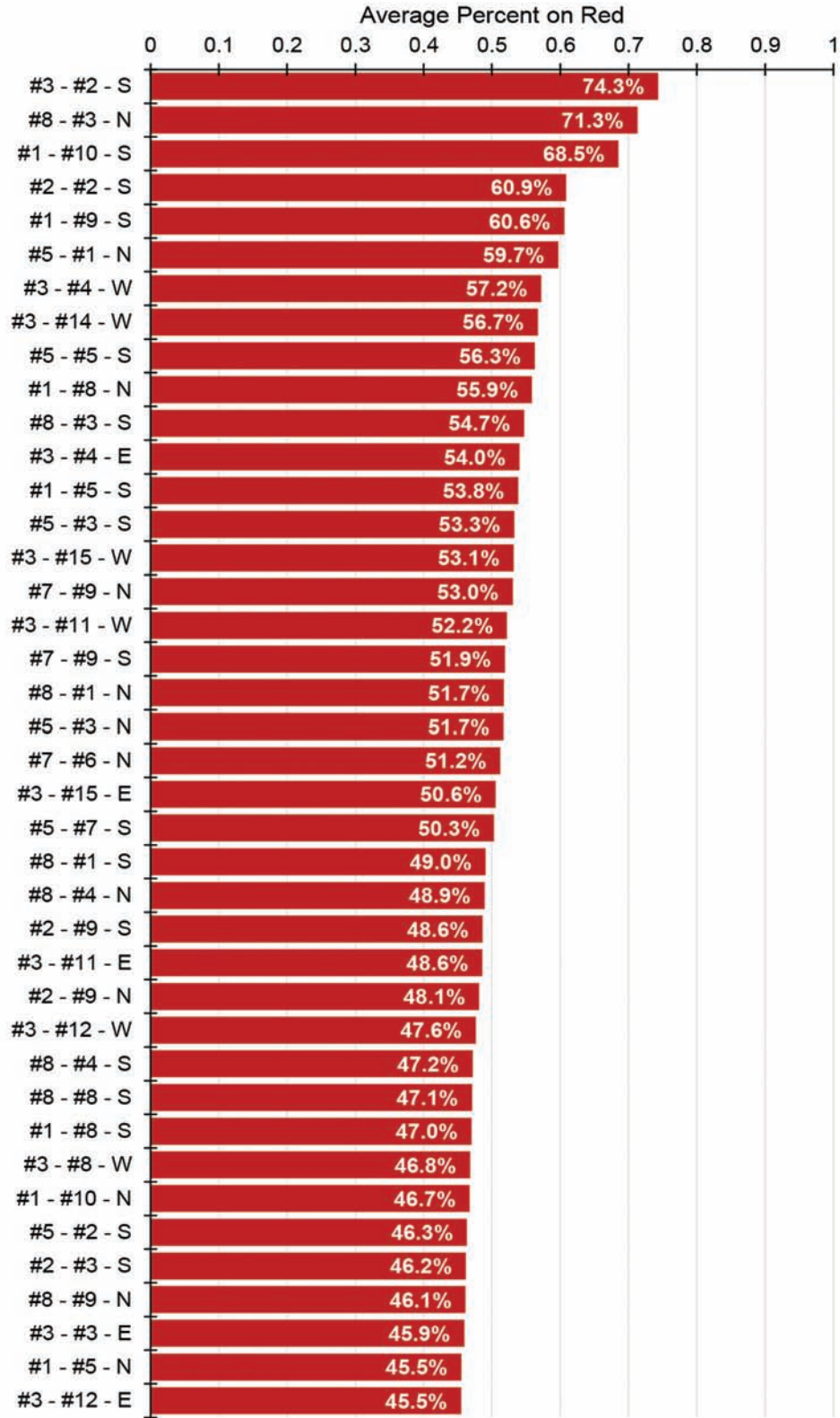


Figure 6.41 Sorted list of average POR values by approach. Data represent the average POR between 10/1/2015 and 11/30/2015. Each row shows the system, intersection, and approach.

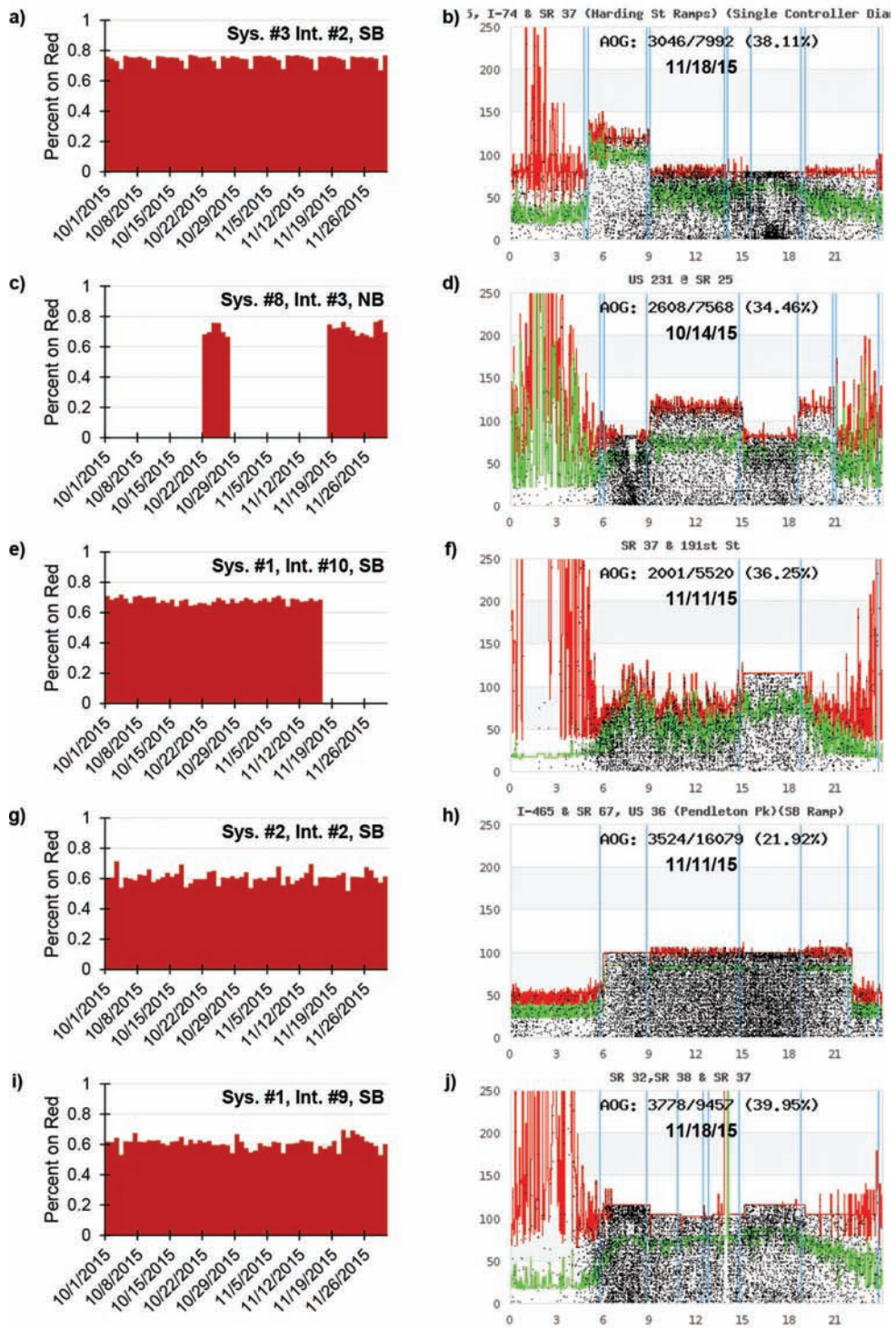


Figure 6.42 The top 5 POR approaches in the INDOT system, 10/1/2015–11/30/2015.

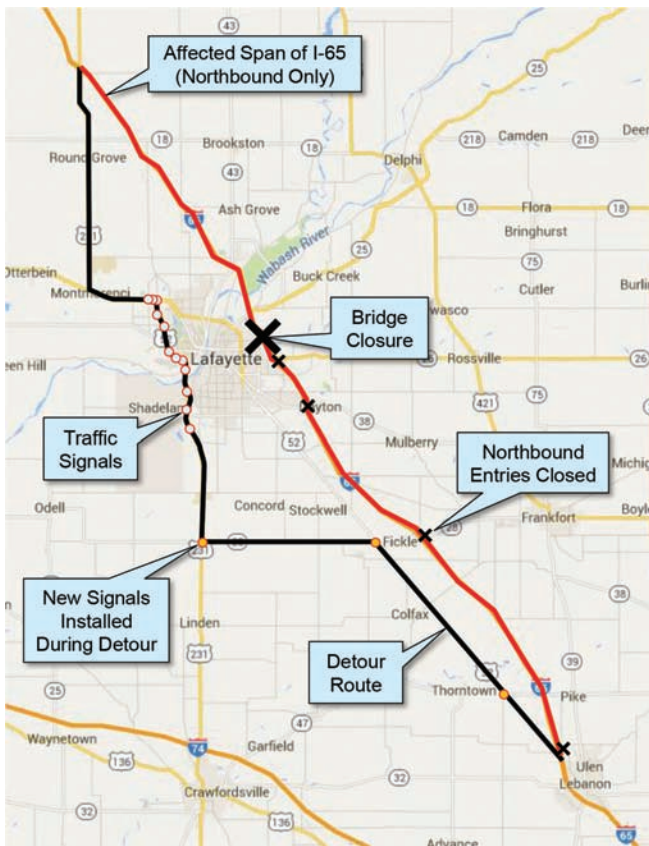


Figure 6.43 Closure of northbound I-65 in August/September 2015, and diversion of traffic onto US 231.

up to 10 miles of very slow traffic, mostly concentrated at the turns at rural intersections along the detour where new signals were installed as quickly as possible. The impact can be seen after August 10, where the purple section of the chart disappears—and also in the upper chart where the pink and blue sections vanish but the orange section enlarges. This shows that after August 10, traffic has been freed up from queues at the rural bottlenecks along US 52 and SR 28, but are now slowing in the signalized portion of US 231. Later, after August 15, the diversion timing plan improves operations further, as the red and yellow portions of the heatmap diminish a bit further.

Finally, the bottom chart in Figure 6.40 shows the traffic volumes measured at several locations along the diversion route using the high-resolution data. These charts show that the volumes appear to actually increase after the beginning of the detour, and hold rather steady through the end. The weekend of August 15–16 has lower volumes than the preceding weekdays, but the following weekend of August 22–23 has considerably higher volumes. That weekend preceded the start of fall semester classes at Purdue University, which is located in West Lafayette.

Performance measure dashboards enabled by automatic collection and processing of high-resolution data also proved helpful in validating the operation of the

detour plan. Figure 6.46 shows a view of the dashboard as it would appear to a traffic engineer. The map indicates the locations of intersections, with the green ones selected to view PCDs. The “panel settings” tab has selected a date and direction for viewing, and it is now possible to scroll up and down in the adjacent panel to view operations through the system. The error message indicates that one of the intersections is missing data. On that particular date, efforts were still underway to retrieve high-resolution data from a few locations.

Figure 6.47 shows the northbound PCDs that are visible on the website for 9/1/2015. The view scans along the detour route, showing the quality of progression at each intersection in order along the roadway. The first intersection (Figure 6.47a) shows random arrivals, as expected given that the upstream intersection is a long distance away. The subsequent intersections (Figure 6.47b–Figure 6.47g) exhibit very good progression for the most part. There is evidence of a secondary platoon at Intersection #4 (River Road), as Figure 6.47c shows. Otherwise, most of the intersections show that the arrivals are coincident with green. This graphical analysis enabled engineers to quickly inspect whether the offsets they had designed for the detour route achieved the intended goal.

Additionally, the engineers responsible for deploying the plan were given a directive by the director of traffic management to achieve as high of a POG as possible along the detour route. The PCDs show that the POG was above



Figure 6.44 US 231 in West Lafayette, Indiana.

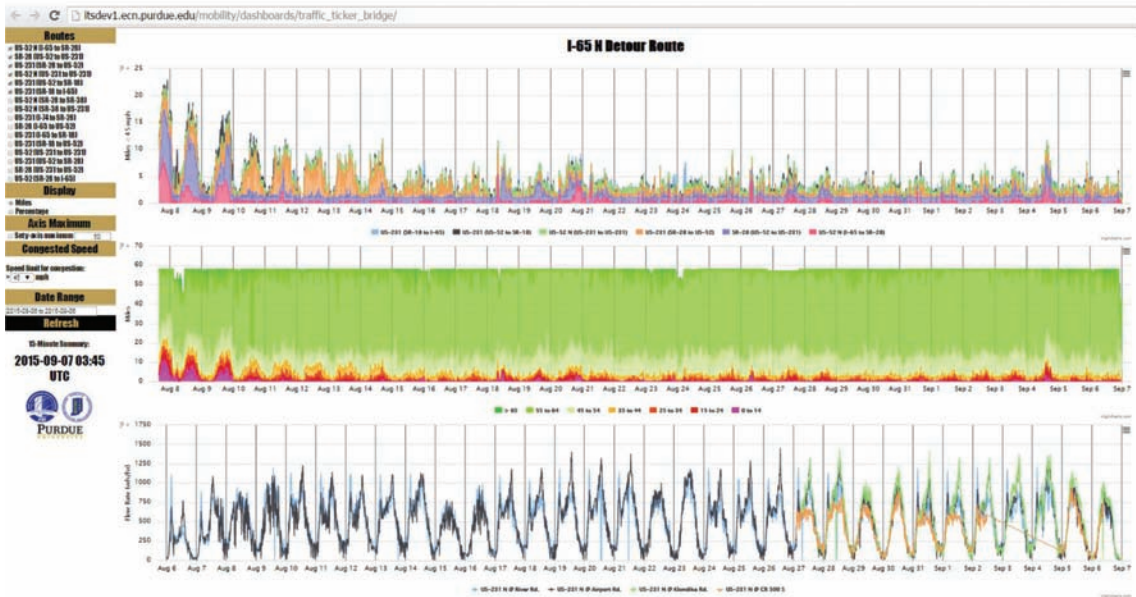


Figure 6.45 Dashboard showing integration of signal data (flow rates measured using high-resolution data) with probe vehicle speed data.



Figure 6.46 View of the INDOT performance measure dashboard.

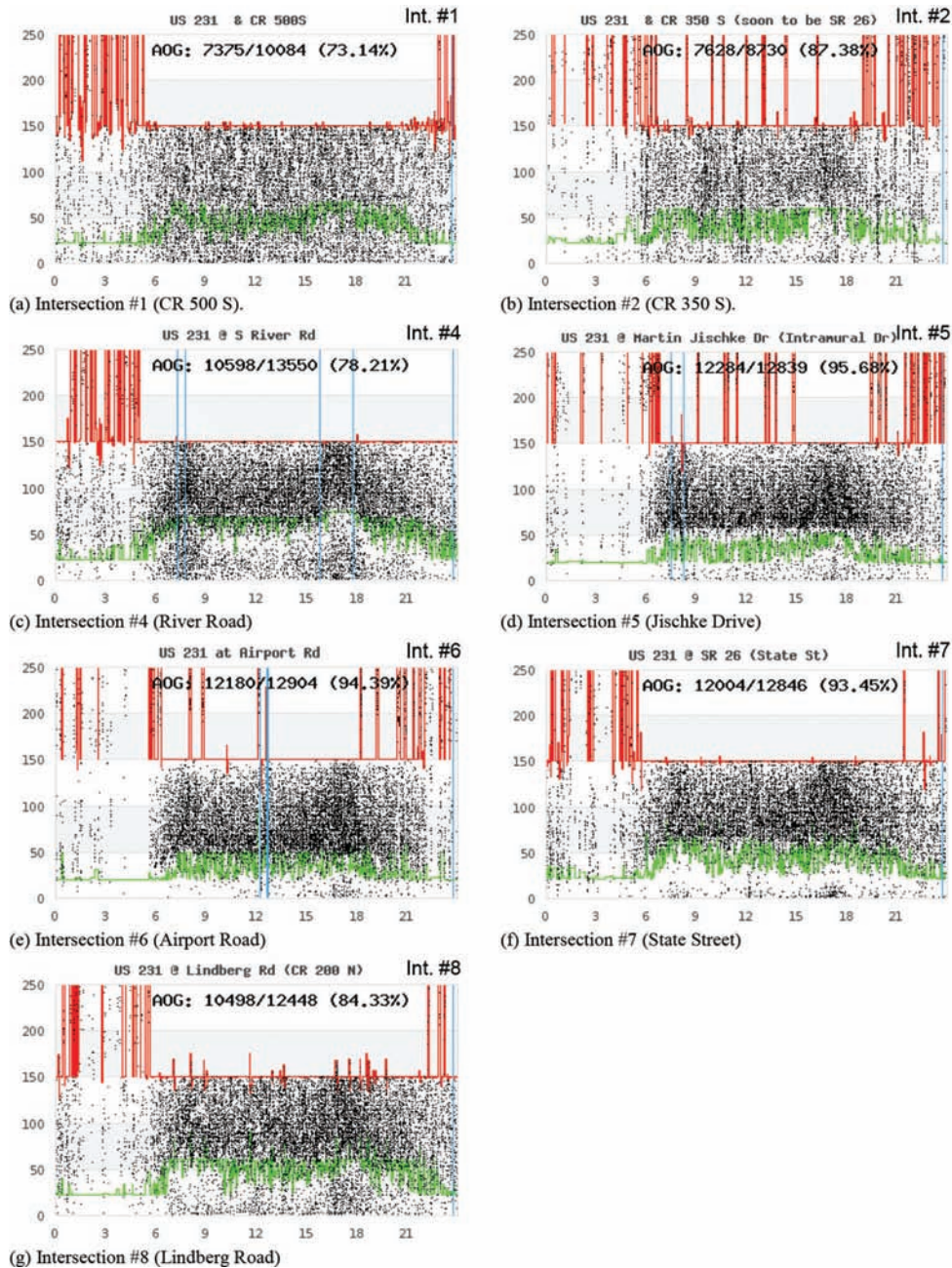


Figure 6.47 Northbound PCDs from 9/1/2015 along US 231, as visible from the INDOT performance measure dashboard.

80% at all of the approaches, except for the entry location and Intersection #4, which still had values above 70%. The PCD at Intersection #4 shows that the green is fully utilized by vehicles, so the POG of 78% is probably about as good as it possibly could be. Three of the approaches had POG above 90%. This is an effective way to document that the operational goal was largely achieved.

Corridor-Level Longitudinal Assessment. At the corridor level, it is feasible to begin viewing the percent on green (POG) for multiple signalized approaches. The US 231 corridor includes 19 approaches in total along

the mainline. Figure 6.48 shows a bar chart that shows the daily POG value for each approach from 7/5/2015 through 11/29/2015. The height of the green bar represents the POG; the black regions show where data are missing, and the gray region marks an approach that does not exist. The vertical lines separate weekdays and weekends.

In terms of the POG, prior to the detour in July, there are some differences between weekday and weekend operation, but overall most intersections have a rather high POG value. The start of the detour on 8/7/2015 has rather little effect on the POG. The

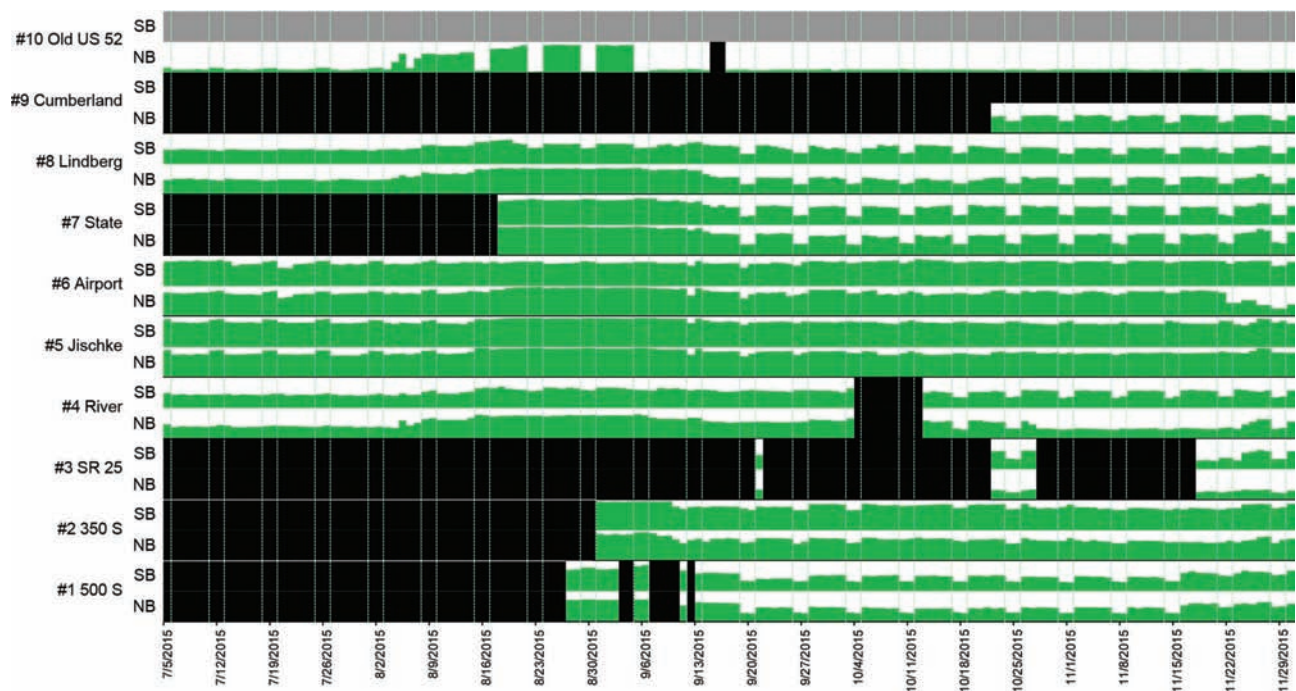


Figure 6.48 Longitudinal POG for US 231 corridor, 7/1/2015–12/1/2015 (overall 24-hour performance shown).

implementation of the detour timing plan increases POG to a very high value for most of the intersections. The most dramatic approach is NB at Intersection #10, which goes from a very low value to a very high value. The reason for this is that this intersection was originally coordinated for the crossing street. Note also that the low POG persists on the weekends, because the detour timing plan only ran during weekdays. After the end of the detour, the POG values generally decrease with the end of the detour timing plan. There is a distinct difference in performance on the weekends. This view also shows the timeline of when intersections were brought online, and when others had data outages.

Figure 6.49 shows a view of AOR per intersection by date from 7/1/2015 through 12/1/2015. These reveal much of the dynamics of the system. The northbound direction (Figure 6.49a) sees a sharp increase in AOR after the start of the detour. Note in particular the spike on 8/5/2015, the decrease for one day (8/6/2015) when the Interstate was reopened, and the increase one day later when the Interstate was closed a second time. Then, after 8/14/2015, there is a sharp decrease in AOG after the deployment of the coordinated timing plan. There are some increases in AOG due to new intersections being brought online. After the end of the detour on 9/9/2015, AOR falls sharply; it then increases again when new timing plans are deployed, although this again is elevated by the stabilization of data collection at Intersections #1 and #2.

The southbound direction (Figure 6.49b) reflects some similar trends, although it did not see any increase in AOR during the initial diversion after 8/5/2015, because no additional traffic was added to this direction.

Figure 6.50 presents a view of AOR by time of day by intersection, summed across the entire analysis period. The view shows the presence of northbound peaks in the morning and afternoon, and a southbound peak in the afternoon.

Drill-Down to Intersections and Approaches. The longitudinal POG chart in Figure 6.48 includes some approaches that have some interesting dynamics. A few of these are worth drilling down to the intersection level to examine these in more detail. First, Intersection #4 (River Road) will be examined to look at changing corridor dynamics during the I-65 detour. Next, Figure 6.48 exhibits some degradation of POG in the last two weeks at Intersection #6 (Airport Road). Last, there is an interesting pattern at Intersection #10 (Old US 52) that will be examined.

The ticker view is helpful for showing trends over time for multiple corridors, but has the limitation that it does not show variation by time of day. An alternative view is shown in Figure 6.51. This graphic is a heatmap of the POG, with the time of day displayed vertically, and day of year displayed horizontally. The color of each cell is scaled between 0% and 100%, with missing data indicated by white.

Figure 6.51a shows a heatmap of northbound POG at Intersection #4. Scanning horizontally across the graph, during the morning hours, POG is consistently higher, whereas between 9:00 and 18:00, it is often lower. Interestingly, during the I-65 diversion, POG is much improved, especially between 8/12 and 9/9. This is most likely due to the use of the detour timing plan, which produced very large green bands in the northbound direction. After 9/16, there is more variation by

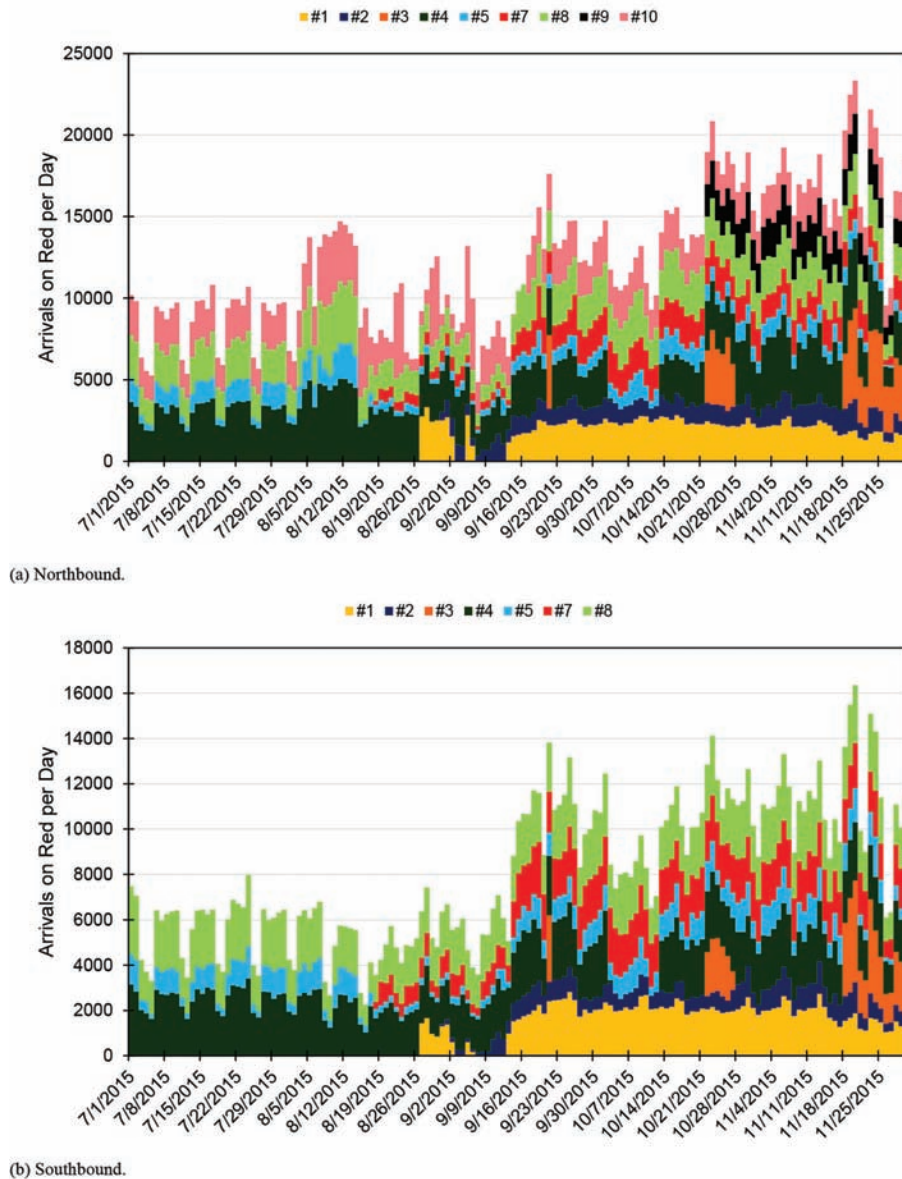


Figure 6.49 AOR per intersection by date, 7/1/2015–12/1/2015.

time of day, and patterns seem to shift considerably as timing plans are adjusted. Some missing data are apparent for several days in early October. The initial timing plan running before 10/28 seems to produce better POG than the one that follows 10/28, although there is a pattern that shows that the weekends do not seem to operate as well. In the southbound direction, as Figure 6.51b shows, there is an apparent problem with low POG during the PM peak, preceding 8/12. On 8/13 and after, with the detour timing plan in effect, this improves. Finally, following the gap in data in October, the pattern falls into one with lower POG on weekends and higher POG on weekdays through most of the day.

Figure 6.52 shows PCDs for southbound and northbound approaches at Intersection #4 that explore these dynamics in more detail.

- On 7/22/2015, prior to the detour (Figure 6.52a and Figure 6.52b), the intersection runs free, except for the 6:00–9:00 period. This accounts for the lower POG after 9:00 in the heatmap.
- The 8/12/2015 data (Figure 6.52c and Figure 6.52d) show conditions after the detour occurred but before the detour timing plan was implemented. The cycle length and the green time have increased in both directions, but outside of the 6:00–9:00 period there does not seem to be any consistent arrival pattern. This explains why the POG values would have improved (with longer green time).
- On 8/19/2015, after the detour pattern was implemented, there is a very clear difference in operations. The southbound direction (Figure 6.52e) sees the arrival of two platoons, which are partially coincident with green, while the northbound direction (Figure 6.52f) successfully captures most of the largest platoon. This would

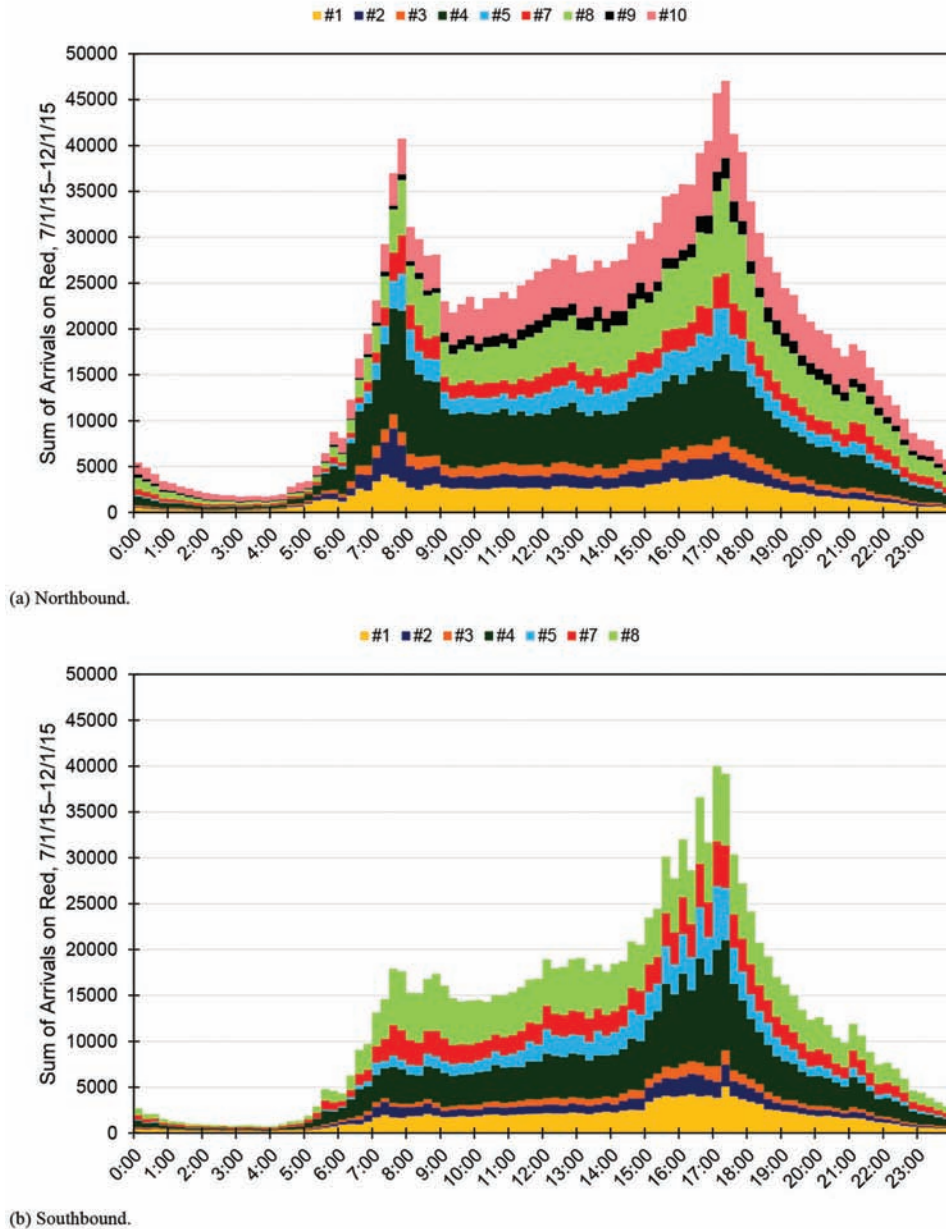


Figure 6.50 AOR per intersection by time of day, 7/1/2015–12/1/2015.

- account for the higher POG values associated with late August and early September in the heatmap.
- On 9/9/2015, after the Interstate was reopened, but the detour pattern was still operating (Figure 6.52g, Figure 6.52h), there is less traffic than one week before. Most likely, the POG values remained the same, as the overall patterns did not change as much.
- Finally, on 9/16/2015, after implementation of a new timing plan for normal conditions, the cycle lengths are reduced and the patterns have changed again (Figure 6.52i and Figure 6.52j).

At some times, the heatmaps (Figure 6.51) show a marked difference between weekend and weekday operations, as shown by the lower POG values for

two days in a row, followed by five days with higher POG values. Figure 6.53 compares PCDs at Intersection #4 for weekday and weekend operations to discover why this is the case. Figure 6.53a and Figure 6.53b show that the intersections are clearly coordinated during weekdays, and Figure 6.53c and Figure 6.53d show that the intersections operate free. This explains the lower POG on the weekends. Also, the progression in the northbound direction appears relatively poor during the week (Figure 6.53b); this corresponds to the higher POG values seen in Figure 6.51a in late October and early November. The heatmap shows that this was improved during the last week of November.

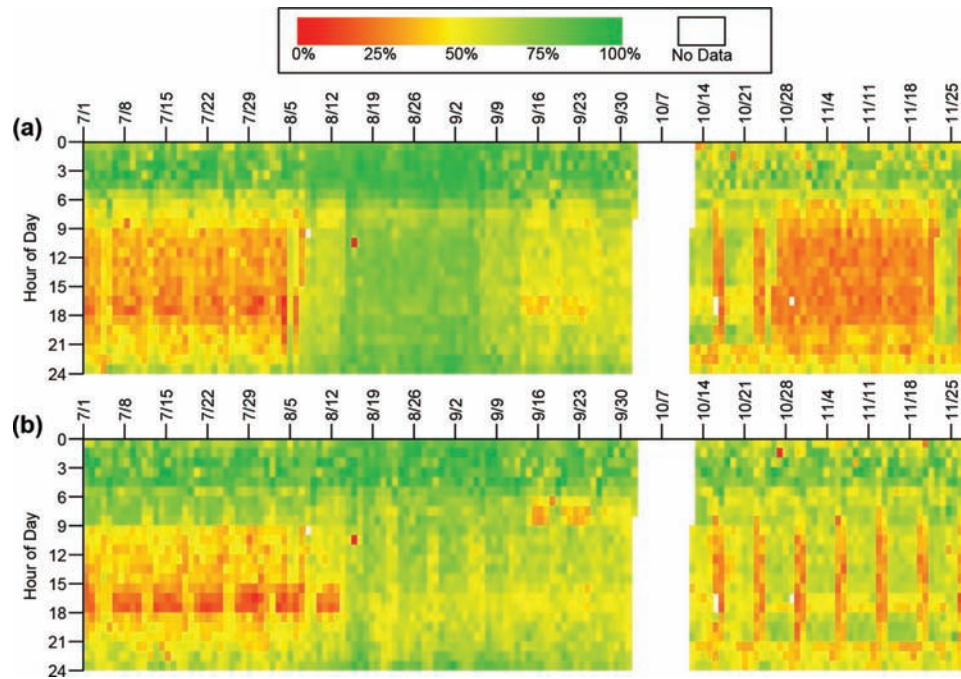


Figure 6.51 POG heatmap for Intersection #4 (US 231 and River Road), 7/1/2015–12/1/2015: (a) northbound; (b) southbound.

In Figure 6.48, it can be seen that Intersection #6 (Airport Road) experiences a fairly dramatic decrease in POG in the last week of the study period. Figure 6.54 presents a heatmap to explore this further. Figure 6.54a shows the westbound direction, which is generally northbound along the corridor, and Figure 6.54b shows the eastbound direction, which is generally southbound along the corridor.

It is worth pausing here briefly to highlight the difference between the local direction and the system direction. Owing to the bend in this arterial route (Figure 6.44), the local intersection is oriented east–west along the mainline, although the arterial itself follows a north–south route. A dashboard or other utility that displays system performance measures will need to find an effective way to negotiate potential ambiguity.

Throughout most of the day and through most of the study period, the POG values are relatively high in both directions, above 75% for many, with only a few intermittent exceptions. Then, after about 11/20, the northbound direction sees POG fall to a lower level (Figure 6.54a). The southbound direction has a high level of POG throughout the entire study period (Figure 6.54b).

Some sample PCDs to explore the reason for the difference are presented in Figure 6.55. Here, we have selected 11/13/2015 as a date where POG was relatively high, and 11/27/2015 as a date where it was relatively low. The earlier date exhibits what seems to be rather good progression in both directions (Figure 6.55a and Figure 6.55b), with the arrivals clustered well within the green in both directions. However, the later date

exhibits what appears to be a very different timing plan. Progression on the local eastbound approach (Figure 6.55c) appears to be relatively good, which corresponds to the green coloring of the heatmap in Figure 6.54b. However, the local westbound approach (Figure 6.55d) is rather poor, particularly in the afternoon. The PCD also shows that green time along the arterial appears to be truncated. Note that even during the overnight period, the signal cycles to some noncoordinated phase regularly, and does not appear to rest in green when demand is light. This suggests that either a side-street phase is in recall. It is unlikely that it would have been programmed as such, so a detector fault is likely. The impact of the side-street detector failure is apparent from the POG data.

The northbound approach at the north end of the system, Intersection #10 (Old US 52), has a pattern during the I-65 detour that appears to alternate between high and low POG between weekdays and weekends in In Figure 6.48.

Drilling down to the intersection, a heatmap view shows the same dynamic occurring, as presented in Figure 6.56. In the northbound direction (Figure 6.56a), there appears to be a very high POG on weekdays and a low POG on weekends. This movement represents the upper end of the arterial, and it was coordinated in this direction during the I-65 detour—but only during weekdays, it would appear. During ordinary conditions, this intersection is coordinated for the east–west direction. Figure 6.56b shows a heatmap of POG for the westbound approach, which looks like the inverse of the northbound. Here, the weekends seem to have higher POG in general (except for the middle of the day) and the weekdays have

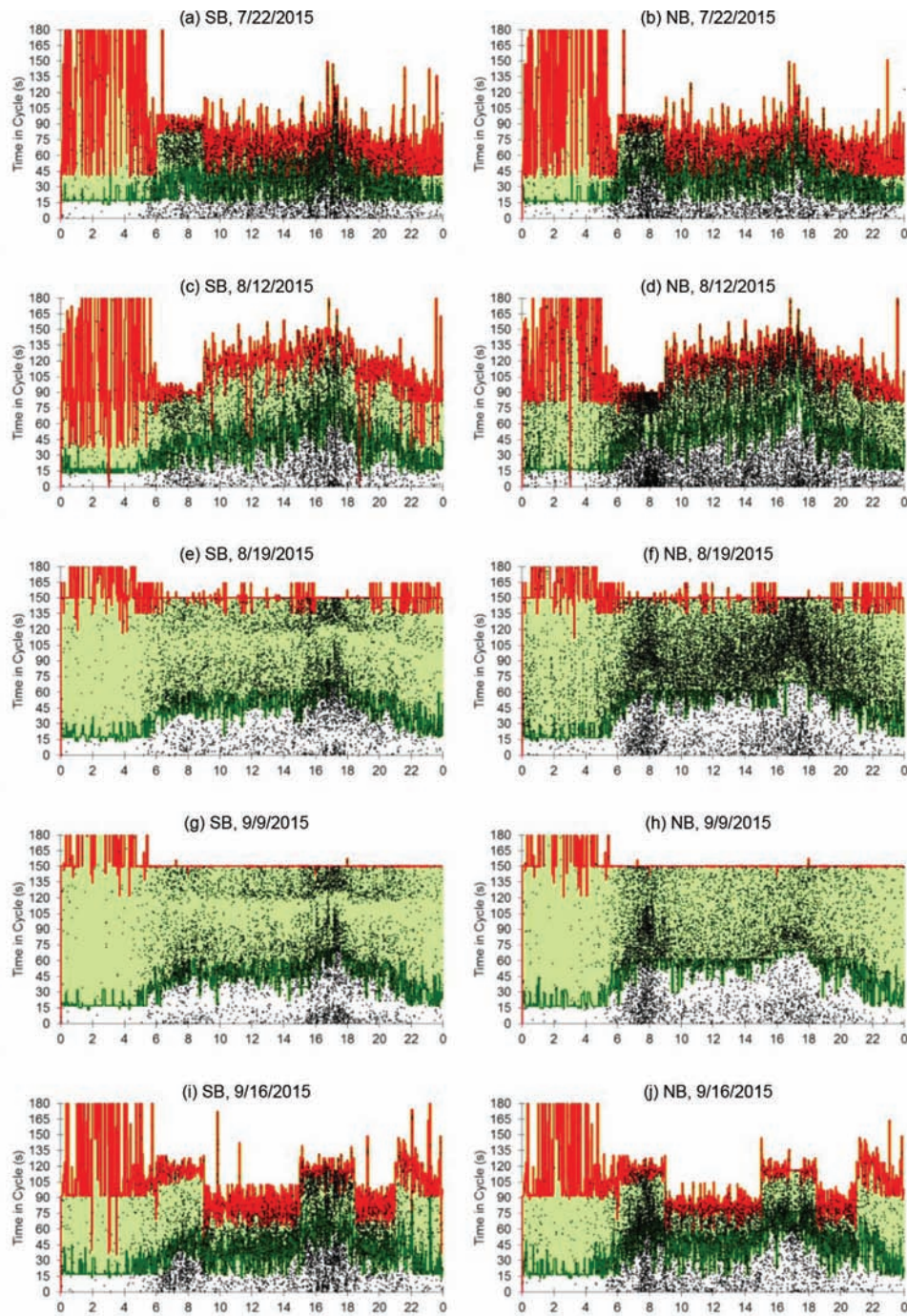


Figure 6.52 Selected PCDs at Intersection #4 (US 231 and River Road), for various times during the I-65 diversion.

lower POG. The heatmap also shows that performance changes over time for the westbound approach, if we compare weekly operation before 8/5 and after 9/9. Earlier in the analysis period, there is some vertical orange and red striping that corresponds to weekends with low POG. Later in the analysis period, the weekends have improved but still appear to be worse than the weekdays.

Selected PCDs for a couple of different dates reveal the impacts of changing the coordinated phases at the

intersection. Figure 6.57a and Figure 6.57b respectively show northbound and westbound coordination during the I-65 detour, on 9/2/2015. The northbound direction exhibits very good progression, although the end of the platoon appears to be truncated at times. However, the overall POG is high, corresponding to the green regions of the heatmap in Figure 6.56a. Arrivals in the westbound direction (Figure 6.57b) appear to be random, although there is another intersection within a mile to

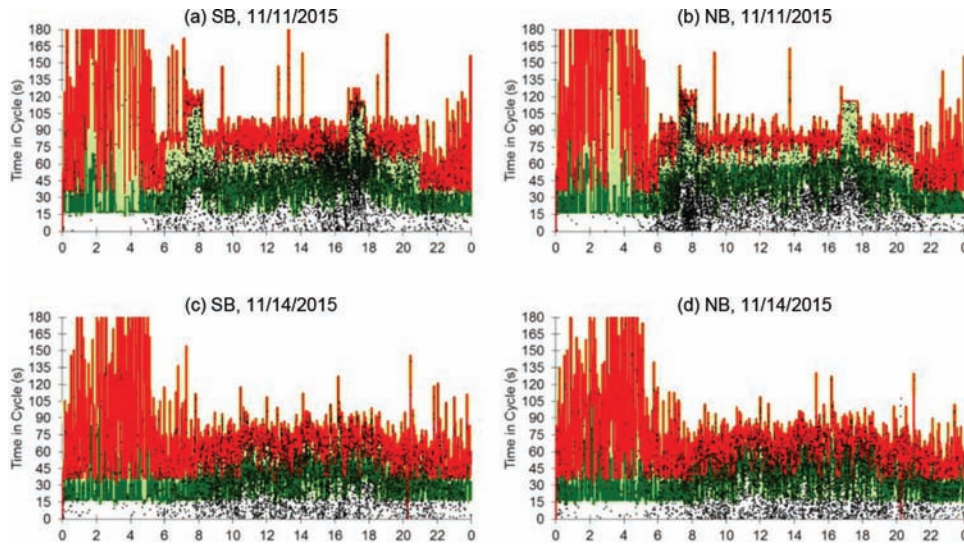


Figure 6.53 Selected PCDs for Intersection #4 (US 231 and River Road) showing differences between weekday (11/11/2015) and weekend (11/14/2015) operations.

the east. It is likely that the neighboring intersection was running free at this time.

Operations after the detour are markedly different. The northbound direction (Figure 6.57c) is no longer coordinated, and the duration of green is much smaller than before. The afternoon period between 15:00 and 19:00 exhibits the diagonal striping that is typical of a cycle length mismatch, as shown earlier in Figure 6.11. However, during the rest of the day, the arrivals are fairly sparse. The westbound direction, in contrast, has very evident platoons as Figure 6.57d shows. These are

well coordinated in the afternoon, but seem to be poorly coordinated at other times of day. Referring back to the heatmap in Figure 6.56b, a dark green trough is evident for weekdays in the vicinity of 9/30.

6.6.4 Summary

This example analysis showed several ways of viewing data on signal progression that go from the system level down to the intersection level, and from several months down to 24-hour periods. These views

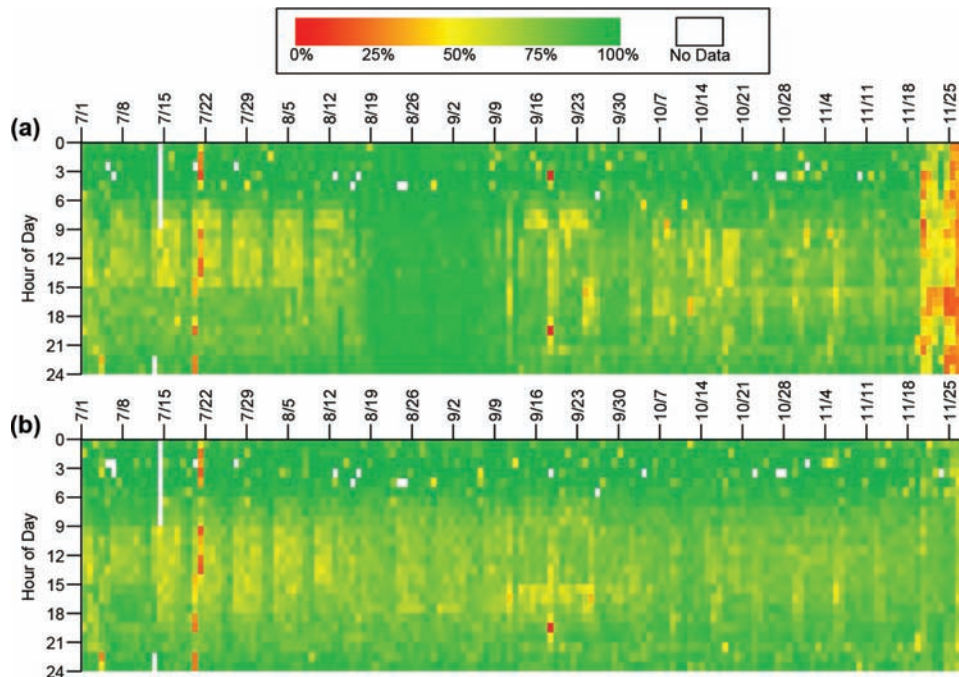


Figure 6.54 POG heatmap for Intersection #6 (US 231 and Airport Rd.), 7/1/2015–12/1/2015: (a) westbound (northbound along the corridor); (b) eastbound (southbound along the corridor).

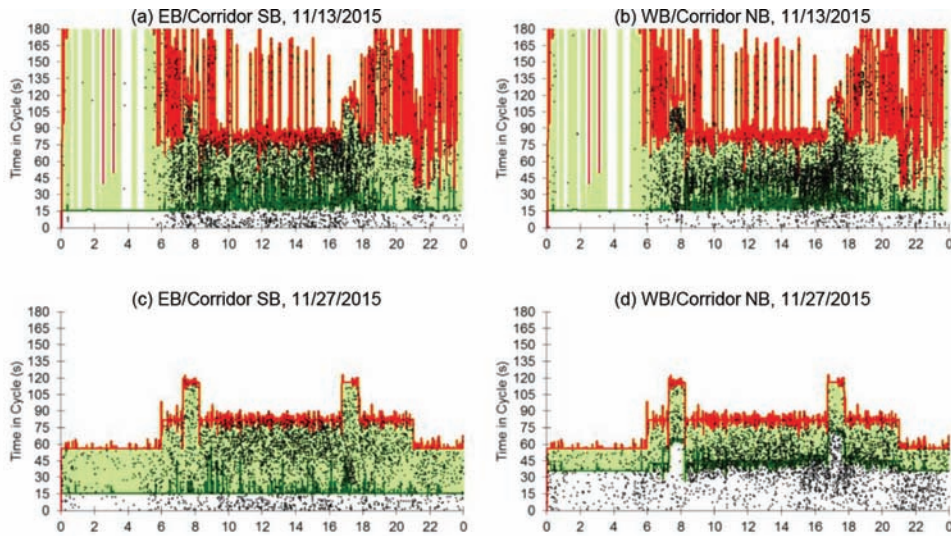


Figure 6.55 Selected PCDs for Intersection #6 (US 231 and Airport Road), showing change in signal timing plan related to reduction in POG observed at a higher level.

would provide valuable intelligence for engineers implementing new timing plans, inspecting opportunities for improving existing ones, and managing a scenario such as the I-65 detour.

6.7 Chapter Summary

This chapter discussed several applications of high-resolution data and travel time data for evaluating and managing signal progression. Basic progression concepts were discussed, and several performance measures

and visualization graphics were presented. Examples of operational diagnostics were presented. Methods of optimizing signal offsets with high-resolution data were exhibited and their impact demonstrated through previous field applications. A case study of network and corridor analysis was presented for the INDOT system and for a selected corridor.

The tools presented here represent a formidable addition to the analysis toolbox of any traffic engineer, and would make a good addition to the features of any central management system or advanced control system.

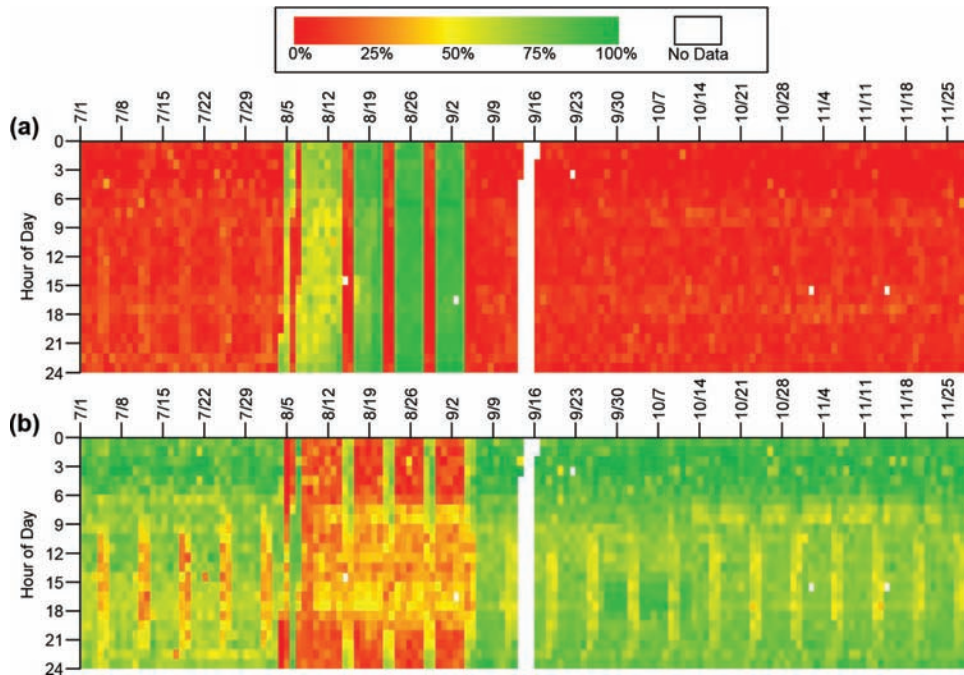


Figure 6.56 POG heatmap for Intersection #10 (US 231 and Old US 52), 7/1/2015–12/1/2015: (a) northbound (arterial left and right turns); (b) westbound (crossing arterial).

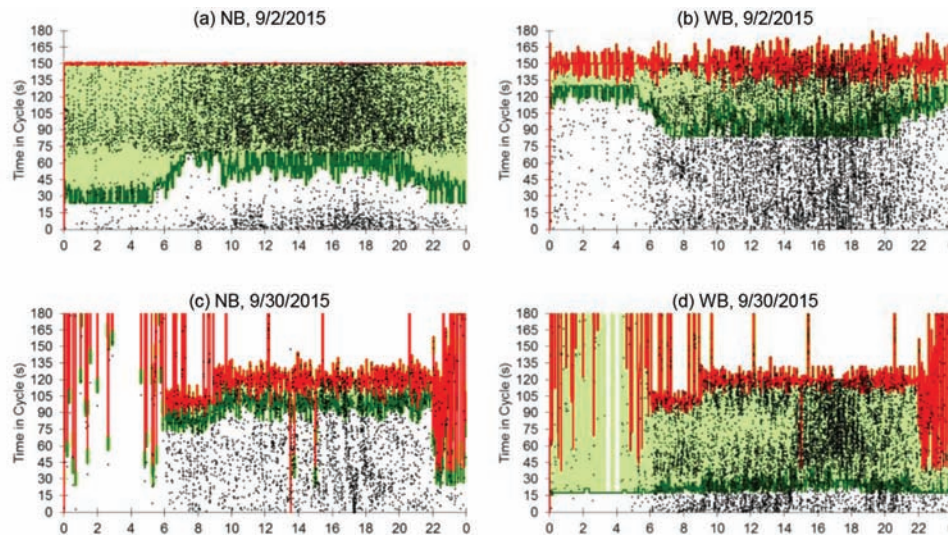


Figure 6.57 Selected PCDs for Intersection #10 (US 231 and Old US 52), showing change from coordination along US 231 to coordination along Old US 52.

The ability to transparently view both the operational details of coordinated signal operations would be helpful to both the improvement of existing plans and also to validate new plans and even new schemes of traffic control. These, combined with external analysis through the means of travel time measurement, form a complete package for analyzing the quality of progression. If further fused with a system for managing signal controllers, they would form a comprehensive utility for maintenance of coordinated signal systems.

7. CONCLUSION

7.1 Report Summary

This report has included a broad discussion of performance measure applications for traffic signal systems, supported by many examples. The opening chapter described the state of the practice and discussed issues related to making agency business processes integrated with performance measures, within the domain of traffic signal operations. There then followed a discussion of the technical requirements needed to implement performance measures.

This was followed by four chapters that presented a variety of uses of performance measures tailored to four categories common to most signal system applications:

- Chapter 3 presented performance measures for evaluating the status of communications systems for traffic signal systems and included a statewide example for the INDOT signal system network.
- Chapter 4 presented performance measures for assessing detector health in traffic signal systems and included both examples of field evaluations and measurement of the impact on mainline and side street phases.
- Chapter 5 presented performance measures for assessing local signal control, including evaluation of capacity allocation, safety performance (red light running), pedestrian service, and diagnostic analysis of preemption

and advanced control settings. The performance measures were demonstrated in several case studies for each of these. A single-intersection case study was presented where operations were thoroughly examined as the intersection had shifting dynamics during a month-long incident on a parallel freeway. Finally, the intersection measures were rolled up to corridor and network views, enabling the identification of problems from the top down.

- Chapter 6 presented performance measures for assessing system control, focusing on traffic progression. This included visualizations of traffic events and development of quantitative performance measures that can be aggregated. The use of the performance measure graphics were demonstrated in use cases including evaluation of common signal operations phenomena, before/after studies, and diagnostic evaluation of advanced control. The use of the same data for system optimization was also presented. The chapter concluded with network and corridor performance measures, again enabling identification of problems from the system level.

With the 2014 Pooled Fund report [2] having provided a catalog of performance measures with details on their computation, the present report has provided more of a “recipe book” to show ways of using the performance measures, with examples that are relevant to concerns commonly shared among many agencies, as discussed in the Basic Service Model (Table 1.3) [7].

7.2 Path Forward for Implementation

At this stage, a well-defined group of signal performance measures now exists, based on high-resolution data from a phase and detector data perspective. Implementation of these will depend on the identification of feasible ways to build out a system to obtain the data. One major advantage is that the methodology can use existing detection to do so in many cases, and will often need only a controller upgrade and perhaps the

addition of network devices (such as a cellular internet access points) to begin data collection. This has largely been the case in the INDOT network. For an agency such as UDOT, which previously invested in an excellent statewide communications system, the upgrading of controllers has been the main implementation step, with the result that UDOT now has performance measures implemented in most of its signal system network. At the time of writing, several other states and local agencies have begun implementing similar systems for signal performance measures.

Establishing communications will be a challenge for some agencies. However, connectivity is becoming ubiquitous as the cost has declined. In the long run, the challenge will likely be met. However, it remains to be seen whether this will continue to be the domain of agencies piecing networks together with incremental addition of network devices, or whether vendors and/or consultants will come to offer connectivity and/or performance measure analysis as a service.

Other agencies will have challenges in achieving functional and accurate detection systems to enable performance measures. High-resolution data can reveal where detection systems have problems, and no detection system is completely exempt from the possibility of failure. To date, such data has not been available before. Time will tell whether increased scrutiny precedes improved performance.

The procurement of a data system to bring in all of the information and to process it is perhaps the biggest gap to bridge for most agencies, particularly smaller ones who might not have access to ample IT resources. Solutions may emerge from some different sources. One possibility is that vendors or consultants will package this as a product or service. Another is that agencies may cooperate to procure regional systems with shared data.

7.3 Future Fusion with Connected Vehicle and Similar Data Types

Detector and phase information consists of simple state data, which for detectors is simply a Boolean contact closure: ON or OFF. In future, these data will likely be augmented by new sources of data that can provide additional information. One potential source is connected vehicle data, which will provide a great deal of information about approaching vehicles. The content of these data is still being determined to some degree, and it will likely be several years before they become available on a widespread basis.

An additional potential data source that might provide similar characteristics is commercial probe vehicle data. To date, the data providers have mainly offered products consisting of aggregated segment speeds and other similar types of analytics based on analysis of the raw data. However, the density of the raw data is approaching the point where it may become useful for evaluating arterial performance on a microscopic level.

Figure 7.1 shows plots of vehicle trajectories based on raw vehicle position data from a section of SR 37 North. The positions of intersections are indicated by the horizontal lines that connect to the inset map. Figure 7.1a shows southbound vehicles and Figure 7.1b northbound vehicles.

The data show the variety of reporting intervals that different vehicles exhibit. There are some vehicles that report their position about once every second (Figure 7.1a, callout *i*), others about once or twice a minute (Figure 7.1a, callout *ii*), and others once every few minutes (Figure 7.1a, callout *iii*).

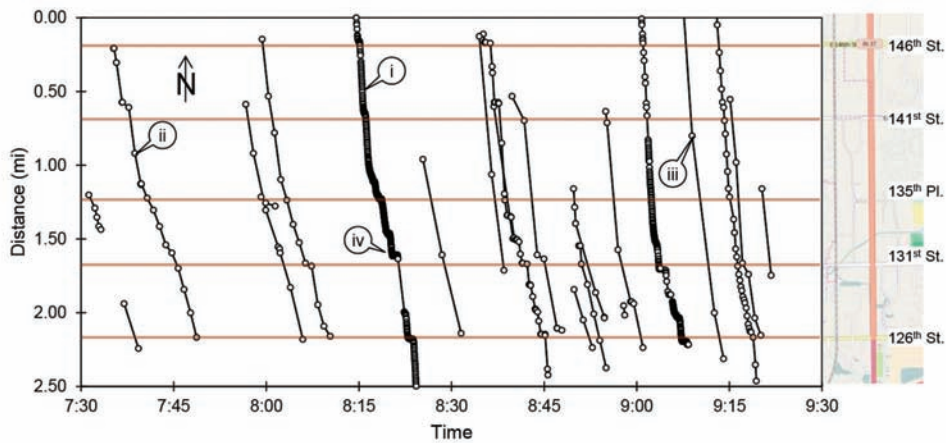
The signal status times have not been reconciled with the data, but it is possible to see bends in the trajectory lines that likely correspond to signal delay, such as in Figure 7.1a, callout *iv*. During the AM peak, southbound travel is more congested heading into Indianapolis, which is clearly reflected in the slightly more horizontal slopes of the southbound trajectories in Figure 7.1a compared with the slightly more vertical slopes of the northbound trajectories in Figure 7.1b.

The southbound direction has 21 trajectories in total across the 2-hour period. Analysis of volume measurements finds that approximately 2,600 southbound vehicles are likely to travel along the corridor during this 2-hour period, which corresponds to a market penetration rate of roughly 0.8%.

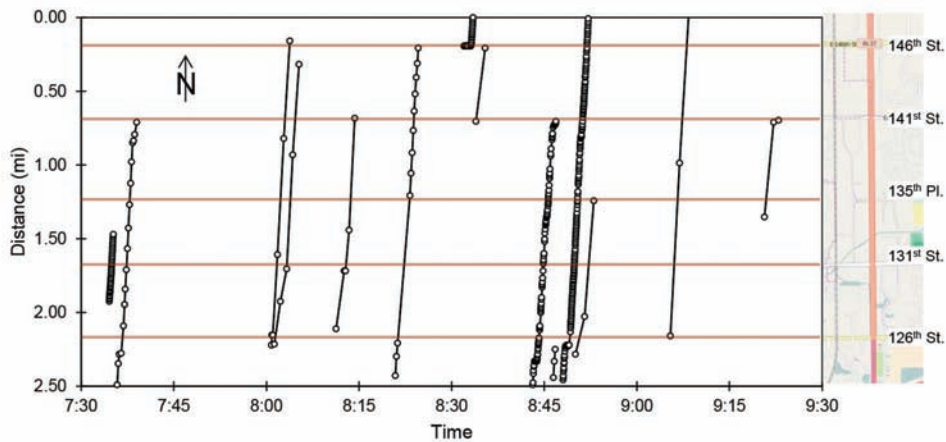
This type of data will greatly enhance signal performance measures by enabling a variety of additional performance measures, once the level of market penetration (i.e., the percentage of vehicles providing data) achieves an appropriate level. Some recent studies [100] have found that about 15% market penetration is needed for some performance measure applications over a 15-minute analysis period. A more recent proof-of-concept study [89] showed that for applications where longer analysis periods are feasible, a market penetration rate as low as 1%, and perhaps even lower, might have useful applications.

Figure 7.2 shows results from this study. Each of the two plots shows a distribution of simulated system performance under offsets that resulted from flow profiles constructed from *subsamples* of the known true vehicle arrival distributions. For example, for market penetration $p = 10\%$, only 1 out of 10 of the arrival detections would be detectable. To “mock up” some sample data, real-world vehicle detections were used, and from these, a subsample was randomly selected using p as the probability of selection. Then, 100 random-sample iterations were executed, to develop a distribution of likely solution performance values for different levels of p , for different analysis periods T . The existing performance is indicated by the blue line, and the green and red lines represent the best and worst possible solutions. The data represent the results of offset optimization for a nine-intersection corridor (SR 37 North).

The spreading of the distribution represents the divergence of the resulting offsets from the solution when 100% of the vehicles are known (i.e., all detection



(a) Southbound vehicles.



(b) Northbound vehicles.

Figure 7.1 Example raw real-world probe vehicle position data from a section of SR 37 North. Data are shown for a portion of Tuesday, 4/14/2015.

times are known). The smaller the market penetration, the broader the distribution and the worse the performance of the median solution.

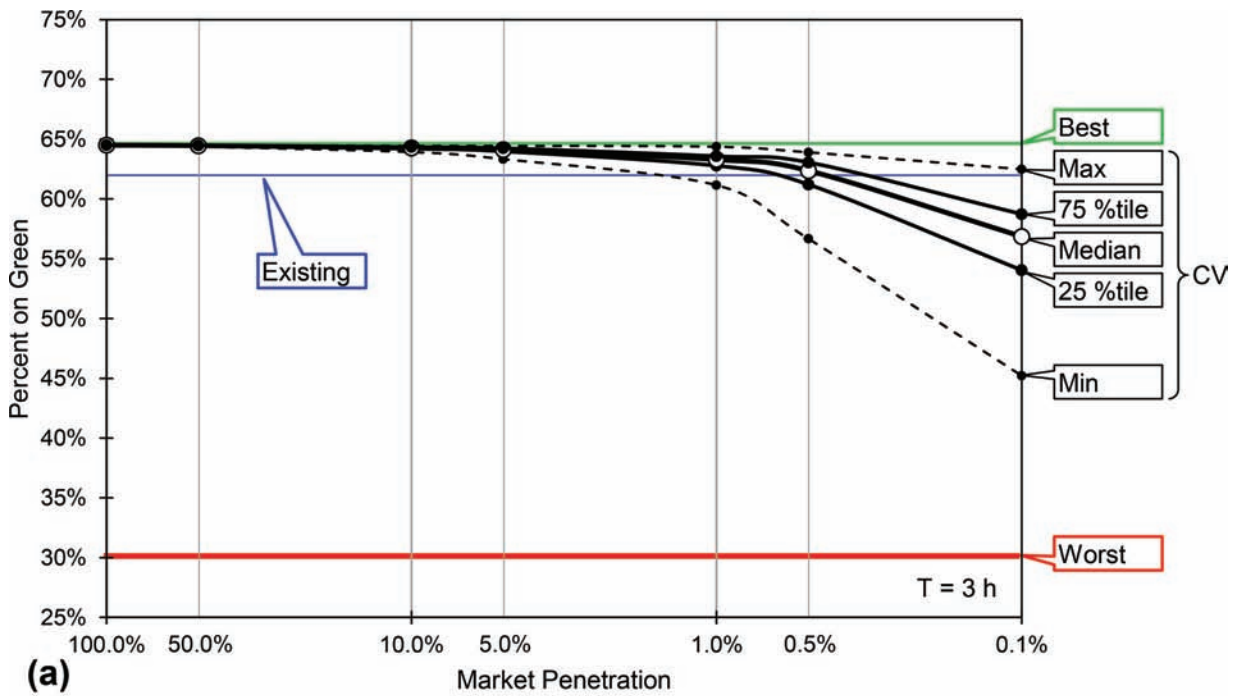
Figure 7.2a shows results for $T = 3$ hours, representing a typical “offline” application such as optimization of offsets. The chart shows that over 75% of the solutions are better than the existing offsets, for market penetrations as low as 1%. Meanwhile, Figure 7.2b shows results for $T = 15$ min, representing an “online” application such as adaptive offset adjustments. In this case, at least 5% of the traffic is needed for 75% of the solutions to be better than the existing offsets.

Should similar results be obtained from actual data, the implication is that it may be possible to achieve offset optimization, and similar types of data-driven analysis and optimization, without any detection systems at all, provided that connected vehicle or analogous data can represent at least 1–5% of the traffic. Future research will pursue this possibility further. One challenge that will need to be overcome is synchronization of clocks between multiple data collection systems.

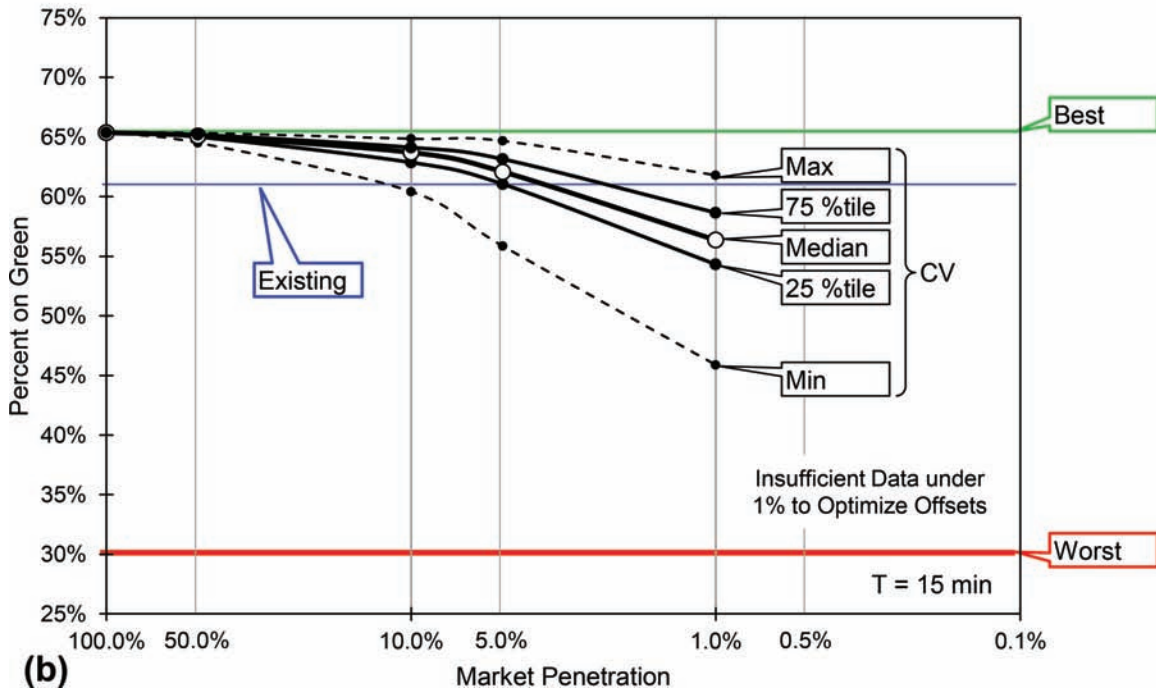
7.4 Report Conclusion

Signal performance measures provide the opportunity for traffic agencies to develop more detailed intelligence about what goes on in their deployed systems than has ever been possible in the past, with the possible exception of any agencies that have had total video coverage of their entire network with 24-hour monitoring. This document has included many different examples, including over 100 graphical exhibits, presenting several very common use cases.

There will likely be continued development of the performance measure methodology, especially concerning the integration of new and emerging datasets such as connected vehicle data, and for modes of traffic additional to those that have been the focus of the research so far. However, with a strong portfolio of performance measures in hand, the focus of research is now pivoting toward implementation, and use of the data to improve operation as well as to support maintenance, planning, design, and agency management. Initially, we would



(a)



(b)

Figure 7.2 Sensitivity of optimization outcome to market penetration p for (a) $T = 3$ h and (b) $T = 15$ min. Lines marked “CV” show the distribution of results for 100 iterations at each trial value of p [89].

expect that the added intelligence provided by performance measures will enable traffic engineers to do traffic engineering more efficiently, by providing relatively rapid feedback to changes made to control policies than is currently available.

Integration of performance data with timing plan design tools would be a valuable extension of the data. The research presented in this report included methods for optimizing offsets and for evaluating changes to splits and cycle length. Future research should connect

the data to the complete evaluation and optimization of timing plans and time-of-day/day-of-week schedules.

Looking further, the data could also be applied to parameter selection for more advanced control systems. This would enable the development of expert systems that go beyond optimizing signal timing in a reactive manner, to making proactive, strategic decisions with network impacts. In the long run, this would eventually support congestion management applications related to dynamic routing and traffic assignment. A detailed record of signal activities and the resulting performance of traffic will be essential to develop intelligent systems for arterial management.

REFERENCES

1. Transportation Pooled Fund Program. (2011). Study detail view. Retrieved from <http://www.pooledfund.org/Details/Study/487>
2. Day, C. M., Bullock, D. M., Li, H., Remias, S. M., Hainen, A. M., Freije, R. S., ... Brennan, T. M. (2014). *Performance measures for traffic signal systems: An outcome-oriented approach*. West Lafayette, IN: Purdue University. <http://dx.doi.org/10.5703/1288284315333>
3. Day, C.M. & D.M. Bullock. (2015). Integrating outcome oriented performance measures into traffic signal operations business processes. In *Proceedings of IEEE 18th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 131–136). New York, NY: IEEE.
4. U.S. Government Accountability Office. (1994). *Transportation infrastructure, benefits of traffic control systems are not being fully realized* (Report RCED-94-105). Washington, DC: U.S. Government Accountability Office.
5. Institute of Transportation Engineers. (1996). *A Toolbox for alleviating traffic congestion and enhancing mobility*. Washington, DC: Federal Highway Administration. Retrieved from <http://www.fhwa.dot.gov/publications/research/operations/its/98436/toolbox.pdf>
6. National Transportation Operations Coalition. (2012). *2012 National traffic signal report card: Technical report*. Retrieved from <http://library.ite.org/pub/e265477a-2354-d714-5147-870dfac0e294>
7. Day, C. M., & Bullock, D. M. (2015). *Performance measures of interrupted-flow roadways using re-identification and signal controller data* (FHWA SBIR DTFH61-14-C-00035). West Lafayette, IN: Traffax, Inc., & Purdue University.
8. Denney, R. W. (2009). *Improving traffic signal management and operations: A basic service model* (Report FHWA-HOP-09-055). Washington, DC: Federal Highway Administration.
9. Balke, K., Charara, H., & Parker, R. (2005). *Development of a traffic signal performance measurement system* (Report FHWA/TX-05/0-4422-2). College Station, TX: Texas Transportation Institute.
10. Smaglik, E. J., Sharma, A., Bullock, D. M., Sturdevant, J. R., & Duncan, G. (2007). Event-based data collection for generating actuated controller performance measures. *Transportation Research Record*, 2035, 97–106. <http://dx.doi.org/10.3141/2035-11>
11. Sharma, A., Bullock, D. M., & Bonneson, J. A. (2007). Input–output and hybrid techniques for real-time prediction of delay and maximum queue length at signalized intersections. *Transportation Research Record*, 2035, 69–80. <http://dx.doi.org/10.3141/2035-08>
12. Liu, H. X., & Ma, W. T. (2009). A virtual vehicle probe model for time-dependent travel time estimation on signalized arterials. *Transportation Research Part C: Emerging Technologies*, 17(1), 11–26. <http://dx.doi.org/10.1016/j.trc.2008.05.002>
13. Liu, H. X., & Ma, W. T. (2014). Using high-resolution event-based data for traffic modeling and control: An overview. *Transportation Research Part C: Emerging Technologies*, 42, 28–43. <http://dx.doi.org/10.1016/j.trc.2014.02.001>
14. Barkley, T., Hranac, R., Fuentes, K., & Law, P. (2014). Heuristic approach for estimating arterial signal phases and progression quality from vehicle arrival data. *Transportation Research Record*, 2259, 48–58. <http://dx.doi.org/10.3141/2259-05>
15. Day, C. M., Smaglik, E. J., Bullock, D. M., & Sturdevant, J. R. (2008). *Real-time arterial traffic signal performance measures* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2008/09). West Lafayette, IN: Purdue University. <http://dx.doi.org/10.5703/1288284313439>
16. Day, C. M., & Bullock, D. M. (2010). *Arterial performance measures, Volume 1: Performance based management of arterial traffic signal systems* (Final Report, NCHRP 3-79A). Washington, DC: National Cooperative Highway Research Program, Transportation Research Board of the National Academies.
17. Day, C. M., Brennan, T. M., Ernst, J. M., Overman, T. W., Sturdevant, J. R., Hainen, A. M., Bullock, D. M. (2011). *Procurement procedures and specifications for performance measure capable traffic infrastructure data collection systems* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2011/18). West Lafayette, IN: Purdue University. <http://dx.doi.org/10.5703/1288284314642>
18. Day, C. M., Sturdevant, J. R., & Bullock, D.M. (2010). Outcome oriented performance measures for management of signalized arterial capacity. *Transportation Research Record*, 2192, 24–36. <http://dx.doi.org/10.3141/2192-03>
19. Day, C. M., Haseman, R., Premachandra, H., Brennan, T. M., Wasson, J. S., Sturdevant, J. R., & Bullock, D. M. (2010). Evaluation of arterial signal coordination: Methodologies for visualizing high-resolution event data and measuring travel time. *Transportation Research Record*, 2192, 37–49. <http://dx.doi.org/10.3141/2192-04>
20. Li, H., Hainen, A. M., Day, C. M., Grimmer, G., Sturdevant, J. R., & Bullock, D. M. (2013). Longitudinal performance measures for assessing agencywide signal management objectives. *Transportation Research Record*, 2355, 20–30. <http://dx.doi.org/10.3141/2355-03>
21. Freije, R., Hainen, A. M., Stevens, A., Li, H., Smith, W. B., Day, C. M., Sturdevant, J. R., & Bullock, D. M. (2014). Graphical performance measures for practitioners to triage split failure trouble calls. *Transportation Research Record*, 2439, 27–40. <http://dx.doi.org/10.3141/2439-03>
22. Lavrenz, S. M., Day, C., Grossman, J., Freije, R., & Bullock, D. M. (2016, January). *Use of high resolution signal controller data to identify red light running* (Paper No. 16-0209). Presented at Transportation Research Board Annual Meeting, Washington, DC.
23. Sunkari, S. R., Charara, H. A., & Songchitruska, P. (2012). Portable toolbox for monitoring and evaluating signal operations. *Transportation Research Record*, 2311, 142–151. <http://dx.doi.org/10.3141/2311-14>

24. Sturdevant, J. R., Overman, T., Raamot, E., Deer, R., Miller, D., Bullock, D. M., ... Remias, S. M. (2012). *Indiana traffic signal high resolution data logger enumerations*. West Lafayette, IN: Purdue University. <http://dx.doi.org/10.4231/K4RN35SH>
25. Live Traffic Data, LLC. (2016). Website. <http://livetrafficdata.com>
26. TrafInfo Communications, Inc. (n.d.). Website. <http://www.trafinfo.com>
27. Schrank, D., Eisele, B., Lomax, T., & Bak, J. (2015). *2015 Urban mobility scorecard*. College Station, TX: Texas Transportation Institute.
28. Maryland Department of Transportation. (2012). *2012 Maryland state highway mobility report*. Baltimore, MD: Maryland Department of Transportation, 2012. Retrieved from http://roads.maryland.gov/ooots/2012_maryland_state_highway_mobility_report.pdf
29. Washington State Department of Transportation. (2012). *The 2012 congestion report*. Olympia, WA: Washington State Department of Transportation. Retrieved from <http://wsdot.wa.gov/publications/fulltext/graynotebook/CR12.pdf>
30. Day, C. M., Remias, S. M., Li, H., Mekker, M., McNamara, M., Cox, E., ... Bullock, D. (2014). *2013–2014 Indiana mobility report: Full version* (Joint Transportation Research Program Indiana Mobility Reports). West Lafayette, IN: Purdue University. <http://dx.doi.org/10.5703/1288284315508>
31. Remias, S. M., Brennan, T. M., Day, C. M., Summers, H., Horton, D. K., Cox, E. D., & Bullock, D. M. (2014). Spatially referenced probe data performance measures for infrastructure investment decision makers. *Transportation Research Record, 2420*, 33–44. <http://dx.doi.org/10.3141/2420-04>
32. Day, C. M., Remias, S. M., Li, H., Mekker, M. M., McNamara, M. L., Cox, E. D., & Bullock, D. M. (2015). Performance ranking of arterial corridors using travel time and travel time reliability metrics. *Transportation Research Record, 2487*, 44–54. <http://dx.doi.org/10.3141/2487-04>
33. Wünsch, G., Bölling, F., von Dubschütz, A., & Mieth, P. (2015). Bavarian Road Administration's use of probe data for large-scale traffic signal evaluation support. *Transportation Research Record, 2487*, 8–95. <http://dx.doi.org/10.3141/2487-08>
34. Urbanik, T., Tanaka, A., Lozner, B., Lindstrom, E., Lee, K., Quayle, S., ... Bullock, D. (2015). *Signal timing manual* (2nd ed.; NCHRP Report 812). Washington, DC: Transportation Research Board of the National Academies. Retrieved from http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_812.pdf
35. Parasuraman, A., Ziethaml, V. A., & Berry, L. L. (1988). A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*, 12–40.
36. Yin, Y. (2008). Robust optimal traffic signal timing. *Transportation Research Part B: Methodological, 42*(10), 911–924.
37. Stevanovic, A., Kergaye, C., & Stevanovic, J. (2011). Evaluating robustness of signal timings for varying traffic flows. *Transportation Research Record, 2259*, 141–150. <http://dx.doi.org/10.3141/2259-13>
38. Gartner, N. H., Deshpande, R. M., & Stamatiadis, C. (2012). Pareto efficiency in signal control: Application to intersections and urban streets (Paper No. 12-0986). In *TRB 91st Annual Meeting compendium of papers DVD*. Washington, DC: Transportation Research Board.
39. International Telecommunication Union. (1994, July). Telecommunication standardization sector of ITU. Information technology—Open systems interconnection—Basic reference model: The basic model. Recommendation X.200. Accessed at <http://www.itu.int/rec/T-REC-X.200-199407-1/en>
40. Bullock, D. M., Day, C. M., Brennan, T. M., Sturdevant, J. R., & Wasson, J. S. (2011). Architecture for active management of geographically distributed signal systems. *ITE Journal, 81*(5), 20–24.
41. Project Management Institute. (2008). *A guide to the project management body of knowledge* (4th ed.). Newtown Square, PA: Project Management Institute.
42. Wiesenfeld, M. T. (2008). *Analysis of design and implementation standards for coordinated traffic signal systems* (MS Thesis). Purdue University, West Lafayette, IN. Retrieved from <http://docs.lib.purdue.edu/dissertations/AAI1480127>
43. Institute of Transportation Engineers. (2008). *Using existing loops at signalized intersections for traffic counts*. Washington, DC: Institute of Transportation Engineers.
44. Day, C. M., & Bullock, D. M. (2012). Characterization of platoon dispersion model with high-resolution signal event data. *Transportation Research Record, 2311*, 16–28. <http://dx.doi.org/10.3141/2311-02>
45. Institute of Transportation Engineers. (2014). *Automated traffic signal performance measures: Case studies* [ITE webinar series]. Washington, DC: Institute of Transportation Engineers.
46. AASHTO. (n.d.). AASHTO innovation initiative: Automated traffic signal performance measures. Accessed at <http://ait.transportation.org/Pages/AutomatedTrafficSignalPerformanceMeasures.aspx>
47. Day, C. M., Brennan, T. M., Premachandra, H., Sturdevant, J. R., & Bullock, D. M. (2011). Analysis of peer data on intersections for decisions about coordination of arterial traffic signal. *Transportation Research Record, 2259*, 23–36. <http://dx.doi.org/10.3141/2259-03>
48. Li, H., Day, C. M., Sturdevant, J. R., & Bullock, D. M. (2016, January). *Scaling detailed high-resolution data split performance measures to statewide system level management* (Paper No. 16-4149). Presented at Transportation Research Board Annual Meeting, Washington, DC.
49. Utah Department of Transportation. (n.d.b). Signal performance metrics: Executive reports. Accessed at <http://udottraffic.utah.gov/signalperformanceexecutive-reports>
50. Middleton, D., Longmire, R., Bullock, D. M., & Sturdevant, J. R. (2009). Proposed concept for specifying vehicle detection performance. *Transportation Research Record, 2128*, 161–172. <http://dx.doi.org/10.3141/2128-17>
51. Day, C. M., Premachandra, H., Brennan, T. M., Sturdevant, J. R., & Bullock, D. M. (2014). Operational evaluation of wireless magnetometer vehicle detectors at signalized intersection. *Transportation Research Record, 2192*, 11–23. <http://dx.doi.org/10.3141/2192-02>
52. Lavrenz, S. M., Day, C. M., Smith, W. B., Sturdevant, J. R., & Bullock, D. M. (2016, January). *Assessing longitudinal arterial performance traffic signal retiming outcomes* (Paper No. 16-0113). Presented at Transportation Research Board Annual Meeting, Washington, DC.
53. Indiana Department of Transportation. (2016). *Procedure for evaluating vehicle detection performance* (ITM No. 934-15P). Indianapolis, IN: Indiana Department of Transportation. 2016. Retrieved from http://www.in.gov/indot/div/mt/itm/pubs/934_testing.pdf

54. Smaglik, E. J., Bullock, D. M., & Urbanik, T. (2005). Evaluation of lane-by-lane vehicle detection for actuated controllers serving multilane approaches. *Transportation Research Record*, 1925, 123–133. <http://dx.doi.org/10.3141/1925-13>
55. Smith, W. B. (2014). *Signalized corridor assessment* (MSCE Thesis). Purdue University, West Lafayette, IN. Retrieved from <http://docs.lib.purdue.edu/dissertations/AAI1565321>
56. Liu, H. X., Wu, X., Ma, W., & Hu, H. (2009). Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C: Emerging Technologies*, 17(4), 412–427. <http://dx.doi.org/10.1016/j.trc.2009.02.003>
57. Wu, X., Liu, H. X., & Gettman, D. (2010). Identification of oversaturated intersections using high-resolution traffic signal data. *Transportation Research Part C: Emerging Technologies*, 18(4), 626–638. <http://dx.doi.org/10.1016/j.trc.2010.01.003>
58. Wasson, J. S., Sturdevant, J. R., & Bullock, D. M. (2008). Real-time travel time estimates using media access control address matching. *ITE Journal*, 78(6), 20–23.
59. Rhea, S., Wells, C., Eaton, P., Geels, D., Zhao, B., Weatherspoon, H., & Kubiawicz, J. (2001). Maintenance-free global data storage. *IEEE Internet Computing*, 5(5), 40–49. <http://dx.doi.org/10.1109/4236.957894>
60. Krasner, G. E., & Pope, S. T. (1988). A cookbook for using the model-view-controller user interface paradigm in Smalltalk-80. *Journal of Object-Oriented Programming*, 1(3), 26–49.
61. Indiana Department of Transportation. Welcome to the LaPorte District. Accessed October 1, 2015, at <https://secure.in.gov/indot/2705.htm>
62. Lavrenz, S. (2015). *High-resolution data-based methods for enhanced asset preservation, mobility, and safety at signalized intersections* (PhD Thesis). Purdue University, West Lafayette, IN.
63. Lavrenz, S., Sturdevant, J., & Bullock, D. (2015). *Strategic methods for modernizing traffic signal maintenance management and quantifying the impact of maintenance activities*. Manuscript submitted for publication.
64. Turner, S. (2007). *Quality control procedures for archived operations traffic data: Synthesis of practice and recommendations*. College Station, TX: Texas Transportation Institute, Federal Highway Administration.
65. Parapar, S. (1973). Computerized reporting for traffic signal maintenance. *Traffic Engineering*, 43(10), 47–51.
66. Blase, J. (1979). Computer aids to large-scale traffic signal maintenance. *Traffic Engineering & Control*, 20, 341–347.
67. Jacobson, L., Nihan, N., & Bender, J. (1990). Detecting erroneous loop detector data in a freeway traffic management system. *Transportation Research Record*, 1287, 151–166.
68. Payne, H., & Thompson, S. (1997). Malfunction detection and data repair for induction-loop sensors using I-880 data base. *Transportation Research Record*, 1570, 191–201. <http://dx.doi.org/10.3141/1570-22>
69. Chen, C., Kwon, J., Rice, J., Skabardonis, A., & Varaiya, P. (2003). Detecting errors and imputing missing data for single-loop surveillance systems. *Transportation Research Record*, 1855, 160–167. <http://dx.doi.org/10.3141/1855-20>
70. Achillides, C., & Bullock, D. (2004). *Performance metrics for freeway sensors* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2004/37). West Lafayette, IN: Purdue University. <http://dx.doi.org/10.5703/1288284313287>
71. Randeniya, D., & Kim, H. K. (2013). Estimation of ITS sensor operational states by analyzing measurements with errors using a hidden Markov model. *KSCCE Journal of Civil Engineering*, 17(7), 1740–1748. <http://dx.doi.org/10.1007/s12205-013-0284-2>
72. Audelo, M., Chou, C., Chen, T., & Nichols, A. (2014). Empirical analysis of controller event data to select vehicle detector fault triggers. *Transportation Research Record*, 2438, 12–22. <http://dx.doi.org/10.3141/2438-02>
73. Pennsylvania Department of Transportation. (2010). *Guidelines for the maintenance and operation of traffic signals*. Harrisburg, PA: Pennsylvania Department of Transportation. Retrieved from <http://www.dot.state.pa.us/public/PubsForms/Publications/PUB%20191.pdf>
74. Brennan, T. M., Griggs, B. D., Grimmer, G., Hainen, A. M., Day, C. M., Sturdevant, J. R., & Bullock, D. M. (2012). Defining design space for parameters of traffic signal timing: Empirical approach. *Transportation Research Record*, 2311, 85–98. <http://dx.doi.org/10.3141/2311-08>
75. Smith, W. B. (2014). *Signalized corridor assessment* (MSCE Thesis). Purdue University, West Lafayette, IN. Retrieved from <http://docs.lib.purdue.edu/dissertations/AAI1565321>
76. Utah Department of Transportation. (n.d.a). Signal performance metrics. Accessed at <http://udottraffic.utah.gov/signalperformancemetrics>
77. Smaglik, E. J., Bullock, D. M., Gettman, D., Day, C. M., & Premachandra, H. (2011). Comparison of alternative real-time performance measures for measuring signal phase utilization and identifying oversaturation. *Transportation Research Record*, 2259, 123–131. <http://dx.doi.org/10.3141/2259-11>
78. Sharma, A., Bullock, D. M., & Bonneson, J. A. (2007). Input-output and hybrid techniques for real-time prediction of delay and maximum queue length at signalized intersections. *Transportation Research Record*, 2035, 69–80. <http://dx.doi.org/10.3141/2035-0>
79. Day, C. M., Premachandra, H., & Bullock, D. M. (2011). Rate of pedestrian signal phase actuation as a proxy measurement of pedestrian demand (Paper No. 11-0220). In *TRB 90th Annual Meeting compendium of papers DVD*. Washington, DC: Transportation Research Board.
80. Hubbard, S. M. L., Bullock, D. M., & Day, C. M. (2008). Integration of real-time pedestrian performance measures into existing infrastructure of traffic signal system. *Transportation Research Record*, 2080, 34–47. <http://dx.doi.org/10.3141/2080-05>
81. Li, H., Day, C. M., Hainen, A. M., Stevens, A. L., Lavrenz, S., Smith, W. B., . . . & Bullock, D. M. (2014). *Field cycle length sweep to evaluate resonant cycle sensitivity*. JTRP Other Publications and Reports. <http://dx.doi.org/10.5703/1288284315509>
82. Brennan, T. M. (2009). *Quantitative assessment methods for railroad preempted intersections* (PhD Thesis). Purdue University, West Lafayette, IN. Retrieved from <http://docs.lib.purdue.edu/dissertations/AAI3402292>
83. Brennan, T. M., Day, C. M., Bullock, D. M., & Sturdevant, J. R. (2009). Performance measures for railroad preempted intersections. *Transportation Research Record*, 2128, 20–34. <http://dx.doi.org/10.3141/2128-03>
84. Brennan, T. M., Day, C. M., Sturdevant, J. R., Raamot, E., & Bullock, D. M. (2010). Track clearance perfor-

- mance measures for railroad preempted intersections. *Transportation Research Record*, 2192, 64–76. <http://dx.doi.org/10.3141/2192-06>
85. McNamara, M. L., Li, H., Remias, S. M., Horton, D. K., Cox, E. D., & Bullock, D. M. (2016, January). *Real-time probe data dashboards for Interstate performance monitoring during winter weather and incidents* (Paper No. 16-0622). Presented at Transportation Research Board Annual Meeting, Washington, DC.
 86. Next Generation SIMulation Community. Accessed at <http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm>
 87. Young, S. (2014, June). *Use of probe and Bluetooth data for arterial performance measures in the I-95 corridor coalition*. Presented at the NATMEC Workshop on Arterial Data and Performance Monitoring, Chicago, IL.
 88. Robertson, D. I. (1969). *Transyt: A traffic network study tool* (Report No. LR 253). Crowthorne, Berkshire, England: Road Research Laboratory.
 89. Day, C. M., & Bullock, D. M. (2016, January). *Opportunities for detector-free signal optimization with limited connected vehicle market penetration: A proof-of-concept study* (Paper No. 16-0112). Presented at Transportation Research Board Annual Meeting, Washington, DC.
 90. Transportation Research Board of the National Academies. (2010). *Highway capacity manual*. Washington, DC: Transportation Research Board of the National Academies.
 91. Smaglik, E. J., Bullock, D. M., & Sharma, A. (2007). Pilot study on real-time calculation of arrival type for assessment of arterial performance. *Journal of Transportation Engineering*, 133, 415–422.
 92. Day, C. M., Lavrenz, S. M., Stevens, A. L., Miller, R. E., & Bullock, D. M. (2016, January). *Extending link pivot offset optimization to arterials with single controller diverging diamond interchange* (Paper No. 16-0111). Presented at Transportation Research Board Annual Meeting, Washington, DC.
 93. Hainen, A. M., Li, H., Stevens, A. L., Day, C. M., Sturdevant, J. R., & Bullock, D. M. (2015). *Sequence optimization at signalized diamond interchanges using high-resolution event-based data*. *Transportation Research Record*, 2487, 5–30. <http://dx.doi.org/10.3141/2487-02>
 94. Hainen, A., Stevens, A., Li, H., & Bullock, D. (2014). *Three-phase operations at a diverging diamond interchange using an unmanned aerial vehicle (UAV) camera*. Purdue University Research Repository. <http://dx.doi.org/10.4231/R7C24TC4>
 95. Day, C. M., Ernst, J. M., Brennan, T. M., Chou, C.-S., Hainen, A. M., Remias, S. M., . . . & Bullock, D. M. (2012). Performance measures for adaptive signal control: Case study of system-in-the-loop simulation. *Transportation Research Record*, 2311, 1–15. <http://dx.doi.org/10.3141/2311-01>
 96. Day, C. M., Brennan, T. M., Hainen, A. M., Remias, S. M., Premachandra, H., Sturdevant, J. R., . . . & Bullock, D. M. (2011). Reliability, flexibility, and environmental impact of alternative objective functions for arterial offset optimization. *Transportation Research Record*, 2259, 8–22. <http://dx.doi.org/10.3141/2259-02>
 97. Day, C. M., & Bullock, D. M. (2011). Computational efficiency of alternative algorithms for arterial offset optimization. *Transportation Research Record*, 2259, 37–47. <http://dx.doi.org/10.3141/2259-04>
 98. Hillier, J. A. (1965). Appendix to Glasgow's experiment in area traffic control. *Traffic Engineering & Control*, 7, 569–571.
 99. McNamara, M., Li, H., Remias, S., Richardson, L., Cox, E., Horton, D., & Bullock, D. M. (2015). Using real-time probe vehicle data to manage unplanned detour routes. *ITE Journal of Transportation*, December, 32–37. Retrieved from <http://library.ite.org/pub/b772541b-04ab-0f6f-a863-34c8217c947e>
 100. Argote-Cabañero, J., Christofa, E., & Skabardonis, A. (2015). Connected vehicle penetration rate for estimation of arterial measures of effectiveness. *Transportation Research Part C: Emerging Technologies*, 60, 298–312. <http://dx.doi.org/10.1016/j.trc.2015.08.013>

Project Partners

This work was supported in part by Pooled Fund Study TPF-5(258), led by the Indiana Department of Transportation (INDOT) and supported by the state transportation agencies of California, Georgia, Minnesota, Mississippi, New Hampshire, Pennsylvania, Texas, Utah, and Wisconsin, the Indiana Local Technical Assistance Program (LTAP), the Chicago Department of Transportation, and the



Signal Performance Measures Workshop, Salt Lake City, Utah, January 26–27, 2016.

Federal Highway Administration Office of Operations and Resource Center. Some material in this monograph was compiled from previous studies that were made possible under National Cooperative Highway Research Program project 3-79a, INDOT State Planning and Research (SPR) projects, Indiana LTAP projects, and USDOT through Small Business Innovation Research (SBIR) projects with Traffax, Inc., and through a joint research project with Marshall University. We would like to thank Rick Schuman and colleagues at Inrix, Inc., for provision of sample vehicle trajectory data in Chapter 7. We are grateful to these sponsors and research partners for their support over the years.

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein, and do not necessarily reflect the official views or policies of the sponsoring organizations. These contents do not constitute a standard, specification, or regulation.

Publication

This report was published in collaboration with the Joint Transportation Research Program and Purdue University. The full content of this technical report is available for download at <http://dx.doi.org/10.5703/1288284315333>. In addition, a print-on-demand bound version of this report can be purchased at that location.

Open Access and Collaboration with Purdue University

The Indiana legislature established the Joint Highway Research Project in 1937. In 1997 this collaborative venture between the Indiana Department of Transportation and Purdue University was renamed as the Joint Transportation Research Program (JTRP) to reflect state and national efforts to integrate the management and operation of various transportation modes. Since 1937 the JTRP program has published more than 1,500 technical reports. In 2010, the JTRP partnered with the Purdue University Libraries to incorporate these technical reports in the University's open access digital repository and to develop production processes for rapidly disseminating new research reports via this repository. Affiliated publications have also recently been added to the collection. As of 2016 the JTRP collection has over 1.3 million downloads, with some particularly popular reports having over 20,000 downloads.