Applying Archived Operations Data in Transportation Planning

A PRIMER

U.S. Department of Transportation
Federal Highway Administration

Notice

Quality Assurance Statement

Cover image: Thinkstock/zhudfeng

Technical Report Documentation Page

<table>
<tr><td>1. Report No.
FHWA-HOP-16-082</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Applying Archived Operations Data in Transportation Planning – A Primer</td><td colspan="2">5. Report Date
December 2016</td></tr>
<tr><td colspan="2">6. Performing Organizations Code</td></tr>
<tr><td colspan="2">7. Authors
Jocelyn Bauer, Richard Margiotta, Michael Pack (alphabetical order)</td><td colspan="2">8. Performing Organization Report No.</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
Leidos, Reston, Virginia
Cambridge Systematics, Inc., Knoxville, Tennessee
Michael Pack, LLC, College Park, Maryland</td><td colspan="2">10. Work Unit No. (TRAIS)</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
Contract No.

DTFH61-12-D-00020</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
U.S. Department of Transportation
Federal Highway Administration
1200 New Jersey Avenue, SE
Washington, DC 20590</td><td colspan="2">13. Type of Report and Period Covered
Primer, July 2014–May 2016</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code
HOP</td></tr>
<tr><td colspan="4">15. Supplementary Notes
Wayne Berman, Government Task Manager, Federal Highway Administration</td></tr>
<tr><td colspan="4">16. Abstract
The purpose of this primer is to assist transportation planners in effectively using archived operations data for developing, analyzing, and evaluating transportation plans and programs. This primer raises planners' awareness of the opportunities afforded through archived operations data and provides guidance on how to take advantage of that data to expand and improve planning practices. It also identifies innovative applications for archived operations data in planning. This primer aims to help planners and their operations data partners overcome the barriers to obtaining and using data, regardless of whether they are just getting started with using the data or have substantive experience.</td></tr>
<tr><td colspan="2">17. Key Words
Archived operations data, planning, transportation, travel time, volume, speed, performance measures, analytic procedures.</td><td colspan="2">18. Distribution Statement
No restrictions.</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page) Unclassified</td><td>21. No of Pages
124</td><td>22. Price
N/A</td></tr>
</table>

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

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS				
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
LENGTH				
In.	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
AREA				
in.^{2}	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
VOLUME				
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
NOTE: volumes greater than 1,000 L shall be shown in m^3				
MASS				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
TEMPERATURE (exact degrees)				
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
ILLUMINATION				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
FORCE and PRESSURE or STRESS				
lbf	poundforce	4.45	newtons	N
lbf/in.^{2}	poundforce per square inch	6.89	kilopascals	kPa

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003).

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS				
SYMBOL	**WHEN YOU KNOW**	**MULTIPLY BY**	**TO FIND**	**SYMBOL**
LENGTH				
mm	millimeters	0.039	inches	in.
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
AREA				
mm^2	square millimeters	0.0016	square inches	$in.^2$
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
VOLUME				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
MASS				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
TEMPERATURE (exact degrees)				
°C	Celsius	1.8C+32	Fahrenheit	°F
ILLUMINATION				
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-lamberts	fl
FORCE and PRESSURE or STRESS				
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003).

Table of Contents

List of Figures

List of Figures (continued)

List of Figures (continued)

List of Tables

List of Acronyms

AADT	annual average daily traffic
ABS	antilock braking system
ALPR	automatic license plate recognition
ATDM	active transportation and demand management
ATMS	advanced transportation management system
AVL	automated vehicle location
ADOT	Arizona Department of Transportation
CAD	computer-aided dispatch
CCTV	closed circuit television
CDTC	Capital District Transportation Committee
CHART	Coordinated Highways Action Response Team
CMP	congestion management process
DMS	dynamic message sign
DOT	department of transportation
DVRPC	Delaware Valley Regional Planning Commission
FAST	Fixing America's Surface Transportation
FDOT	Florida Department of Transportation
FFS	free flow speed
FHWA	Federal Highway Administration
FMS	freeway management system
GDOT	Georgia Department of Transportation
GPS	Global Positioning System
HAR	highway advisory radio
HCM	Highway Capacity Manual
HPMS	Highway Performance Management System
HOV	high-occupancy vehicle
I-TMS	Internet-Traffic Monitoring System
ICM	integrated corridor management
iPEMS	Iteris Performance Measurement System
IT	information technology
ITS	intelligent transportation systems
LRTP	long-range transportation plan
MAG	Maricopa Association of Governments
MATOC	Metropolitan Area Transportation Operations Coordination
MIST	Management Information System for Transportation

MPH	miles per hour
MPO	metropolitan planning organization
MSTM	Maryland Statewide Transportation Model
MWCOG	Metropolitan Washington Council of Governments
NHS	National Highway System
NJDOT	New Jersey Department of Transportation
NPMRDS	National Performance Measures Research Data Set
O-D	origin-destination
ODOT	Oregon Department of Transportation
PCPHPL	passenger cars per hour per lane
PeMS	Performance Measurement System
PHV	peak-hour volumes
POMT	Port of Miami Tunnel
PORTAL	Portland Oregon Regional Transportation Archive Listing
PSU	Portland State University
PTI	planning time index
QC	quality control
RFP	request for proposal
RITIS	Regional Integrated Transportation Information Systems
ROI	return on investment
RTMS	remote traffic microwave sensors
RWIS	road weather information systems
SHA	State Highway Administration
SHRP 2	Strategic Highway Research Program 2
TDF	travel demand forecasting
TMC	transportation management center
TTI	travel time index
UMD	University of Maryland
VMT	vehicle-miles of travel
VPP	Vehicle Probe Project
WSDOT	Washington State Department of Transportation

1. INTRODUCTION

Making the Case – Using Archived Operations Data for Transportation Planning

Effective transportation planning and investment decision making depends on timely, comprehensive, and accurate data. Traditionally, data for planning has come from manual collection techniques and approximations from model output. While these sources are still used, transportation planners at the State, metropolitan, and local level are beginning to leverage an increasingly prevalent kind of transportation data—archived operations data.

Archived operations data is information collected and stored to support the monitoring and management of the transportation system. Enormous amounts of data on the performance of the transportation system are generated daily. For example, there are over 4 billion probe-based road segment speeds generated in the United States each day. California alone produces nearly 30 million sensor measurements for speed and volume daily.[1] Archived operations data can include traffic, transit, bike, pedestrian, construction, and weather information that is usually collected in real-time by intelligent transportation system (ITS) infrastructure, such as in pavement inductive loop detectors, radar detectors, remote traffic microwave sensors (RTMS), Bluetooth, and E-ZPass or other unique identifier tag readers. It also includes incident or event information entered into electronic logs by transportation or public safety personnel. Although not exhaustive, Table 1 contains many of the common types of archived operations data that are collected.

> Transportation planners at the State, metropolitan, and local level are finding they are able to rely less on assumptions and modeled data, solve a broader range of problems, and make more effective decisions with archived operations data.

Historically, operations data was used almost exclusively "in the moment" to help manage the system in real-time, and much of the data generated from these systems was either not archived or was archived in ways that made it extremely difficult to access and work with. Most operations data archives were only used for auditing purposes and were only accessible by one or two individuals within the agency's operations or information technology (IT) group.

However, with shrinking budgets; a focus on performance-based planning; a greater commitment to data quality; and advances in sensors, collection techniques, and archiving technologies, agencies are realizing the value of these data collection efforts beyond day-to-day operations and are tapping into opportunities to make better use of existing operations data. This has led them to enhance operations data archiving in ways that make it more readily available to larger and more diverse user groups, such as planners. Transportation planners at the State, metropolitan, and local level are finding they are able to rely less on assumptions and modeled data, solve a broader range of problems, and make more effective decisions.

1 Pack, Michael, University of Maryland Center for Advanced Transportation Technology Laboratory, Personal Communication, 2015.

Table 1. Types of archived operations data.[2]

Archived Operations Data Type	Description
Traffic volume, speed, class, and occupancy from point and probe data sources	Information collected by agencies and third parties from roadway sensors that could include inductive loops, side-fired sensors (acoustic, microwave, etc.), radar, and video. Also includes data from probe-based systems—either agency-owned or third-party supplied.
Event, work zone, and incident information	Information entered by each agency into its own incident management system. Data typically includes incident location, type, severity, information about the vehicles involved and their status, to whom are notifications made and which responders are on-scene, lane closures, response plans or detours, and messages on dynamic message signs (DMS) or highway advisory radio (HAR).
Weather data	Weather alerts, temperature, precipitation types and rates, wind speeds, radar data, and other information from the National Weather Service, third parties, the media, etc. Also includes weather and pavement surface conditions that agencies gather from their roadway weather information systems.
Device operational status	Data on the operational status of roadway devices from each agency, including traffic detectors, DMS, traffic signals, HAR, roadway weather information systems, and closed circuit television (CCTV) cameras, where available.
Managed lane status	Data on when high-occupancy vehicle (HOV) restrictions are in effect, direction of reversible lanes by time of day, and price of high-occupancy toll lanes by time.
Surveillance video	Live CCTV feeds focused on roadways, assets, or pedestrians.
Transit alerts	Transit alerts, service disruptions, and other information transmitted by transit providers—both public and private.
Automated vehicle locations (AVLs)	The locations and status of freeway service patrols, transit vehicles, or other assets equipped with AVL hardware.
Signal status	Operational status of signals at intersections or ramp meters, such as operational, maintenance mode, flashing, or offline.
Signal timing plans	Signal timing plans, current or future timing schemes.
Computer-aided dispatch (CAD) information	Data from public safety computer-aided dispatch (CAD) systems, such as fire, emergency medical services, and law enforcement. Can include dispatch requests, incident types, severity, responder requests, or even lane status.

2 Pack, Michael and Ivanov, Nikola, National Cooperative Highway Research Program Synthesis 460, *Sharing Operations Data Among Agencies* (Washington, DC: 2014). Available at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_460.pdf.

Table 1. Types of archived operations data. (Continued)

Archived Operations Data Type	Description
Static, descriptive information	Data from public safety CAD systems, such as fire, emergency medical services, and law enforcement. Can include dispatch requests, incident types, severity, responder requests, or even lane status.
Decision-support response plans	The various actions that the departments of transportation (DOTs) are likely to take to help minimize congestion impacts and clear roads more quickly. Could include pre-programmed DMS messages, signal timing plans, traveler information strategies, detours, etc., that are grouped together into a single, cohesive "plan of action" ready to implement. The sharing of these response plans can help agencies to better coordinate so that one agency's response plan is not in conflict with another's.
Parking data	Location of parking facilities, number of spaces occupied and available, time and duration of parking space utilization, current fees, restrictions, and data on how to reserve a space.
Travel time	Often a derivative of speed data, travel time data represents the number of minutes it takes a person to travel from one location to another. Travel times are often divided into road segments where the start and end points of the segments are intersections or key features, such as bridges or tunnels. Vehicle travel time data can be derived from point sensor speed data. It also can be directly measured by probes, such as license plate recognition, toll tag transponders, Global Positioning Systems, and cell phone tracking. Alternatively, it can be estimated and predicted from other data sources.
Freight movements	Mixture of data related to the origin-destination (O-D) of various shipments or types of shipments, statistics on the type of goods being shipped, the mode by which the goods are shipped, value of the goods, quantity of goods, type of shipping container, and safety records.
O-D data	Tells operations personnel and planners where trips begin; where they end; and, sometimes, the routes that are taken. This data can be valuable for planning purposes and is useful for real-time operations when trying to measure the impact of various traveler information strategies and the impact of incidents on arterials and other secondary roads. Private data aggregator services routinely make O-D data anonymous.
Routing data	Data that can be used by both emergency first responders and the traveling public to determine the fastest route, shortest path, etc., from one point to another. Comprised of road network data, turning restrictions, speed limits, and other information related to route types and distances.

Benefits of Archived Data to Transportation Planners

Archived operations data provides numerous benefits to planners. Archived operations data is more representative of existing conditions than modeled data, can serve multiple purposes within a transportation agency, and enables new types of analyses to support better planning and investment decisions. Key benefits of using archived operations data in planning are discussed below.

Fosters Performance-based Planning and Programming

The use of archived operations data enables performance-based planning and programming at metropolitan planning organizations (MPOs) and State departments of transportation (DOTs) through the use of system performance data to guide decision making. As explained further in Chapter 5. Planning Opportunities for Archived Operations Data – Basic to Innovative, archived operations data are a critical part of many performance-based planning and programming activities, including:

- Setting outcome-based objectives.
- Developing and tracking performance measures.
- Identifying system performance needs and problems.
- Analyzing and evaluating scenarios, programs, projects, or strategies.
- Prioritizing and selecting programs and projects.
- Monitoring and evaluating the impacts of implemented programs and projects.

Provides a More Complete Picture of System Performance

Comprehensive real-world data can supplant synthetic data generated by a model and improve decision making in some applications. Because archived data is continuously collected, it overcomes the sampling error inherent in using small samples of data collected manually, which provides planners with a better reflection of reality. Large sets of operations data that are archived across long time periods are valuable in assessing the true variability and range of values. They can be used to evaluate more useful performance measures and provide verification that models are not unintentionally emulating unusual conditions.

Visualization of archived operations data inspires important insights into the performance characteristics of the transportation system, which supports more effective decisions on where to spend limited transportation dollars.

Opens Up New Types of Analyses to Support the Planning Process

The highly detailed nature of archived operations data allows for many types of analyses previously unavailable to planners, such that the planning process is able to address a wider variety of issues. For example, archived data facilitates the inclusion of reliability into long-range transportation plans (LRTPs). Travel time reliability could not be adequately measured or modeled without archived data, because it requires continuous data collection to see patterns over time.

Another benefit of using archived data over traditional data is that it allows the impacts of relatively infrequent events to be documented and analyzed. Such events may include a major winter storm, full closure of a roadway due to a severe incident, evacuations prior to hurricanes, and days when air quality standards are exceeded. This also enables planners to identify the likely causes of performance problems by correlating multiple data sets over time.

Enables More Sophisticated Modeling

Archived data can improve existing analysis tools and enable new ones by providing more complete input data than have been traditionally used, as well as more extensive data for calibration and validation. Planners also are able to calibrate more detailed models, such as activity-based and dynamic traffic assignment models, that benefit from archived operations data. The use of archived operations data at agencies around the United States enables the use of tools developed under various Strategic Highway Research Program 2 (SHRP 2) projects that present new reliability and capacity modeling tools and analysis methods.

Definitions of Reliability

Travel time reliability is typically defined in one of two valid ways.

- The variability in travel times that occur on a facility or for a trip over the course of time.
- The number of times (trips) that either "fail" or "succeed" in accordance with a predetermined performance standard or schedule.

In both cases, reliability (more appropriately, unreliability) is caused by the interaction of the factors that influence travel times:

- Fluctuations in demand (which may be due to daily or seasonal variation or by special events).
- Traffic control devices.
- Traffic incidents.
- Inclement weather.
- Work zones.
- Physical capacity (based on prevailing geometrics and traffic patterns).

These factors produce travel times that are different from day to day for the same trip.

Examples of Planning Use of Archived Operations Data

Washington State Department of Transportation:

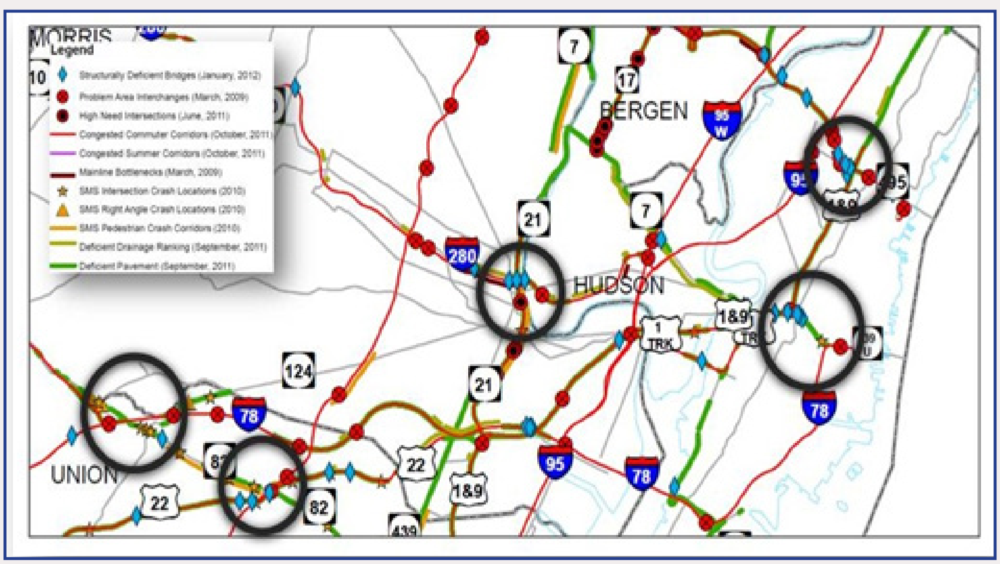
The Washington State DOT (WSDOT) has multiple applications for archived operations data to support their planning and overall performance management processes. WSDOT conducts before and after studies of congestion mitigation projects or programs as funds allow, typically with loop detector data in the Puget Sound region or automatic license plate recognition technology if funded by the project, and then provides the results as part of the Gray Notebook or in the Annual Congestion Reports. The agency also uses data from the NPMRDS. WSDOT recently reported on the effects of tolling operations and its traffic incident management program.

Houston-Galveston Area Council:

The Houston-Galveston Area Council used archived operations data to investigate how the timing of evacuation decisions affects traffic patterns. Archived operations data was used to identify potential improvements during the Hurricane Katrina and Hurricane Rita evacuations and to advance future plans for elements, such as the timing of government evacuation announcements and employee dismissals.

New Jersey Department of Transportation:

The New Jersey DOT (NJDOT) uses archived operations and safety data to visualize and identify candidate project areas (Figure 1).



Source: New Jersey Department of Transportation.

Figure 1. Map. New Jersey Department of Transportation management system integration of candidate project areas.

Examples of Planning Use of Archived Operations Data (Continued)

Maryland DOT State Highway Administration:

Archived operations data is used to support travel demand forecasting and simulation model calibration (estimating model parameters) and validation (confirming model performance). The Maryland DOT State Highway Administration (SHA) uses probe-based speed data from a third party data vendor and the I-95 Corridor Coalition's Vehicle Probe Project (VPP) Suite to calibrate its traffic simulation models, which serve as the basis for analyzing future scenarios for long-term planning. SHA also uses speed and traffic count data to calibrate existing conditions for travel demand and mesoscopic models, including the Maryland Statewide Transportation Model (MSTM).

Primer Purpose and Overview

The purpose of this primer is to assist transportation planners in effectively using archived operations data for developing, analyzing, and evaluating transportation plans and programs. This primer addresses the needs of planners from organizations with very little archived operations data, as well as organizations with an overwhelming amount of data. It raises planners' awareness of the opportunities afforded through archived operations data and provides guidance on how to take advantage of that data to expand and improve planning practices. It also identifies new and innovative applications for archived operations data in planning. This primer helps planners and their operations data partners overcome the barriers to obtaining and using data, regardless of whether they are just getting started with using the data or have substantive experience.

This primer helps address the numerous reasons why archived operations data has not been fully adopted for transportation planning and programming. For example, planners may not be aware that the data exists or may have difficulty accessing it. They also may not know how to mine, manipulate, manage, or use the data to support their work. In addition, data collected for real-time system monitoring can be extremely large. While distilling the data to inform investment decisions is achievable, it may require new expertise or tools.

The remainder of this document is organized as follows. Readers are encouraged to jump to the section of this primer that speaks to their specific need at any given time.

Chapter 1: Introduction – introduces the reader to the application of archived operations data for planning, including the benefits, a range of uses in planning, and real-world examples.

The purpose of this primer is to assist transportation planners in effectively using archived operations data for developing, analyzing, and evaluating transportation plans and programs.

Chapter 2: Meeting a Range of Planning Needs with Archived Operations Data – provides readers with an overview of a spectrum of planning activities that are enabled or enhanced with archived operations data. The case studies in Chapter 5 give an in-depth look at how archived operations data can be used for individual planning activities.

Chapter 3: Conquering the Challenges of Using Archived Operations Data – provides concrete guidance on handling institutional and technical challenges associated with obtaining, archiving, analyzing, and incorporating operations data into transportation planning activities.

Chapter 4: Obtaining Archived Data that Planners Need – informs readers about data sources, collection techniques, format, quality, and archiving.

Chapter 5: Planning Opportunities for Archived Operations Data – Basic to Innovative – offers rich case studies of archived operations data used to support planning and programming, which are often drawn from real-world experiences of planning organizations across the United States.

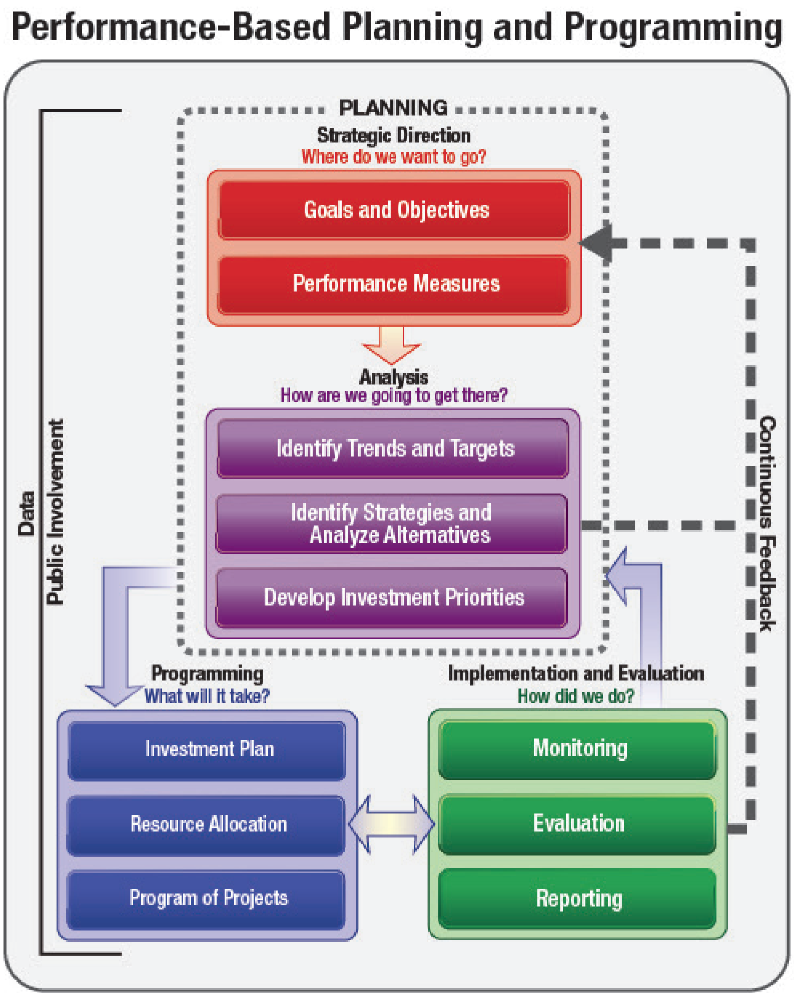
Chapter 6: Getting Started – presents readers with tips and a checklist on how to get started using archived operations data, whether they have a little or a lot of data.

2. MEETING A RANGE OF PLANNING NEEDS WITH ARCHIVED OPERATIONS DATA

Archived operations data can be used to support existing planning methods, as well as enable new ones. This chapter introduces the planning activities that can be improved through the use of archived operations data. Many of these planning applications will be revisited using detailed case studies in Chapter 5, Planning Opportunities for Archived Operations Data – Basic to Innovative.

Performance-Based Planning and Programming

Performance-based planning has received a lot of attention from transportation agencies in the past 5 years and is a basic underpinning of the Fixing America's Surface Transportation (FAST) Act requirements for State departments of transportation (DOTs) and metropolitan planning organizations (MPOs). Figure 2 shows the basic process of performance-based planning and programming.[3] Conceptually, performance-based planning is similar to the congestion management process (CMP) in that agencies identify specific objectives and needs, take actions, and continuously monitor conditions with performance measures.



Source: Federal Highway Administration.

Figure 2. Diagram. Flow chart of performance-based planning and programming.

3 Federal Highway Administration, *Performance Based Planning and Programming Guidebook*, Report Number: FHWA-HEP-13-041 (Washington, DC: 2013). Available at: http://www.fhwa.dot.gov/planning/performance_based_planning/pbpp_guidebook/.

For the most part, undertaking performance-based planning would be impossible without archived operations data. Below are examples of how archived operations data enables key elements of performance-based planning:

- **Monitoring and evaluating system performance to identify performance issues, prioritize projects, and develop reports.** Archived travel time data forms the basis for computing a wide variety of congestion, reliability, and freight performance measures. The data is usually available at high spatial and temporal resolution (e.g., 1 minute measurements by individual link of roadway location—where segments can be as small as just a few hundred meters) and can be aggregated to many other levels (e.g., peak period by route). For example, Texas Agricultural and Mechanical University's annual Urban Mobility Report now uses archived operations data to produce congestion and reliability metrics by urban area.[4] Many agencies also produce mobility performance reports using this data.[5]

  Travel time data is not the only useful type of archived data for performance-based planning. Analysis of travel times reveals trends and identifies problems, but it does not indicate what is causing them. Additional operations data on factors that may be causing congestion can be used to understand travel time trends. This data includes incident and work zone characteristics (especially duration by amount of blockage), weather patterns, and demand fluctuations. Figure 3 shows an example of how incident data can be used to "drill down" into the details of an event to explain outcomes—in this case, travel time reliability. Likewise, other contributors to congestion, such as work zones, can be broken down into smaller components to examine in more detail and compare against archived data. This not only helps to explain changes in the outcomes of an event, but also can lead to changes in response activities. In the case shown in Figure 3, by examining an incident timeline from an agency's advanced transportation management system (ATMS), it is possible to identify which incident response actions take excessive time to complete. Several incident data aggregation and visualization tools have been developed that helps planners identify trends in operations, environmental conditions, geographies, and road usage, which can affect safety. As one example among many, see Indiana's Mobility Performance Report (available at: http://docs.lib.purdue.edu/imr/).

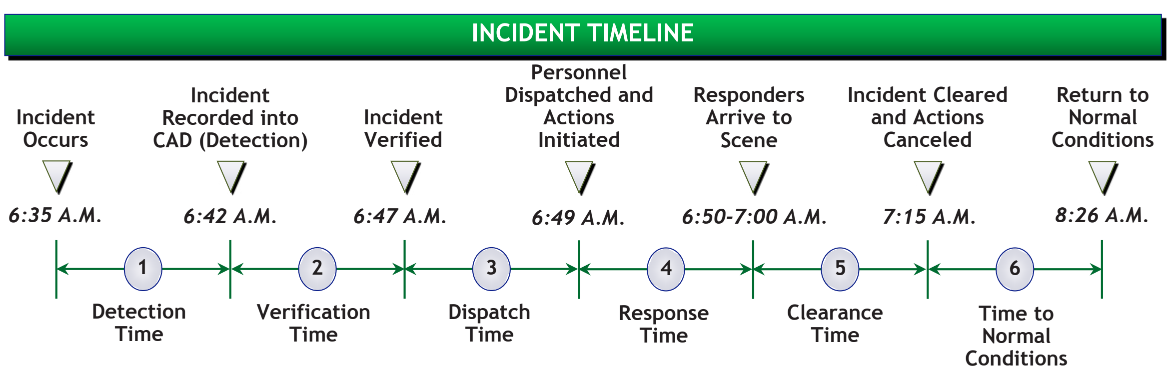


Figure 3. Graphic. Timeline displaying time taken for each activity during the incident.

4 Texas Agricultural and Mechanical University, 2012, *Annual Urban Mobility Report* (College Station, Texas: 2010). Available at: http://mobility.tamu.edu/ums/.

5 As one example among many, see Indiana's Mobility Reports, available at: http://docs.lib.purdue.edu/imr/.

- **Setting performance targets.** The FAST Act requires agencies to set performance targets and to monitor progress toward those targets. Archived data provides information on past trends that are useful in establishing targets (i.e., "it is good to know where one has been before deciding where to go."). When combined with forecasts, trend data provides a picture of what targets are attainable. Archived operations data also can be used to understand why targets were or were not achieved.
- **Evaluation of completed projects.** A key component of performance-based planning is the ability to conduct evaluations of completed projects and to use the results to make more informed investment decisions in the future. State DOTs and MPOs routinely invest a great effort in forecasting the impact of capital, operating, and regulatory improvements on general and truck traffic trips, but they seldom have the opportunity to evaluate the outcomes of their investments, except in isolated cases where special studies are conducted. The availability of archived data is changing this situation—special data-collection efforts no longer have to be undertaken. Archived data enables the establishment of an ongoing project and policy evaluation program.

  There are multifold advantages to instituting a formal evaluation program. Successes can be highlighted and failures can be enlightening. When a project does not perform as planned, the evaluation may indicate remedial action that can be taken. In addition, by documenting the measured impacts of a project or strategy, that information can be used by analogy when estimating the potential benefits of similar projects in the future.
- **Trip-based mobility monitoring and accessibility.** Although archived origin-destination (O-D) travel time data is just becoming available, it is expected to become more prevalent as technology evolves. O-D data makes it possible to monitor the performance of trips taken by travelers. By contrast, the current state-of-the-practice is to monitor performance from the facility's perspective, because measurements are taken at the facility level. A comprehensive mobility measurement program will involve using both trip- and facility based measures because they each inform analysts about the nature of mobility in a region. Many transportation investments, as well as land use and development policies, are oriented to the entire trip-making process; therefore, understanding the full trip performance experienced by travelers is important. The use of emerging operations strategies, such as active transportation and demand management (ATDM) and integrated corridor management (ICM), also require information on the entire trip-making process. Additionally, measuring accessibility of employment, recreation, shopping, and other destinations is a logical extension of trip-based mobility performance.

Planning Model Development: Analyzing Alternatives and Identifying Future Performance Issues

Archived operations data enables planners to significantly improve the performance of their existing planning models and expand the range of models they can use. Long-range transportation plans (LRTPs), as well as most of the studies conducted by planners, utilize models to identify deficiencies and compare the impacts of alternative investments. The travel demand forecasting (TDF) and traffic simulation models used by planners for these purposes are growing increasingly sophisticated in both their formulation and data requirements. Archived operations data provides a convenient source of data for these planning models. A major advantage of archived operations data is its high degree of temporal and spatial coverage. Traditionally, data to operate, calibrate, and validate models was based on relatively small samples. Archived operations data provides much greater detail in terms of coverage. The following list describes opportunities for archived operations data to support planning models.

- **Model inputs.** Several inputs to TDF models and transportation system performance monitoring can be developed from archived data:
  - Peak-hour volumes (PHV), annual average daily traffic (AADT), peak-hour factor (K-factor), and directional factor (D-factor) can be derived from continuously collected volume data. Link capacities can be empirically derived. Improved vehicle-miles of travel (VMT) data for TDF and air quality models, as well as for safety analysis, also can be developed with archived data.
  - From continuously collected speed and travel time data, free flow and peak-hour speeds and travel time distributions can be established for individual facilities. For air quality models, hourly speeds by link for the base condition can be developed with archived data.
- **Model calibration and validation.** The major outputs from TDF models are volumes and average speeds. These can be checked against archived data. Also, interzonal travel times can be developed from archived data; this can be used to calibrate the trip distribution phase of TDF modeling.
- **Emerging modeling frameworks**. As TDF and traffic simulation models grow increasingly sophisticated, so do the data requirements to run the models. This is especially true for reliability analyses. Recently, traffic simulation models have been applied to reliability studies by making multiple runs, where each run (or scenario) represents various combinations of conditions that influence reliability. For example, demand levels may be set at low, medium, or high, and the presence of an incident may be defined by different blockage and duration levels. In the scenario approach, the probability of each occurrence for each combination must be known so that the results can be properly combined via weighting. Archived data for the facilities being studied is analyzed to determine the probabilities.

The Strategic Highway Research Program 2 (SHRP 2) was a national research program managed by the Transportation Research Board from 2006 to 2015 to find solutions for critical national transportation challenges: highway safety, congestion, and road and bridge renewal. SHRP 2 research produced several analytical procedures for estimating reliability (Table 2). These procedures are beginning to be put into practice by transportation agencies across the United States. The products from all of these research efforts require inputs derived from archived operations data. As noted in Table 2, several SHRP 2 Reliability products require link capacity values as input. Modelers tend to use calculated rather than empirical link capacities; however, during the pilot testing of the SHRP 2 Reliability products, it was found that using empirical capacities is critical to obtaining reasonable results.

While the use of archived operations data for planning is uncharted territory to some agencies, more and more agencies are using archived operations data to enhance and expand their planning capabilities. The next chapter will help agencies that have a range of experience to overcome the institutional and technical challenges commonly encountered.

Table 2. Granularity of Strategic Highway Research Program 2 reliability-oriented analysis tools.[6]

Strategic Highway Research Program 2 Project	Analysis Scale (in order of increasing complexity)	Input Requirements from Archived Data
Development of Improved Economic Analysis Tools (Project Number C11)	Sketch planning; system or project level.	Incident duration and frequency.
Evaluation of Cost-Effectiveness of Highway Design Features (Project Number L07)	Detailed sketch planning; mainly project level.	Incident and work zone lane-hours lost; hours of year with rainfall and snowfall above thresholds.
Establishing Monitoring Programs for Mobility and Travel Time Reliability (Project Number L02)	Performance monitoring and project evaluations using empirical data.	Demand variability; presence of incidents, work zones, and weather tagged to individual facilities.
Incorporation of Travel Time Reliability into the Highway Capacity Manual (HCM) (Project Number L08)	Project planning using HCM scale of analysis.	Volumes on links and ramps, demand variability, incident duration by blockage type, and link capacities and free flow speeds.

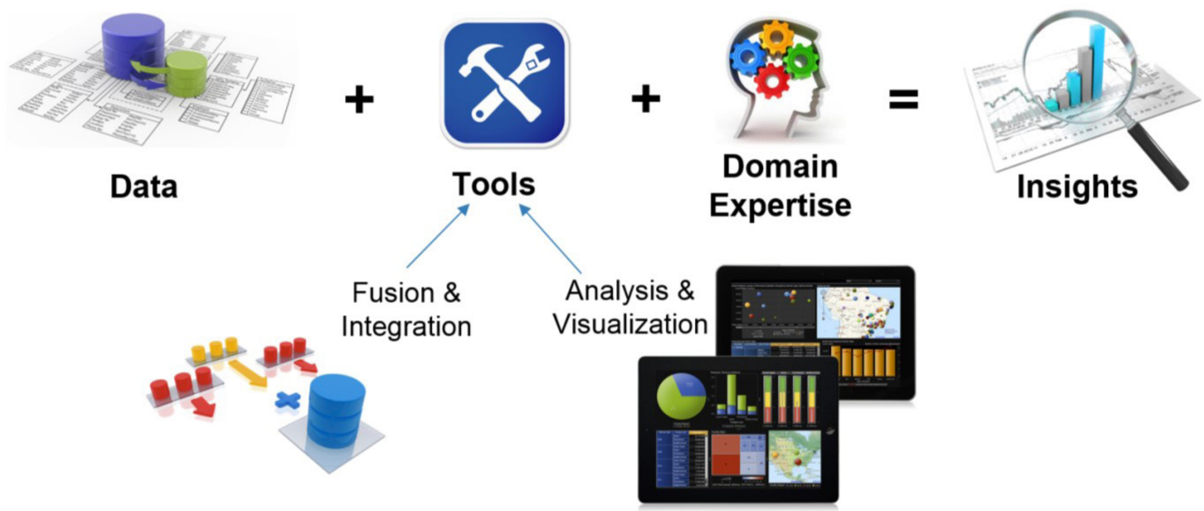
6 Additional information on any of the SHRP 2 solutions listed in Table 2 can be found at the Transportation Research Board's Strategic Highway Research Program 2 Web page: http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/Blank2.aspx.

Table 2. Granularity of Strategic Highway Research Program 2 reliability-oriented analysis tools. (Continued)

Strategic Highway Research Program 2 Project	Analysis Scale (in order of increasing complexity)	Input Requirements from Archived Data
Understanding the Contribution of Operations, Technology, and Design to Meeting Highway Capacity Needs (Project Number C05)	Project planning using mesoscopic simulation scale of analysis.	Volumes on links and ramps, demand variability, and link capacities and free flow speeds.
Partnership to Develop an Integrated Advanced Travel Demand Model with Mode Choice Capability and Fine-Grained, Time-Sensitive Networks (Project Number C10)	Regional planning using linked travel demand and mesoscopic simulation analysis.	Incident duration and frequency.
Incorporating Reliability Performance Measures in Operations and Planning Modeling Tools (Project Number L04)	Regional planning using linked travel demand and mesoscopic or microscopic simulation analysis.	Volumes on links and ramps, demand variability, incident duration by blockage type, and link capacities and free flow speeds.

3. CONQUERING THE CHALLENGES OF USING ARCHIVED OPERATIONS DATA

Data, tools, and domain expertise are the three components required to effectively use archived operations data in planning. Without these three basic components, users of archived data will be unable to derive the meaningful insights that help to make informed decisions that intelligently move an agency forward (Figure 4).



Source: University of Maryland CATT Laboratory.

Figure 4. Diagram. Components of an effective archived operations data archive.

Data. Chapter 1 provided an overview of the various types of archived operations data. Data captured by an agency or third-party may include speeds, volumes, accidents, freight movements, transit schedule and ridership, agency assets, and others. However, simply knowing that this data exists and is in an archive someplace is not sufficient; stakeholder buy-in and understanding how to access the data also are needed.

Tools. Often overlooked, the tools that are available to agency staff are the most critical element in making use of archived operations data. These can include tools that fuse "siloed" data from disparate sources, the tools that fill in gaps ("missing data"), and those that identify or screen outliers. Other important tools support analytics and visualization that help the agencies "see" into the data—asking questions, identifying issues, deriving meaning from the data, and communicating those insights to others.

Domain expertise. With the right tools and the right data, anyone can review a data set and start to ask questions, but to ask the right questions, to interpret the resulting analysis, and to otherwise act on the information in the archive, domain expertise is needed. Domain experts include transportation professionals with specific technical skills in areas such as traffic operations, transportation planning, or statistical analysis.

There is a mixture of technical and political challenges within an institution associated with creating an effective archive. All can be overcome with varying degrees of effort. The remainder of this chapter focuses on some of the more prevalent institutional challenges associated with the data, tools, and domain expertise of an operations archive.

Institutional Challenges

A common misconception about using archived operations data is that the technical challenges associated with the collection, fusion, and management of the data, along with the development of tools, are the greatest hurdles. In reality, these technical challenges are almost always easily solvable by qualified technical personnel. However, the institutional challenges (e.g., political, ideological, etc.) can create significant issues that delay and/or prevent operations data from being leveraged fully. The extent to which each of these issues may exist in each agency will vary greatly.

Technology policy. Many agencies have strict guidelines on what technologies can and cannot be used within its own enterprise. For example, some agencies use only Microsoft products and platforms, such as .NET and MS SQL; whereas, other agencies forbid open-source software for fear of being unable to find available support for a product or that misuse of the open-source technologies could lead to litigation.

Resource constraints. An agency may want to build and maintain their archive internally, but quickly find that they lack the capacity to do so. This may not necessarily be a lack of skills, but rather, a lack of skilled individuals. An agency who takes on a significant development effort with only one or two superior coders or minimal information technology (IT) support may find its program at a standstill if the specialized personnel leave the agency or are otherwise committed to other agency projects.

Storage and processing capacity. Technical barriers are common to agencies that are building archives in-house. Certain archives, especially those from connected vehicles or probe-based data sets, can grow at a surprising rate. The Florida Department of Transportation (FDOT), for example, collects data from approximately 8,200 speed/volume sensors. This sensor network produces the equivalent of 23.6 million data points every day and over 8.6 billion records each year. Sensor networks like the one FDOT uses produce only a fraction of the data generated by probe data sources. Using FDOT again as an example, the data the agency purchases from a third-party probe data provider produces nearly 34,000 records every single minute. That translates to 48.9 million records per day or 17.8 billion records per year. Due to State procurement and hiring policies, very few departments of transportation (DOTs) have dedicated and responsive IT teams that can plan out, install, and configure expensive hardware or cloud services to accommodate these massive archives for long periods of time.

Tools and accessibility. Even when hardware can be purchased and installed, developing the appropriate analytics software and databases that make the data easily accessible to end users can be a significant hurdle for agencies. For an agency to build successful tools independently, a healthy mixture of software engineers, architects, user interface and user experience design specialists, developers, and project managers is necessary. The tools will need to be maintained over time; therefore, complete documentation and staff that can be called upon over the course of many years to keep the tools up-to-date are needed. Because of the high barrier to entry and upkeep costs, many agencies are now choosing to either purchase existing tools, or leverage tools that other agencies or universities have already paid to develop. This effectively creates a pooled-fund approach to software development and maintenance. This approach is becoming easier to use for those agencies who are unaccustomed to purchasing services and for those who have historically not adopted tools and products that have not been invented in-house or even in-State.

Networking bandwidth. Though less common, agencies may have issues with networking and bandwidth capacity. Large data sets that stream from various sources will need to be compiled and streamed to others. The ability to send large sets of data from field devices or to and from third-party providers can be difficult with existing infrastructure.

Security concerns. Working with agency security boards or committees can be another challenge that often seems insurmountable. Getting the appropriate permission to set up online data access or to deploy specialized software, for example, can range from 1 month to several years, or even be denied if there is not an advocate.

If present, most technical issues are solvable if there is an internal advocate. Network capacity can be expanded at relatively low costs. Other IT and hardware procurement issues can be bypassed by outsourcing to contractors or through the purchase of software-as-a-service or IT-as-a-service.

Internal resistance. Internal resistance may develop when: (1) trying to convince the owner of the data or archive to share it with others, (2) trying to convince those who have not traditionally relied on operations data to trust its quality, and (3) trying to convince potential users of the value of the operations data compared to traditional (or currently used) data sets.

Source: Thinkstock/Dynamic rank.

Figure 5. Photo. Technician.

- **Sharing**—internal conflict over the ownership and control of data, information, and the capability that comes with it is not uncommon. It is important to understand the various motivations for wanting data versus the motivations for providing data. It is the responsibility of the owner and administrator of a particular data set to keep that data safe and secure (i.e., to be the steward). What motivation is there for releasing that data to larger audiences? At best, the data steward may be recognized for his or her efforts to maintain the data or get a verbal "thanks." More likely, however, the data steward is fearful that, in releasing the data to others, the data may become scrutinized excessively (or inappropriately) by outside groups. Thus, the data steward becomes bombarded with questions. If the data is incorrect, or even misinterpreted, or if there is a security breach or something is damaged, his or her job could be jeopardized. It can seem like a lose-lose situation for the steward. For the majority of agencies, there is little to no apparent benefit to the steward to make the data more widely available.
- **Trusting**—for those who have yet to make use of archived operations data or those who have biases towards other types of data, deciding to look at operations data can be difficult. There is an old school of thought that goes: *"Data collected for planners is of higher quality than that for the operations folks. Those operations people only care about in-the-moment conditions, and do not take the same quality control initiatives that planners take on their data."* This is essentially a lack of trust in the quality of someone else's data, which is not perceived as being collected for the same purposes or need as someone else. This lack-of-trust issue is so pervasive that it is not uncommon to find more than one set of detectors at a single location: one for the planners, one for operations personnel, and a third for the private sector or other agency. Each group thinks the other group is not doing a good enough job in quality control (QC) or maintenance on their device. However, getting past the trust issue can be easier than many other challenges. A quick verification and validation process performed on the data can show where issues exist. Many third-party archive applications (e.g., Regional Integrated Transportation Information System (RITIS) and Iteris Performance Measurement System (iPeMS) have data quality checking modules that can quickly identify issues, patterns, and causality in just a few minutes. These tools can verify or deny the existence of an issue, and help to encourage fixing the problem. From an economic standpoint, it is often much more cost-effective to pool the resources of several departments to fix the quality of the operations data set rather than to continue to expend resources to collect data multiple times at the same location.
- **Valuing**—when existing processes are in place for collecting, processing, and cleansing non-operations data, proving that the operations data has enough value to warrant switching to use it (or augmenting existing data with the operations data) can seem daunting. If a business unit in an organization already has a well-formed process for collecting, processing, and managing data from another source, then what is the impetus for trying something different? Often, it is simply a matter of better marketing and demonstrating the benefits of enhanced or augmented data by the operations groups. More and more, operations data sets, especially those from probes, are proving to be highly reliable, of finer granularity, and to have greater spatial coverage than traditional planning

> data sets. The fact that they are continually collected also is a positive aspect of these data sets—even though that does present additional data management burdens. Ultimately, an internal advocate is needed to clarify the benefits and quality of the data and to make the case for either switching to it or augmenting existing data with it.

Additional arguments have been used to overcome issues with sharing, trusting, and valuing operations data with varying degrees of success. There is the sensible argument, which debates the philosophical and ethical arguments of sharing data and pooling resources. There is the legal justification, which attempts to leverage interagency agreements or executive orders to change behavior. There is the funding argument, which attempts to pay an agency or business unit to switch their data collection and sharing practices in an effort to get past the "the cost is too great" excuse. There is the "shame" argument, which attempts to show how everyone else in the organization is sharing data and leveraging resources, thus isolating those individuals who are resisting. The problem with these tactics is that they do not address the "fear factor" that many individuals have about changing their behavior and sharing their data with others.

By showcasing the tools and capabilities that will be made available when the operations data is archived and accessible, others will want to make use of the data and share it as well.

There is the "make a friend" tactic, which leverages personal connections to try to instill change. This tactic often is successful as it is communications-based. However, if only the leadership within an organization is convinced, but the individuals who daily work with the data are unconvinced, then as soon as the ally within the organization gets promoted, retires, or otherwise leaves the organization, the resistance returns.

Ultimately, every manager and analyst in every agency has questions that need answers and is looking for insights into their own data. If these managers and data stewards can be shown the power of certain visualization and analytics tools that they will gain access to as part of the archive, they usually will understand the value of the data, realize that it is worth the supposed risk of using the data, and make it available to others through sharing. This is the "build it and they will come" approach. By showcasing the tools and capabilities that will be made available when the operations data is archived and accessible, others will want to make use of the data and share it with others. Tools can be built in-house with some risk, as described later. However, existing tools and applications developed by third parties or other agencies also can be showcased. Leveraging existing tools to showcase the power of operations data often is the quickest, cheapest, and least-risky approach to demonstrating capabilities, securing stakeholder buy-in, and getting a process up and running in as little time as possible.

Challenges in Changing Planning Methods and Products

The use of archived operations data may require planners to adjust their technical methods and modify their products. Historically, travel demand forecasting (TDF) models have been used not only to forecast future performance but also to estimate current performance where performance measurement data were lacking. With archived data, the use of models to “measure” current performance is no longer necessary. Therefore, products that formerly relied on models to derive current conditions now have to be converted to use measured data. Switching from models to data for these applications requires planners to develop a strategy. Some of the issues that need to be addressed include:

- **Coverage.** Will the archived data cover the roadways of interest? If not, how will the non-covered roadways be handled?
    - Example: Congestion management process (CMP). An agency only has archived operations data of speeds and volumes for area freeways, yet the current CMP monitors signalized arterials and collectors. The agency may decide to purchase travel time data from a commercial vendor.
- **Reporting period.** How frequently does the archived data need to be obtained: monthly, quarterly, or annually?
- **Maintaining continuity in trends**. If a trend line exists from the models, how will the switch to archived data affect it?
    - Example: CMP. An agency has been producing a biennial mobility performance report that is based on applying its TDF model. It is changing to vendor-supplied travel time data. A test revealed that, for the current year, the performance measures developed with the data are significantly different from the modeled measures, resulting in a discontinuity in trends. A possible solution is to use the results from the current year tests to adjust the previous years’ performance. This can be done by noting the percent difference between modeled and measured data for the current year to develop adjustment factors by time of day and highway type.
- **Definitions of performance measures.** Will the measures developed with archived data represent the same underlying phenomenon as the modeled measures?
    - Example: Delay estimates using TDF models. Delay computations require three inputs: a reference (usually free flow) speed, actual speed, and the volume of vehicles exposed to a speed. TDF models are used to develop these estimates; all three delay inputs are available at the link level; therefore, delay is computed for each link and summed. If vendor supplied travel times are used, there are numerous sources of incompatibility: the vendor and TDF networks usually will not align; reference speeds may be calculated differently; and, most significantly, companion volumes to the travel time estimates do not exist. Likewise, the estimates of speeds for air quality models would benefit from using both travel time and volume data so that the proper weighting can be applied.

The most desirable solution is to conflate the TDF and vendor travel time networks. This ensures that delay estimates are made on the same spatial basis and makes the volumes from the TDF model available for use with the travel time measurements. Note that conflation is a complicated process that requires both networks to use a common geolocation referencing system (e.g., latitude/longitude). If the TDF network does not have “points of contact” (either attributes or shared line work), the vendor network may be conflated to other traffic count databases that do provide that information, such as the Highway Performance Monitoring System (HPMS).

- **Acceptance of new performance measures**. How will the introduction of new performance measures enabled by archived data fit into current products?
  - Example: Including travel time reliability in the long-range transportation plan (LRTP). Improving travel time reliability is increasingly appearing as a stated goal of LRTPs. When this is the case, methods for monitoring and forecasting should be in place. Reliability performance can be measured directly with archived data, but what role does archived data play in forecasting reliability? Recent experience suggests there are at least two ways to use archived data. First, some metropolitan planning organizations (MPOs) have used archived data to reformulate the volume-delay functions used in their TDF models. By using continuously collected travel time data, the effects of the factors that cause unreliable travel (e.g., weather, work zones, incidents, or demand fluctuations) are accounted for in the revised functions. Second, archived data can be used to develop relationships between model-produced average travel time or speed values and reliability metrics. Both of these applications are discussed in more detail in Chapter 5.

4. OBTAINING ARCHIVED DATA THAT PLANNERS NEED

This chapter provides information aimed to help planners understand where they can obtain archived operations data, which data elements are particularly useful for planning, and basic information on how data can be processed to become useful information. This chapter also identifies critical considerations for agencies looking to implement a data archive to support planning.

Locating and Accessing Archived Operations Data

Planners who are interested in obtaining archived operations data have several options. Depending on the type of operations data desired, planners may find it within their organization, other public agencies (e.g., State departments of transportation (DOTs), metropolitan planning organizations (MPOs), local traffic departments), university research laboratories, or from the private sector. If critical data is not available, planners may be able to get the resources needed to collect the data; although, this primer does not provide guidance in data collection. Chapter 6 provides a detailed list of specific sources for archived operations data, including some that are available for no cost. For example, the Federal Highway Administration (FHWA) provides the National Performance Measures Research Data Set (NPMRDS), which contains archived speed and travel time data for the National Highway System, to MPOs and State DOTs for free. The four cases below provide examples of where planners at MPOs and State DOTs are obtaining archived operations data to advance their planning activities.

Capital District Transportation Committee

The Capital District Transportation Committee (CDTC) is the MPO for New York's Capital Region operating out of Albany, New York. CDTC maintains a summary of traffic volume data for Capital District roadways in its *Traffic Volume Report*. The *Traffic Volume Report* includes data for all State and non-State roadways with functional classifications greater than local. CDTC collects data for this report through different mechanisms, including obtaining traffic counts from the MPO's member agencies and the New York State DOT's Management Information System for Transportation (MIST) data set.[7] Table 3 contains highlights of the operations data obtained by CDTC and its use in planning.

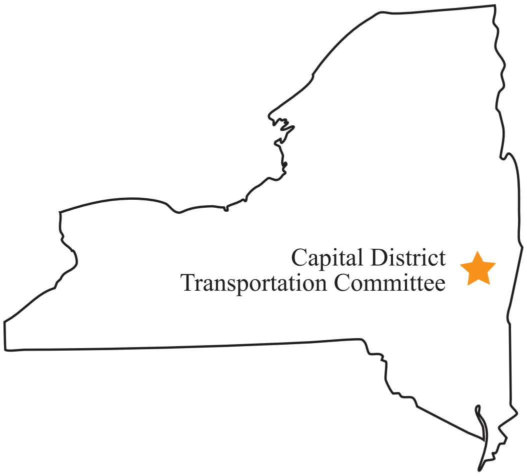


Figure 6. Map. Outline of New York State.

7 Capital District Transportation Committee (CDTC), *CDTC Traffic Volume Report Data*, September 2011. Available at: http://www.cdtcmpo.org/tvr/tvrdes.htm.

Table 3. Highlights of operations data available to Capital District Transportation Committee.

Data Type	Data Source	Who Collects the Data	Planning Activity Supported
Volume	Traffic counts, loop detectors.	Local agencies.	Performance measurement.
	Management Information System for Transportation (MIST), loop detectors.	New York State Department of Transportation (DOT).	Performance measurement, model improvement.
Speed	MIST, National Performance Management Research Data Set (NPMRDS), Global Positioning System vehicle probes.	Private sector.	Performance measurement.
Incidents	MIST, crash.	New York State DOT.	Identifying high-incident locations.

CDTC has found value in using archived operations data for identifying transportation needs and problems. Specifically, it has access to the MIST data set that provides volume, speed, and incident data along I-87 and I-90. The committee has used its travel demand model to identify and quantify congestion problems. The model is focused on AM and PM peak-hour traffic on a typical weekday. The MIST data provides a more comprehensive understanding of weekday congestion patterns on the Interstate system. For example, MIST data demonstrated that average operating speeds on the Interstate system, in the absence of incidents, were generally higher than predicted by the model. The Northway (I-87) had a similar amount of delay as I-90, but the MIST data indicated that reliability was worse on the Northway, consistent with a popular perception that the Northway has worse congestion. The MIST data also highlighted that congestion was especially severe on Fridays in the summer, when vacation travel mixes with commuter travel. This perspective was overlooked by focusing on typical PM peak-hour data. MIST data allows the MPO to understand and weigh the costs and benefits of investments to address non-recurring and recurring delay. It contributed to CDTC establishing a strong priority for operations and incident management and a strong policy in the congestion management process and regional transportation plan that expressways in the region will not be widened to address congestion except in the context of managed lanes.

> The MIST archived data including volumes, speed, and incident data allow the CDTC to understand and weigh the costs and benefits of investments to address non-recurring and recurring delay.

CDTC has used archived operations data to better understand its travel demand model. Comparing the outputs from their model with the archived data, CDTC learned that its models were not as effective in predicting delay as they had believed. CDTC intends to use archived data to improve these models.

CDTC contracted with a consultant to develop a tool to assist in the analysis of archived operations data. A previous challenge to analyzing the data was that the MIST data set was large and consisted of many unlinked tables within the database. Prior to the development of the tool, conducting analysis was a labor-intensive process requiring coordination between many of these different tables to understand the data.

Maricopa Association of Governments

The Maricopa Association of Governments (MAG) is the MPO for the metropolitan Phoenix area. The primary sources of archived data used at MAG come from the Arizona Department of Transportation's (ADOT's) Freeway Management System (FMS) and private sector probe data. The ADOT FMS data provides volume, speed, and vehicle classification, but is limited to approximately 69 percent of the urban freeway system. The data purchased from the private sector and NPMRDS data provides speeds for the for the entire freeway and arterial system, but does not provide volumes. MAG has purchased probe data for the last 4 years. MAG has been using the NPMRDS data for travel model validation as well as for calculating congestion-related performance metrics. Table 4 identifies the operations data types obtained by MAG and how MAG uses the data to support its planning activities.

Figure 7. Map. Outline of Arizona.

Table 4. Highlights of operations data available to Maricopa Association of Governments.

Data Type	Data Source	Who Collects the Data	Issues with Data Format or Quality	Planning Activity Supported
Classification	Arizona Department of Transportation (ADOT) Freeway Management System (FMS) loop detectors.	ADOT	Covers only 69 percent of freeway lane miles.	Model calibration, validation, baseline for activity-based model, scenario planning, program rebalancing, project evaluation and ranking.

Table 4. Highlights of operations data available to Maricopa Association of Governments. (Continued)

Data Type	Data Source	Who Collects the Data	Issues with Data Format or Quality	Planning Activity Supported
Volume	ADOT FMS counts.	ADOT	Covers only 69 percent of freeway lane miles.	Performance measurement, model improvement.
	Counts.	Local agencies	None identified.	Same as above.
Speed	ADOT FMS.	ADOT	Covers only 69 percent of freeway lane miles.	Same as above.
	Private-sector-purchased probe data and National Performance Measures Research Data Set (NPMRDS).	Private sector.	Supplements data on freeways from ADOT FMS and covers arterials. Only provides speed and not volume.	Same as above.

In 2012, MAG developed in-house an interactive performance dashboard that allows all audiences to access the various multimodal performance outcomes at the system and corridor levels. The dashboard, called MAGnitude, is supported through a partnership between ADOT, MAG, and Valley Metro. It provides reliability, delay, congestion, volume, speed, and other measures on a map-based interface that can be tailored to specific corridors or system wide. Users can look at data each year from 2009 to 2015 on maps, bar charts, and infographics. Archived ADOT FMS data, MAG counts, safety, and private-sector probe data fuel the website.

▸*Maryland DOT State Highway Administration*

The Maryland DOT State Highway Administration (SHA) is a transportation business unit of the Maryland DOT that owns and operates the State's the State's highways outside of Baltimore City. SHA's data collection program is comprised of both private-sector and public-sector data to evaluate existing projects and programs and identify short-term and long-term improvements. It applies multiple analysis and visualization tools to inform transportation decision-making and "tell a story" with archived operations data that clearly communicates the need for projects and outcomes of implementation in high-priority locations.

Maryland DOT SHA gathers archived operations data from multiple sources:

- The State's advanced transportation management system (ATMS) platform, called Coordinated Highways Action Response Team (CHART), collects detailed incident and event data in real-time. This same system also collects speed and volume data from deployed intelligent transportation systems (ITS) sensors, road weather information systems (RWIS) data, construction and lane closure information, virtual weigh station data, and other ITS-related information.
- In addition to ATMS-collected data, SHA collects other vehicle classification, volume, and speed data from traffic counts, floating car runs, and automatic traffic recorders. This data is housed in the agency servers and publically accessible through the SHA Internet-Traffic Monitoring System (I-TMS) website. Data housed in the traffic count databases are collected through a uniform process for the whole network for FHWA Highway Performance Management System (HPMS) requirements. Traffic data also is collected on a project-by-project basis.
- Probe-based speed and travel time data are purchased from the private sector and accessed through the Vehicle Probe Project (VPP) Suite.
- For arterials and lower-functional-class roadways, SHA obtains volume data and truck classification data.
- There is also Bluetooth coverage for speed data along certain segments of the SHA network, including I-270, which has high-occupancy vehicle (HOV) lanes.

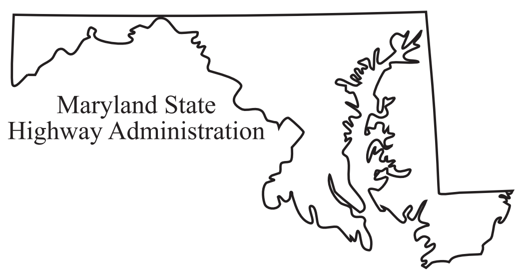


Figure 8. Map. Outline of Maryland.

All of the above-mentioned incident, volume, speed, and ITS device data are fused and archived through the Regional Integrated Transportation Information System (RITIS) and SHA systems. A list of archived operations data types and sources available to SHA are described in Table 5.

Table 5. Highlights of operations data available to Maryland Department of Transportation (DOT) State Highway Administration (SHA).

Data Type	Data Source	Who Collects the Data	Planning Activity Supported
Classification	Internet-Traffic Monitoring System (I-TMS), traffic counts, automatic traffic recorders.	Maryland DOT SHA.	Feasibility studies, project planning, highway design, performance monitoring.
Volume	Coordinated Highways Action Response Team, traffic counts, automatic traffic recorders.	Maryland DOT SHA.	Monitor/report congestion on arterials, model development.

Table 5. Highlights of operations data available to Maryland Department of Transportation (DOT) State Highway Administration (SHA). (Continued)

Data Type	Data Source	Who Collects the Data	Planning Activity Supported
Speed	Private sector vehicle probe speed data.	University of Maryland (UMD) - Center for Advanced Transportation Technology Laboratory.	Calibrate models, evaluate programs or projects, performance reporting.
	Floating car runs.	Maryland DOT SHA.	Calibrate models.
	Bluetooth.	Maryland DOT SHA.	Performance reporting, origin-destination patterns.
	I-TMS, traffic counts, automatic traffic recorders.	Maryland DOT SHA.	Evaluate programs or projects.
Incidents	I-TMS, traffic counts, automatic traffic recorders.	Maryland DOT SHA.	Safety and reliability analysis for project planning.

Oregon Department of Transportation

The Oregon DOT (ODOT) manages over 19,000 lane-miles of highway. Among other ITS equipment, ODOT has deployed over 384 highway cameras, 150 freeway ramp meters, 126 portable message signs, 117 road and weather information stations, and 220 permanent variable message signs.[8] Along with the corresponding sensors and support systems, many of these deployments assist ODOT in collecting and using archived operations data.

Portland State University (PSU) archives data from ODOT's freeways, including volume, speed, occupancy per lane, and vehicle classification. ODOT recently upgraded its ramp meter firmware and software to collect vehicle classification data by length and speed (four length categories with average speed per category) for all ramp meter locations. In addition, PSU archives all available traffic signal data, including Bluetooth and bus information from the arterials in Oregon, under the Portland Oregon Regional Transportation Archive Listing (PORTAL) effort. ODOT is actively adding new locations to install Bluetooth sensors on arterials and freeways to get a better understanding of travel times on the arterials. Once available, the intent is to archive this data in PORTAL. ODOT has also purchased private sector probe speed data. Table 6 lists the types of archived operations data available for ODOT planning and its sources.

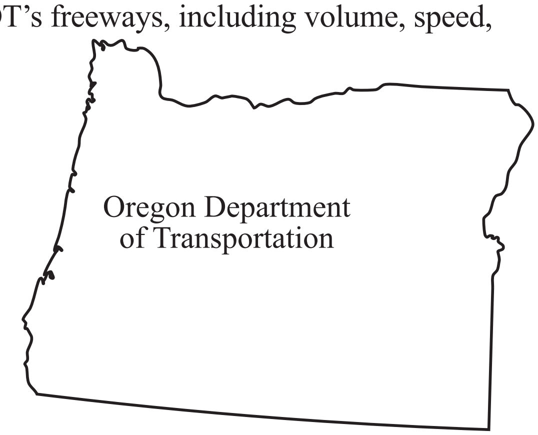


Figure 9. Map. Outline of Oregon.

8 Oregon Department of Transportation, *Key Facts 2012*. Available at: http://www.oregon.gov/ODOT/docs/keyfacts_2012.pdf.

PSU manages and conducts all quality control (QC) and quality assurance of the archived data. ODOT has a steering committee to provide oversight for the PORTAL effort. While some of the ODOT data systems have built-in processes that alert ODOT to potential issues (e.g., sensor damaged or not functioning), PSU monitors the quality of the PORTAL data and sends alerts to ODOT when it detects data that may not be accurate (e.g., unrealistic travel speeds).

Table 6. Highlights of operations data available to Oregon Department of Transportation.

Data Type	Data Source	Who Collects the Data	Planning Activity Supported
Volume	Loop detectors.	Oregon Department of Transportation (ODOT).	None reported.
Speed	Loop detectors.	ODOT.	None reported.
Vehicle Classification	Ramp meters.	ODOT.	None reported.
Incidents	Crash data system.	ODOT.	Oregon Driver and Motor Vehicle Services crash reporting forms.

Obtaining Third-party Data and Tools

Third-party data is becoming increasingly popular among DOTs and MPOs as a way to more economically collect data on a wide geographic scale. Examples of third-party data can include:

- Probe-based speed/travel time data from companies like HERE, INRIX, TomTom, etc.
- Origin-destination (O-D) data.
- Incident data.
- Weather data.

The procurement of these data sets can dramatically improve capabilities within an agency, making it easier to measure system-wide performance, identify problems, prioritize projects, and have greater situational awareness.

Whenever an agency is considering the procurement of third-party data—whether that be probe-based speed or travel time data, incident and event data, weather data, etc.—the most critical component of any procurement will be contract language controlling acceptable use. Far too often, agencies procure data without specifying terms of use in the request for proposal (RFP). If terms of use are not specified, then the data provider will specify their own terms of use. When this happens, the following can occur:

- Acceptable use can be too restrictive to be of value to the agency (e.g., by limiting use to internal, real-time applications).
- Agencies may be restricted from archiving the data at all.

- Agencies may be restricted from using third parties to archive the data.
- Agencies may not be able to share the third-party data with any researchers (internal or external) or with consultants working on behalf of the agency on specific projects.
- Each respondent to an RFP could have widely varying acceptable use terms that could drastically change the cost models of each respondent, which can make comparing responses and offerings in a reasonable way difficult.

When developing RFPs, agencies frequently overlook their position of power. If the agency proposes its own desired acceptable use terms, the private-sector data providers will respond and price the data accordingly.

The I-95 Corridor Coalition has set a standard for acceptable use when acquiring third-party probe data, and their language could be considered in other State procurements of probe and other data (weather, incident, etc.). A copy of the coalition's data use agreement can be found online (http://i95coalition.org/wp-content/uploads/2015/02/VPPII_DUAv9_signed.pdf?dd650d) and as can a copy of the actual RFP (http://i95coalition.org/wp-content/uploads/2015/02/RFP-83794N-FINAL-2.pdf?dd650d).

Source: Thinkstock/Rex Wholster.

Figure 10. Photo. Interstate highway shield for I-95.

As archived operations data sets continue to increase in size, quality, and availability, planners are realizing that is data alone are not sufficient to spark innovation and use in the planning process. Sophisticated analytics technologies, decision support tools, accessibility mechanisms, or other visualization tools are needed to leverage these archives, understand systems performance, make informed investment choices, and realize the full benefits of these archives. Visualization has far-reaching potential for communicating transportation needs to leaders who must prioritize budgets, enhance the ability of transportation organizations to deliver timely and ever-more complex programs within those budgets, educate the engineers who will make it all happen, and communicate to the traveling public the implications of transportation investments.

Numerous archived operations data analytics tools are available today. These tools can be purchased as-is, with existing data already incorporated into them from around the world. They also can be purchased and integrated within an agency's existing archived data sets and can be highly customized to meet agency-specific goals and criteria. Examples of these products include the Iteris Performance Measurement System (iPeMS) suite of products and the University of Maryland Center for Advanced Transportation Technology (CATT) Laboratory's RITIS suite of products.

Data Elements Useful for Planning

Intelligent Transportation System Roadway (Point) Detectors: Volume, Speed, and Lane Occupancy

ITS roadway detectors have been deployed on freeways for several decades. Detectors provide lane-specific measurements of volume, speed, and lane occupancy. The data is transmitted from the field as 20- to 30-second summaries and is archived at levels up to 15 minutes (5-minute aggregations are very common). Detectors are spaced at roughly1/2-mile intervals in dense urban areas and include mainline lanes, as well as HOV and ramp lanes in some systems. A variety of technologies are used to collect the data, including inductance loops, radar, and video image processing, but they all result in the same basic data elements. Speeds are measured over very short distances—the length of the detection zone—so they are essentially "spot" speeds (time mean speeds). The speeds are usually transformed to travel times by assuming a length of highway where the speed is constant, usually half the distance to the next upstream and downstream detectors.

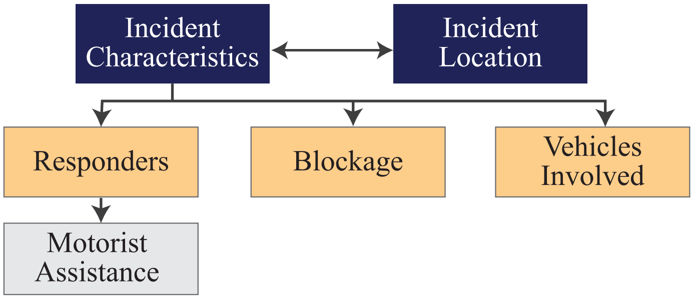
On signalized highways, use of data from roadway-based detectors is far less direct. Some detectors are placed at midblock locations (away from signals) and measure the same conditions as freeway detectors. However, midblock speeds are not indicative of signalized highway performance, as the effect of the signals is not included.

Probe Detection - Travel Times

Toll tag reader and roadway-based vehicle re-identification technologies can provide travel times measured between two fixed points. Global Positioning System (GPS) vehicle re-identification provides travel times between two points in time or instantaneous speeds. These measurements are usually "snapped" to a roadway segment, as vendor-supplied travel time data is. As discussed above, companion volume measurements are not available with these technologies, as only a sample of vehicles are collected.

Incident Data

Incident characteristics are an extremely important data source for conducting planning analyses. This data describes incident location, characteristics, and the nature of blockages (duration of lane or shoulder blockages). Although incident data is being collected today by transportation and emergency response agencies, complete uniformity has not yet been achieved. In general, the types of data available are shown in Figure 11.



Source: Transportation Research Board, National Cooperative Highway Research Program.

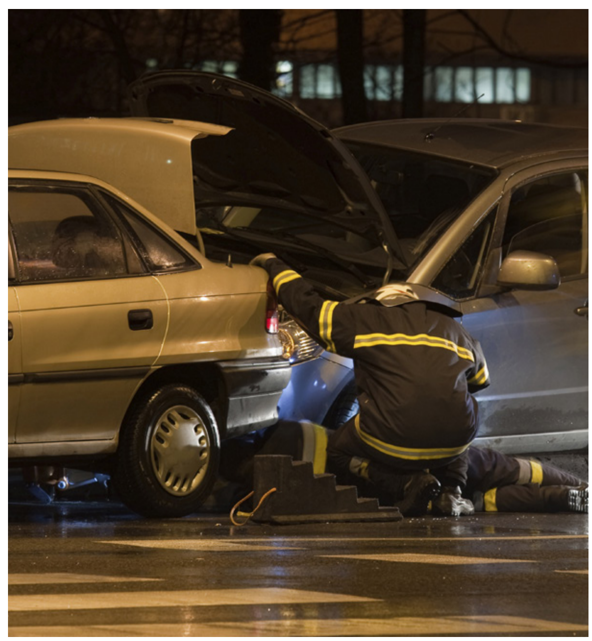
Figure 11. Diagram. Incident data model for performance monitoring.[9]

9 Cambridge Systematics et al., *Guide to Effective Freeway Performance Measurement: Final Report and Guidebook*, Web-Only Document 97, Transportation Research Board, Washington, D.C., 2006. Available at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w97.pdf.

The most important data elements in these categories are:

- Incident type (crash, disabled vehicle, fire, debris, and abandoned vehicle).
- Incident collision type (fixed object, overturn, vehicle/side, vehicle/head-on, and vehicle/rear-end).
- Incident "timeline" data (time stamps for the start of the incident, detection, verification, on-scene arrival, lane/shoulder open, and all clear).
- Incident severity ("KABCO" injury scale).
- Incident location (route, travel direction, milepost, or GPS coordinates).
- Incident blockage.[10]
  - Cross-section feature affected (lane, partial lane, right shoulder, left shoulder, median, and off of maintained way).
  - Lane type (through/general purpose, through/HOV, auxiliary, on-ramp, and off-ramp).
  - Begin/end time.
- Number of involved vehicles.

The duration of the blockages caused by incidents is the most important piece of information needed for planning applications, but others affect adjacent traffic flow conditions. Also, some incident data may be incomplete, and the type of incident may be used as an indicator of the lane closure impacts (e.g., for data-checking purposes). This data is usually as readily available from databases, especially those that already have incident blockage/duration information; therefore, there is no additional effort involved in assembling this data.



Source: Thinkstock/csakisit.

Figure 12. Photo. Incident clearance.

Possible sources of data include the following (these are not mutually exclusive, and multiples may be used in combination to feed data into the system):

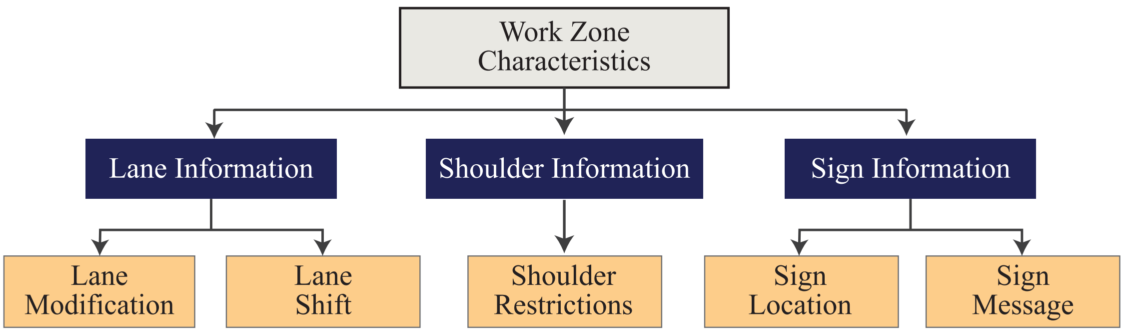
- Service patrol/incident response team data entry – Personal digital assistant or laptop in the vehicles.
- Transportation management center (TMC) operator data entry – Automate the conversion of web information to database entries.
- Police computer-aided dispatch (CAD) files.

10 Ideally, blockage is updated every time it changes during a given incident.

▸ *Work Zone Data*

For planning applications, work zone data is useful in understanding congestion patterns and explaining why certain congestion trends emerged. Work zone data is sometimes collected as part of incident data, but also independently by construction and maintenance personnel. Figure 13 shows the basic categories and structure for work zone data. Major types of data on work zone characteristics are lane information, shoulder information, and sign information. For each of those categories, there is data such as lane modification, lane shift, and shoulder restrictions.

Work zone characteristics – The actual and planned changes in the roadway environment created by the work zone. These are used to measure the extent of work zones in time (duration) and space (amount of existing highway removed for the work zone) and their impact on safety and mobility. They also can be used in traveler information services to alert motorists to expected work zone conditions. In general, the types of information that should be collected are work zone type, longitudinal characteristics and extent (including details on transition zones and tapers), duration of work zone characteristics, and major cross-section characteristics of the work zone. Both actual and planned changes should be tracked separately.



Source: Transportation Research Board, National Cooperative Highway Research Program.

Figure 13. Diagram. Work zone data for planning applications.[11]

Work zone activity – The activities related to traffic management and construction/rehabilitation in a work zone. These are used to assess mobility and safety impacts of traffic control plans and motorist guidance as well as improvements in construction planning and execution. Data includes specifications in traffic control plans, times traffic control plans are in effect, traffic control device placement in the field and times used (e.g., pavement markings, dynamic message signs (DMS) and static signage, positive guidance devices, barriers, and other new technologies), construction and rehabilitation field activities (e.g., crew size by task, task duration, and equipment used on-scene), and time-of-day and where in the work zone the work occurred. *This level of detail is usually not necessary for planning applications unless evaluations of work zone management practices are undertaken*.

11 Cambridge Systematics, et al., *Guide to Effective Freeway Performance Measurement: Final Report and Guidebook*, Web-Only Document 97, Transportation Research Board, Washington, D.C., 2006. Available at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w97.pdf.

Special Event Data

These kinds of events are somewhat different than work zone events in that they occur repeatedly and more often in the same location (e.g., sports stadiums, concert halls). Only a minimum amount of data on special events is required and is used primarily to document surges in traffic volumes:

- Event name and location.
- Start/end dates and times.
- Approximate attendance (may be segmented by time periods).
- Changes in operating policies, including start/end dates and times: HOV restrictions, contraflow, and ramp meter timing.

As with work zone data, special event data is useful in explaining and understanding congestion patterns; that is, continuously collected volume data will detect the volume increases due to special events, but planners will want to tag those surges as special events for performance reporting.

Weather Data

Unlike volumes, travel times, incidents, and work zones, the vast majority of weather data collection is conducted by non-transportation agencies. There has been limited activity in developing performance measures for weather activities; of those that do exist, they relate to the duration/extent of weather events, weather prediction versus actual weather that materialized, and snow and ice removal activities. However, the influence of weather on traffic flow, safety, and customer satisfaction is substantial. Details on the nature of weather events that influence traffic characteristics, travel decisions by users, and activities by agencies need to be collected:

- Type of weather – Precipitation type, fog, dust, and high winds.
- Duration of the event.
- Intensity of the event – Depends on type of weather: rainfall/snowfall rates, visibility, and wind speed.
- Geographic coverage of the event – Ideally, roadway segments affected by the weather event.
- Treatment strategies – Treatment strategies related to snow and ice control require coordination with maintenance personnel. These include pretreatment of roadway surfaces, as well as removal of snow and ice. Other treatment strategies that are more directly controlled by operations personnel include supporting access for emergency and maintenance vehicles and using RWIS and forecast data to identify areas for pretreatment.
- Control strategies – Speed limit control, retiming of ramp meters, and traffic signals to accommodate reduced capacity and vehicle speeds; road closures; lane use control; vehicle type restrictions; and aggressive incident management (e.g., more service patrols than normal).
- Advisory strategies – Highway advisory radio (HAR) and DMS weather messages, 511 weather messages, website weather messages, and weather messages to private Internet service providers.

Signal Control Data

This data includes:

- Operational status of signals at intersections or ramp meters, such as operational, maintenance mode, flashing, or offline.
- Signal timing plans and current or future timing schemes.

Source: Thinkstock/Carsten Reisinger.

Figure 14. Photo. Traffic signal.

For planning applications, signal timing and phasing can be used to set the capacity used in travel demand forecasting (TDF) and mesoscopic simulation models. In addition, recent work has highlighted how data from signals and approaches can be used to develop certain types of arterial performance measures. Day et al. developed an approach based on using high-resolution controller event data.[12] They developed a portfolio of performance measures for system maintenance and asset management; signal operations; non-vehicle modes, including pedestrians; and travel time-based performance measures for assessing arterial performance. The focus of this work is clearly on the signal operation and maintenance, but several aspects are of interest to measuring arterial corridor performance.

Other Operations Data Useful for Evaluations

Planners are increasingly relied upon to conduct evaluations of completed projects and new policies. In fact, an ongoing evaluation program is a hallmark of performance management. Several other sources of operations data are available for these evaluations, even if this data is not useful for other traditional planning applications:

- **Device operational status –** Data on the operational status of roadway devices, including traffic detectors, DMS, traffic signals, HARs, RWIS, and closed circuit television (CCTV) cameras, where available.
- **Managed lane status –** Data on when HOV restrictions are in effect, direction of reversible lanes by time of day, and price of high-occupancy toll lanes by time.
- **Parking data –** The location of parking facilities, number of spaces occupied and available, time and duration of parking space utilization, current fees, restrictions, and data on how to reserve a space.
- **Incident activity: institutional –** The institutional setting, arrangements, and protocols for incident management are extremely important determinants of performance; these will determine how long incidents are in effect (incident duration), thereby influencing performance. Institutional activities may include:
    - Agency participates in a formal multi-agency regional or Statewide program to coordinate management of traffic incidents that contains strategic planning, a program plan, and annual work plan.

12 Day, C.M., E.J. Smaglik, D.M. Bullock, and J.R. Sturdevant, *Real-Time Arterial Traffic Signal Performance Measures*, Publication FHWA/IN/JTRP-2008/09, Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 2008. doi: 10.5703/1288284313439.

- Agency participates in a team that meets on a regular basis to evaluate and improve coordinated incident response and to address traffic problems as well.
- Agency has developed a concept of operations document for incident management.
- Multi-agency contact list has been developed in the area containing the names, phone numbers, pager numbers, and other pertinent information for the appropriate response personnel.
- Incident Management (Incident Command) System used on-scene to manage traffic incidents.
- Legal specification by State law or formal agreement as to who is in charge at the scene of a traffic incident (Incident Commander).
- A plan has been developed and adopted by responding agencies for staging and parking response vehicles and equipment at a traffic incident site in a manner that minimizes lane blockage and facilitates the re-opening of lanes.
- Respondents protected through law or court opinion for liability claims for damages to vehicles or cargoes during clearance activities so long as the removal was not done in a careless or grossly negligent manner.
- TMC and Service Patrol active incident management times (e.g., 24 hours, 7 days).
- State or local jurisdiction has a law that requires drivers involved in a property-damage-only accident (where vehicles can be driven) to move the vehicles from travel lanes to a safe location to exchange information or wait for police.
- Quick clearance laws or policies.
- Length of time abandoned vehicles are allowed to remain on a freeway shoulder (assuming they are not an imminent hazard).
- Laws or policies regarding the removal of stalled or abandoned vehicles from freeway shoulders are in effect.
- Policies and procedures are in place to facilitate quick removal of heavily damaged vehicles and non-hazardous cargoes.
- Procedures are in place for removing crash victims from incident scene.
- Transportation agency and public safety agency are colocated in a TMC.

■ **Incident activity: physical –** The equipment and resources necessary to conduct effective incident management also will have an effect on incident duration and performance:

- Incident detection algorithms (used with traffic sensors).
- Method of incident notification to all responders (communication between agencies).
- CCTV coverage (percent of study segment capable of being viewed with CCTV).
- Degree of wireless communication for on-scene management.
- Service patrol vehicles per centerline mile by time period.
- TMC floor personnel per centerline mile.

Connected Vehicle Data

Data from connected vehicles, once considered a futuristic data source, is now available through both third-party vendors at the national level and from connected-vehicle pilot sites in several States. Presently, at the national level, commercial vehicles rolling off of the assembly line have the ability to provide connected vehicle data in real-time. This can be information about driver behavior and vehicle diagnostics, including:

- Speeding.
- Harsh braking/acceleration.
- Antilock braking system (ABS) engagement.
- Stability control.
- Temperature sensors.
- Wiper use.
- Headlight use.
- Air bag deployments.
- Emissions data.
- Oxygen sensors.
- Seat belt use.
- Tire pressure.

The same providers also are able to furnish real-time intersection turning movements and near-real-time O-D data in an anonymized way. This real-time operations O-D data reflects actual trip-making patterns between a subset of nodes. At the smallest scale, this data can include turning movements at interchanges. At medium scales, it can track corridor distribution of traffic. At the largest scale, it reflects region-to-region trip demand.

Source: Federal Highway Administration.

Figure 15. Image. An illustration of interconnected infrastructure.

Having access to this type of data has significant impacts on real-time operations, but can also have major implications for planners and researchers. Having increased access to O-D data and safety condition information from connected vehicles will enable new safety analytics, new methods of problem identification, and more extensive model validation data sets.

Data Sources and Collection Techniques to Ensure Usefulness in Planning

Archived operations data exists at very detailed temporal and spatial scales. Continuous data from ITS roadway detectors can be archived at as low a level as 20-second intervals by lane. Currently, available data from commercial travel time vendors exists at 1- to 5-minute intervals by roadway link. As vendors increase the sophistication of their products, additional detail may become available. Data from traffic signals can include phase-by-phase and, sometimes, second-by-second conditions. Processing this data into a useable form poses several challenges.

What levels of aggregation should be used for storing the data? Aggregation is a tradeoff between computational efficiency and storage capacity versus level of detail available for applications. Developers should carefully consider what level of aggregation best balances these tradeoffs. Sound archiving practice suggests that all raw or originally collected data should also be maintained for a period of time, even if it is offline.[13] When aggregating for storage, data is often transformed by combining it with other data (e.g., delay is a combination of travel time and demand). For both straight aggregation and transformations, it is useful to document the "lineage" of the data so that users can better understand its structure. This documentation requires that structured metadata be maintained as part of system that houses the data.[14]

What QC procedures should be used to check the data for accuracy? True QC procedures would be based on comparison to independently collected data, even if it is on an audit basis, as is done for the I-95 Corridor Coalition.[15] However, most agencies do not have the resources to conduct continuous data audits, although some limited one-time tests may be conducted. Instead, the majority of QC processing must be done post-hoc with reasonableness checks, comparison to traffic flow theory, and outlier detection. Once potential data errors are detected, it must be decided if the data should be deleted or flagged as erroneous. (Flagging is preferred so as to preserve the original data for other uses.) Then, a decision has to be made if the erroneous data are to be replaced with imputation.[16] Imputation carries its own set of issues, and users must decide if imputed data should be used. For example, performance measures that are "counts" or "sums" require complete data for the time period in question (e.g., traffic volume and delay). Other measures, such as speed, may be treated as a sample and computed without imputed data, if users determine a sufficient sample size exists.

How should the data be turned into measures? The starting point for developing performance

13 As specified in *ASTM Standard Guide for Archiving and Retrieving ITS-Generated Data*. Available at: http://www.astm.org/Standards/E2259.htm.

14 As specified in A*STM Standard Practice for Metadata to Support Archived Data Management Systems*. Available at: http://www.astm.org/Standards/E2468.htm.

15 Hamedi, M., Haghani, A., Zamir, K., and Wang, Z., *I-95 Corridor Coalition Vehicle Probe Project: Validation of HERE Data, Report for North Carolina (#7) US-29 and US-74*, February 2016. Available at: http://i95coalition.org/wp-content/uploads/2015/02/I-95-CC-Validation_Report-NC07-HERE-FINAL2.pdf?dd650d.

16 See Chapter 3 of the Strategic Highway Research Program *2 L02 Final Report.* Available at: http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_S2-L02-RR-2.pdf.

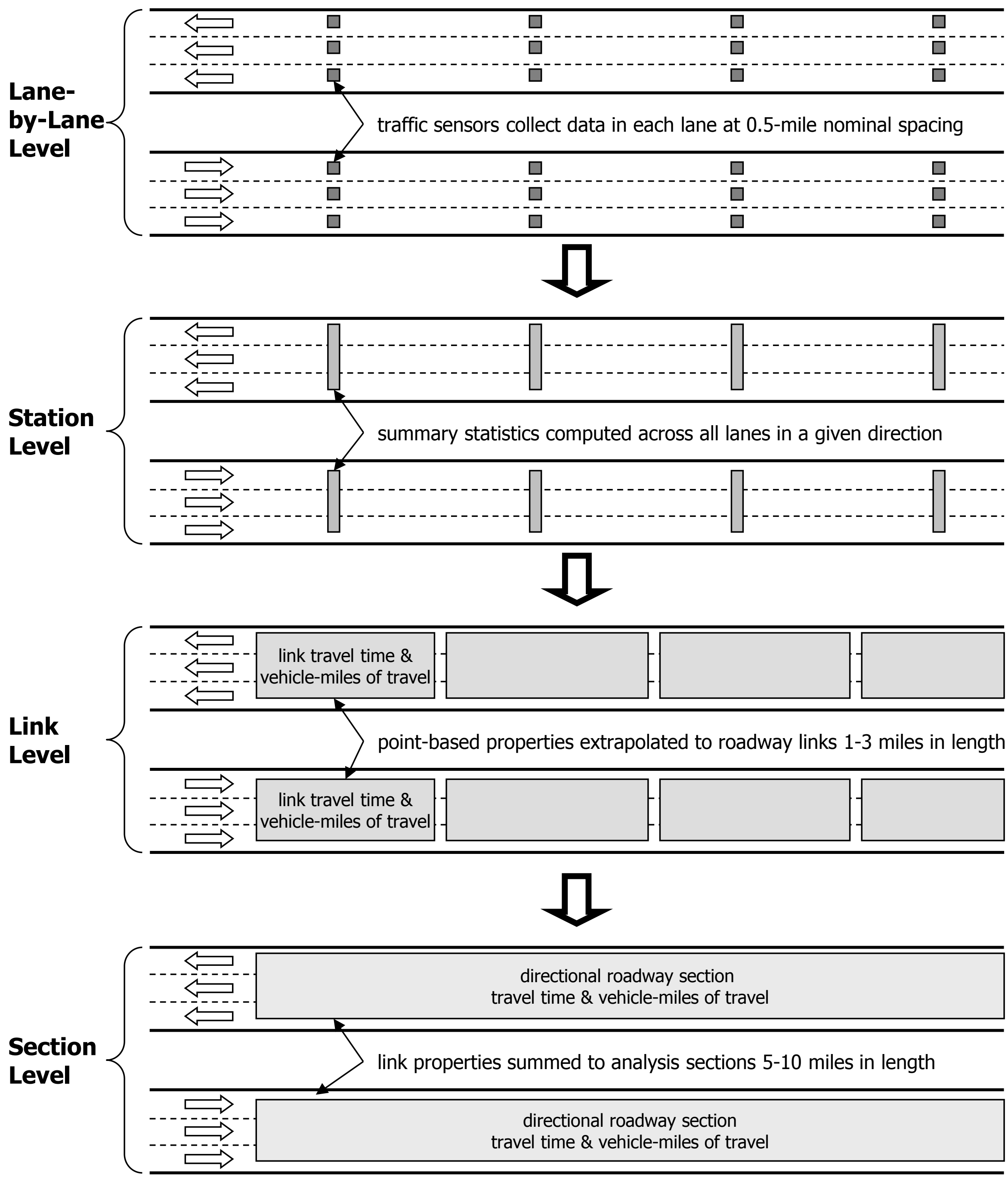
measures for mobility is travel time data. Over the past 15 years, travel time data availability has increased exponentially, starting with data from ITS roadway detectors and continuing to the current generation of vehicle probe data from private vendors. However, ***none of this data represents directly measured travel times*** as defined by measuring the passage of individual vehicles as they traverse the network. To create travel times from this data, transformations are required. For roadway detectors and the current generation of privately-supplied vehicle probes, this means assuming that the measurements of speeds apply equally over some distance, usually the length of the reporting segment. Users should be aware that the travel times so created are an approximation of actual travel times. These restrictions do not apply to data collected from individual vehicle probes, such as GPS-equipped "floating cars," roadway systems that measure the passage of individual vehicles via electronic signatures (e.g., Bluetooth readers), and toll tag systems.

In the near future, it is expected that directly measured travel times will become available from private sources. As of this writing, some vendors are developing travel times for origin-destination pairs. In the longer term, data from connected or automated vehicles could supply origin-destination travel times. As this data becomes available, it should replace the current generation of data used for performance measurement. Until that time, the following guidance is offered.

Performance measures and data inputs to models typically are summarized temporally and spatially. For example, mobility performance often is computed for "facilities," usually several miles in length and encompassing several ITS roadway detectors or probe data links. Two methods can be used to compute performance measures on such facilities. The first method computes the measures for each spatial unit, and then combines them to get facility performance (e.g., speed and travel time distributions). The problem posed by this method is that, while it is possible to combine means from links on the facility (the harmonic mean would be used for speed), statistical theory indicates that variances and percentiles cannot be combined because speeds (or travel times) on successive links are correlated.[17] The second method is based on first developing travel times (and vehicle miles of travel (VMT) where available) for the entire facility by the temporal level in the original data. The distributions from which measures are produced are then based on the entire facility, not stitched together from its components. Figure 16 presents a graphic summary for how facility travel times are computed from ITS roadway detectors reporting data by individual lanes.[18] (Note: A computation procedure for developing the measures using this approach will likely be included in Chapter 24 of the 2016 update to the Highway Capacity Manual (HCM).)

17 A summary discussion and extension of this phenomenon can be found in the Strategic Highway Research Program *2 L04 Final Report*. Available at: http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_S2-L04-RR-1.pdf.

18 Turner, S., R. Margiotta, and T. Lomax, *Monitoring Urban Freeways in 2003: Current Conditions and Trends from Archived Operations Data*, Report No. FHWA-HOP-05-018, December 2004. Available at: http://mobility.tamu.edu/mmp/FHWA-HOP-05-018/.



Source: Federal Highway Administration.

Figure 16. Diagram. Computing facility travel times from intelligent transportation system roadway detectors.

However, the second method is not a complete reflection of travel time conditions on the facility, because it is not based on the actual passage of vehicles over time. That is, it is a single "snapshot" of conditions for a given time period. To more closely replicate true travel times, the "virtual probe" (also known as "vehicle trajectory") method can be used to synthesize travel times. The virtual probe method "traces" the vehicle trip in time and applies the link travel time corresponding to the precise time in which a vehicle is expected to traverse the link. For example, a section travel time that begins at 7:00 AM will use a link travel time for 7:00 to 7:05 at the trip origin, but could use a link travel time from 7:05 to 7:10, or 7:10 to 7:15 at the trip destination. The virtual probe method attempts to more closely model the actual link travel times experienced by motorists as they traverse the freeway system.

For relatively short segments, the difference in results of the snapshot versus the virtual probe method is thought to be small. As the length of the facility increases, the differences are exaggerated by the time/space offset. For example, the Florida DOT (FDOT) recently developed a monitoring program for long-distance freight trips in the state based on the trajectory method to address this problem.[19]

The virtual probe method also can be applied to signalized arterials to develop performance measures from signal system and control data. This data is "event-based," as it relates to either the presence or absence of a vehicle at a point on or near the signal approach, or the status of the signal phasing. Liu et al. used the virtual probe method with both vehicle-actuation and signal phase change data in a synchronized manner.[20] This data is time-stamped, thus allowing the reconstruction of the history of traffic signal events along the arterial street. In the simulation, the virtual probe vehicle's trajectory is traced in time and space, and its status at any point in time is dependent on the underlying data. When the vehicle "completes" its trip, total travel time is recorded. In the absence of directly measured travel times and delays, this approach would work for measuring arterial performance.

Another major issue in the development of mobility performance measures is developing a complete data set from which to compute the measures. Such a complete set would include simultaneous measurements of both speed (or travel time) and traffic volumes. Volumes are used to compute some measures (VMT and delay) and are used to weight the results for sections with varying amount of traffic. ITS roadway detectors provide this structure, but vendor-supplied vehicle probe data does not; this data should be treated as a sample. As mentioned above in the section, Challenges in Changing Planning Methods and Products, vehicle probe and traffic count data sets are conflated to provide the complete set of data. Conflation brings its own problems, especially when many-to-many spatial relationships require allocation of the traffic data. Further, nearly all traffic counts are daily counts, which have to be decomposed into the time periods represented by the probe data. For this process, default temporal distributions are used that are developed from continuous count stations.[21]

19 Cambridge Systematics, Inc., Trip-Oriented Mobility Monitoring Framework Development Final Report, Prepared for Florida Department of Transportation, November 5, 2014. Available at: http://floridampms.com/Final%20Reports/Task%2012%20Trip%20Based%20Travel%20Time%20Reliability.docx.

20 Liu, H.X., W. Ma, X. Wu, and H. Hu, *Development of a Real-Time Arterial Performance Monitoring System Using Traffic Data Available from Existing Signal Systems*, Prepared for Minnesota Department of Transportation, December 2008. Available at: http://www.lrrb.org/PDF/200901.pdf.

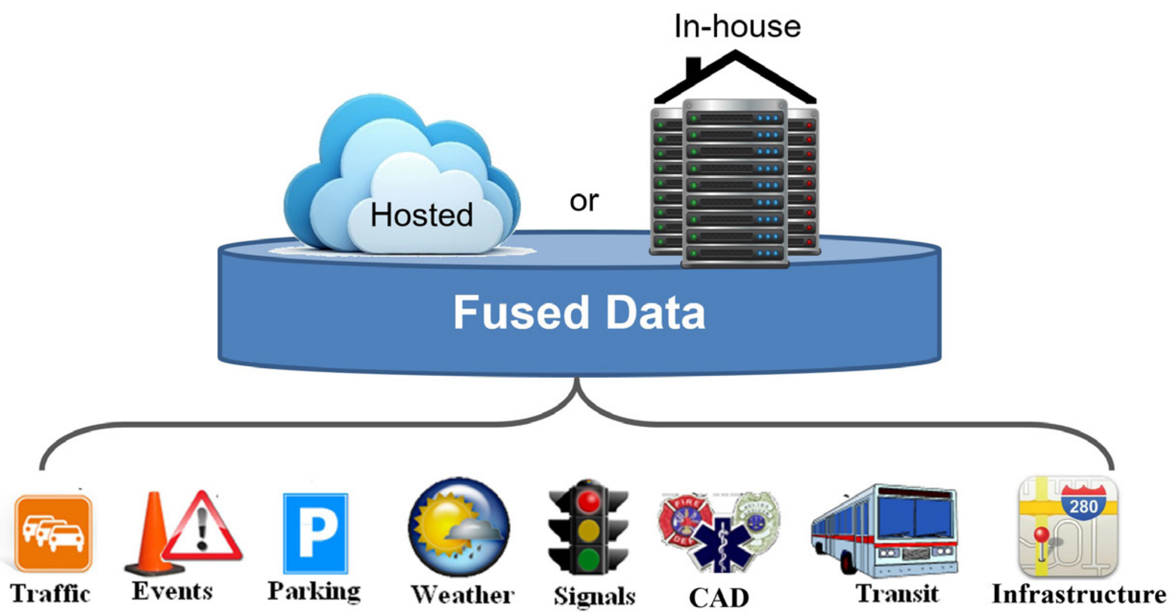
21 See Appendix A of Texas Agricultural and Mechanical University's *2012 Urban Mobility Report*. Available at: http://d2dtl5nnlpfr0r.cloudfront.net/tti.tamu.edu/documents/mobility-report-2012-appx-a.pdf.

Data Archives: Implementation Options and Common Mistakes Associated with Each

There are two options from which to choose when deciding how to implement an operations data archive (Figure 17):

- Agency builds/creates in-house.
- A third party specializing in data archiving or offering an existing product for sale can be procured. That third party can handle all aspects of the data archive, including ongoing maintenance.

In addition, it may be possible to find another agency that is willing to build and deploy the data archive; however, that agency becomes responsible for operations and maintenance moving forward.



Source: University of Maryland CATT Laboratory.

Figure 17. Diagram. Operations data archives built in-house or hosted by a third party.

There are pros and cons to each of these approaches, some of which may not be obvious at first.

In-House

If an agency believes it has the capacity and drive to develop, deploy, and host its own archived operations data, then it is important to keep the following in mind:

- Identify the data (where is it archived and who is responsible for it).
- Meet with potential users.
- Identify needs (but realize positive feedback may not initially occur).
- Let the needs dictate the technology (rather than vice versa).
- Focus more on usability, the user interface, and the questions or problems to be addressed.

Costs for databases vary greatly. Licensing of the database engine alone can range from no cost (for open-source database technologies) to several hundred thousand dollars per year. Databases are well suited for storing smaller data sets that are relational in nature. However, with massive amounts of data (hundreds of terabytes, petabytes, etc.), it begins to become quite expensive to implement fast search. To combat these growing costs and the need for speed and failover capabilities, Google developed some algorithms that allowed data to be divided into smaller "chunks" that were mapped to many computers. When calculations needed to be made, the data "chunks" were brought back together and processed rapidly. These algorithms were eventually open sourced and named Hadoop, which is available on many cloud platforms. Many public and private-sector companies have now adopted Hadoop-style technologies.

▸ *Third-party Hosting*

Often, one or more archives will be implemented through consultant support. This could be as simple as drafting a new task order for an existing ATMS provider, or it may be more complex and involve a new procurement going out to bid. If bidding a new procurement, there are a few things to consider, including:

- Avoid over-specifying implementation technologies. For example, it may be best not to tell a consultant that they must deploy something in "the cloud," or that they must use a specific type of relational database technology (e.g., Oracle, MsSQL, etc.).
- Consider how available an archive needs to be. If it is not being used for 24-hour/7-day mission critical operations, then it is likely not worth the extra expense to have a fully redundant/ hot-swappable system. Ensure data is backed up, but it may not be necessary to spend additional money on a hot-swappabl system with just over 5 minutes of downtime annually.
- Dedicate more time to working on requirements for usability, functionality, and recruiting multiple user groups to determine expected usage.
- Attempt to procure the services of a consultant who has performed similar work for other agencies. The archive may need customization and tailoring, but a proven provider is often more reliable than a standard consultant.
- Recognize that initial startup will be costly. There are several private-sector and university providers that have excellent archiving, fusion, and analytics products. Some of these systems work across borders and across multiple agencies. Consider adopting similar technologies or products as neighboring jurisdictions when possible so that shared experiences, knowledge, and benefits from shared resources can be leveraged.
- Avoid "black box" solutions that do not explain the underlying technologies, algorithms, or methods used to calculate the performance measures, despite the notion that it is prudent to purchase existing products. Ensure the chosen provider has documented procedures that can be shared with engineers and analysts. Some providers have multi-State or multi-agency steering committees that collectively drive the features of the archive products to ensure they are constantly meeting user needs.

Common Mistakes:

Mistakes made while procuring new data collection systems. When procuring a new system (e.g., ATMS, 511, or other), ensuring the "home" agency has access to and ownership of any data that is created is essential. With software-as-a-service-type contracts for some of these systems, it is not uncommon for commercial off-the-shelf products to claim ownership of all data created in a particular system. This is a particular issue for transit system operators who have purchased systems from third parties. These third parties have historically claimed all of the automated vehicle location (AVL), routing, and other data as "proprietary," and have either refused to allow it to be archived for other agency use or have charged exceedingly high fees to make the data available to other archives and for other use.

All agency systems (e.g., ATMS, AVL, 511, CAD, or other) should be capable of sharing, in real-time, all data fields in their database in some sort of data feed to trusted third parties (e.g., other DOTs, public safety agencies, trusted university partners). This includes any device configuration information, lookup codes, incident information, detector data, operator actions, static and dynamic Geographic Information System information (e.g., device locations, mile marker locations, etc.), or video. Making these statements within individual systems procurement documents will assist in avoiding large future expenses or being unable to access data in the desired ways.

Data ownership and acceptable use. As mentioned above, it is imperative for an agency to ensure ownership of all data entered into any of its systems. In other words, the agency should ensure that it is freely available to share the data with any third parties it so chooses. The one exception to this requirement is with third-party data that the agency may choose to purchase from a private company (e.g., Waze, SpeedInfo, Traffic.com, INRIX, HERE, TomTom). Whenever possible, the agency should negotiate terms of use for sharing third-party or privately collected data, so that the agency is free to share the data with any other non-commercial entity (e.g., other operational agencies, universities, public safety agencies). For agency-created and -owned data, the agency should always make every effort to be able to share their data with external parties without the need for complex legal agreements. If a data sharing agreement is necessary, the DOT should only reference Federal laws and should remain silent on governing law so that inter-State data sharing will be possible.

Persistence of data. All data feeds to an archive should be persistent, such that eventual data delivery to the data archive is guaranteed and no data is lost. Should there be a network connectivity issue or a power failure at the archive, the providing system should continue to hold the data in a queue until such time as the connection can be re-established and all of the data can be sent.

5. PLANNING OPPORTUNITIES FOR ARCHIVED OPERATIONS DATA – BASIC TO INNOVATIVE

This chapter showcases five important planning activities that are enabled by archived operations data:

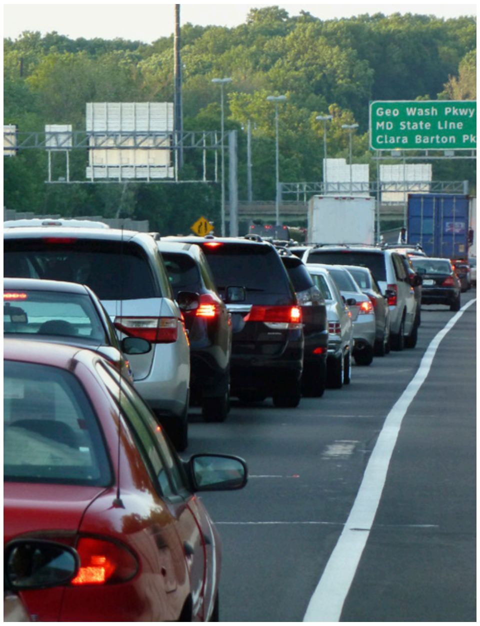
- Identifying and confirming the problem.
- Developing and reporting mobility performance measures.
- Using archived operations data for loading into and calibrating reliability prediction tools.
- Performing before and after studies to assess projects and program impacts.
- Identifying causes of congestion.

The sections below introduce each planning activity by identifying the need for archived operations data for this activity, the types of relevant data, analytical procedures that transform the archived data into useful information, and the impact that using archived operations data for that planning activity can achieve. Following the introduction to each planning activity, one or more case studies are provided to demonstrate, in detail, how archived operations data can be used and what agency outcomes were achieved. Each case study highlights the use of visualization and its influence on the planning activity.

Problem Identification and Confirmation

Need

All agencies have limited dollars available to improve mobility, safety, and aging infrastructure; therefore, most agencies attempt to identify the "worst" problems in their jurisdiction (i.e., problems that have the greatest impact, and most likely, the greatest potential for a high return on investment (ROI) should countermeasures be put into place). An agency needs an efficient way to identify, quantify, and rank safety and mobility issues, and to determine the cause of these issues. Because of the dynamic nature of transportation, it is important that an agency be able to perform these types of analysis efficiently and cost-effectively many times throughout a year to respond to external political influences and/or to public concern. The results of these analyses can then drive the long-range transportation plan (LRTP), influence funding decisions, or simply be used to respond to public concerns over why a particular project was chosen over another.



Source: Thinkstock/tomwachs.

Figure 18. Photo. Congestion on freeway.

Relevant Data

Historically, continuously collected permanent count station data and police accident reports would have been influential in this type of analysis. In some cases, temporary count station data may have been used, or density data may have been collected by aerial photography surveys. Both the count station and police accident data, while collected in real-time, will often have many months of lag between when the data is collected and when it becomes available to analysts.

Aerial photography surveys may be collected only for a few hours once every few years, and only at select locations where problems have been anecdotally noted. Temporary count stations are deployed only once every few years on select roadways, and they rarely collect data for more than 1 week. Because these data sources are not collected continuously, agencies must rely on modeling and forecasting techniques, which may not account for seasonal variations, incidents, or other dynamic conditions. Archived operations data, however, can help to reduce or even eliminate some of the challenges associated with the above data sources. Real-time flow data can be collected from sensors, toll tag readers, license plate readers, or probe data from third-party vendors (e.g., HERE, INRIX, or TomTom). Advanced transportation management system (ATMS) platforms also collect real-time event data and lane closure information from construction events, disabled vehicles, accidents, and special events.

Analytical Procedures

Operations data from probes or other high-density sensors can be used to identify congested areas or bottlenecks. When this data is integrated with ATMS event data, it becomes possible to identify if the congestion is recurring or non-recurring.

Flow data from probes and dense sensor networks can sometimes be challenging to work with because of the sheer volume of data, complex linear referencing systems that might not match up with agency centerline road databases, and limited agency technical and resource capacity. Agencies with strong, integrated information technology (IT) departments, dedicated data analysts, sufficient data storage hardware, and access to statistical software, such as SAS (Statistical Analysis System), can handle this type of analysis given time and proper funding.

However, many agencies subcontract their analysis to consultants, universities, or transportation departments, or they purchase off-the-shelf congestion analytics software and data integration services tailored for transportation planners.

Impact

Regardless of the method used to analyze operations data, agencies often find that the use of operations data can allow them to identify problems based on more timely data, which helps to instill greater confidence in the decisions that are made as a result. Also, when probe-based data is used for the analysis, the data has greater coverage than traditional loop detectors. Therefore, there is more confidence that problem identification analysis is covering “all roads” instead of just detected roads. Lastly, agencies who purchase off-the-shelf congestion analytics software find that the speed of analysis increases so significantly that resources are freed up and the agency is able to respond to requests much faster.

Case Study: Problem Identification and Confirmation at the New Jersey Department of Transportation

Overview

The New Jersey Department of Transportation (NJDOT) uses a combination of archived operations data and data processing and visualization tools to identify performance issues on the transportation system and develop easy-to-understand performance measures with visualization that speak to senior leadership, elected officials, and the public, in alignment with Fixing America's Surface Transportation (FAST) Act and organizational goals. NJDOT is leveraging a mixture of archived operations probe data, volume data, weather data, and event data to help identify congestion issues on the State's roadway network in part to create problem statements for consideration in its project delivery process.

The archive of these data sets has been made available to NJDOT planners through a web-based visual analytics tool known as the Vehicle Probe Project (VPP) suite. This suite gives planners the ability, with minimal effort, to automatically detect and rank the worst bottleneck locations in a county or entire State, determine if the congestion is recurring or non-recurring, determine causality, measure the economic impact, and produce graphics that can be inserted into analysis documents and presentation materials to be used with decision makers. The VPP suite is an outcome of a project directed by the I-95 Corridor Coalition that began in 2008 with the primary goal of providing Coalition members with the ability to acquire reliable travel time and speed data for their roadways without the need for sensors and other hardware. The VPP has evolved since that time and now also provides tools to support the use of the data and integrates several vehicle probe data sources.[22]

Data

NJDOT is using four archived data sets: probe, volume, event, and weather. The probe data is derived from a private-sector data provider and is archived at 1-minute intervals from a real-time feed that is originally used for travel time and queue estimation. The probe data is represented as the average speed across the length of a road segment for each 1-minute period.

The volume data used by the agency is derived from Highway Performance Management System (HPMS) data that is also provided back to the agency through a third-party data provider.

The event data derives from NJDOT's OpenReach system, which is their ATMS and 511 platform. This data is entered by operations personnel into their ATMS platform in real-time. Types of information recorded include the type of event (e.g., construction, accident, special event, disabled vehicle, police activity, etc.), location, and lane closure information.

The weather data is a mixture of both weather alerts and radar that derives from freely available National Weather Service data feeds.

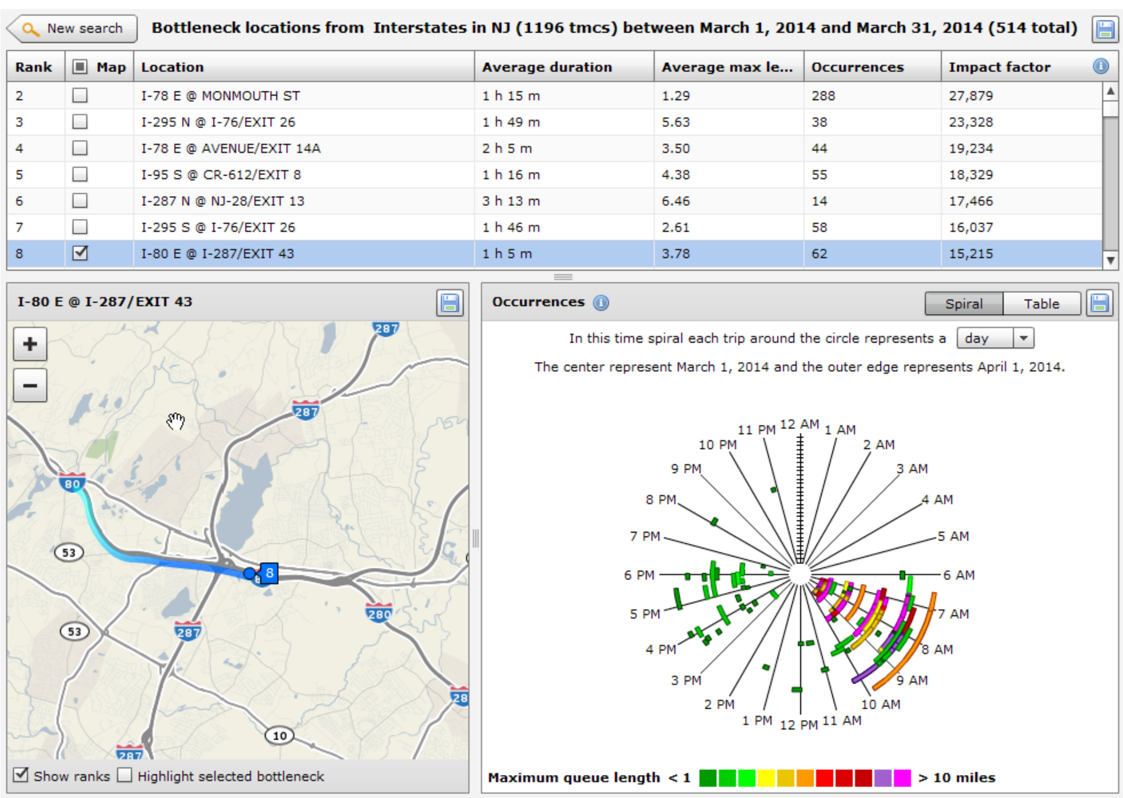
All of the above-mentioned data sources are fed into the Regional Integrated Transportation Information System (RITIS) archive and the VPP suite of analytics tools housed outside of the agency. RITIS also began as a project under the I-95 Corridor Coalition and is an automated data fusion and dissemination system with three main components, including: (1) real-time data feeds, (2) real-time situational awareness tools, and (3) archived data analysis tools.

22 I-95 Corridor Coalition, Vehicle Probe Project web site. Available at: http://i95coalition.org/projects/vehicle-probe-project/.

Agency Process

NJDOT identifies and analyzes congestion issues on the system and develops communication materials using the following basic process with the VPP suite of analytics tools:

1. Rank the worst congested (Figure 19) locations within the State by:
   a. Road class.
   b. Time of day.
   c. Day of week.
   d. Date range.
   e. Geography (statewide, by county, by zip code, by corridor, or by user-defined regions).
2. Determine if the congestion is recurring or non-recurring (Figure 20).
3. If non-recurring, identify causality through event data analysis (Figure 21), including:
   a. Weather events.
   b. Accidents.
   c. Construction.
4. For each congested location or corridor, study the economic impact of that congestion through user delay cost analysis (Figures 22 and 23).
5. Develop compelling reports and graphics that are presented to high-level decision makers to communicate the impact of the problem in a meaningful way (Figures 24 - 27).

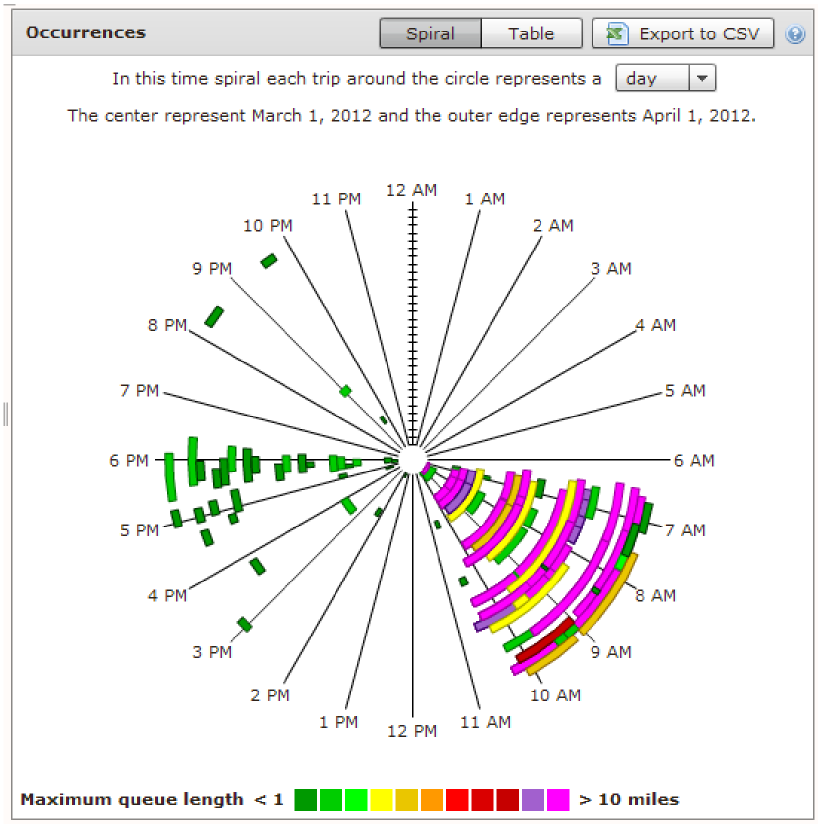


Source: Bottleneck Ranking Tool from the Probe Data Analytics Suite. University of Maryland CATT Laboratory.

Figure 19. Screenshot. Ranked bottleneck locations for all of New Jersey during March 2014.

The bottleneck ranking tool allows an agency to “rank” the worst congested locations in any geographic region (multi-State, State, county(ies), city, etc.) based on archived data (e.g., probe-based speed data, weather data, and ATMS incident logs), which informs users on certain statistics (e.g., on how significant each congested location is, when the congestion occurs, how it may have been affected by incidents and events, etc.).

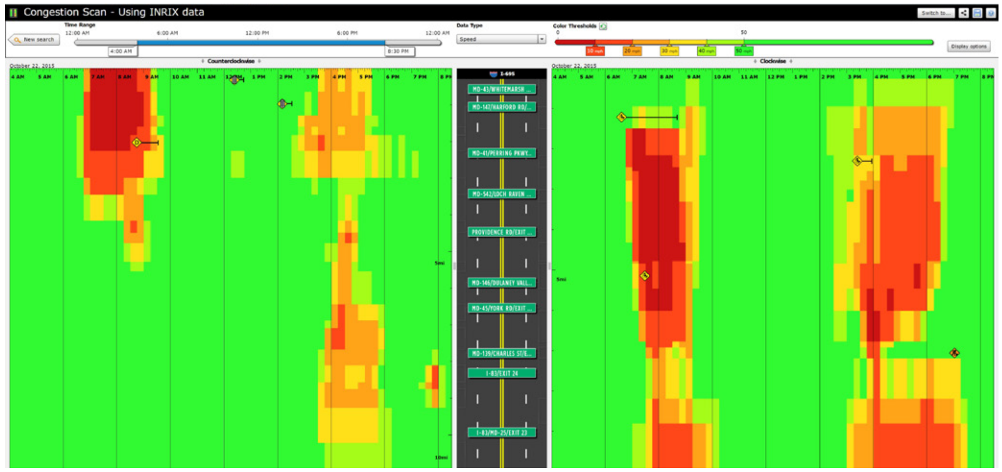
The graph in Figure 20 is one of the ways NJDOT views information on bottlenecks using the VPP suite. This graph shows the Monday through Friday pattern of congestion in the mornings (6:30AM through 10:30AM) with little to no congestion on the weekends. The spiral graphs are useful in identifying temporal patterns in the data. The repeated patterns that temporal data exhibit over longer time spans are often more illuminating than the linearity of the data and can be used in tandem with event and weather data to determine if the bottleneck location is recurring or non-recurring. The spiral graph renders data along a temporal axis that spirals outward at regular, user-defined intervals (e.g., days, weeks, or months). Bottleneck events are rendered as bands along the axis, creating visual clusters among data that contribute to patterns. The particular bottleneck shown in Figure 20 is rendered on a daily cycle—each revolution in the graph is a single day where the inner-most revolution is March 1, and the outermost revolution is March 31. Bottlenecks at this location are more prevalent in the morning between the hours of 6:30AM and 10AM, with a few less-severe bottlenecks occurring in the evening between 5PM and 6PM. The color denotes the maximum length of the bottleneck. The 5-day work week pattern also is shown in this bottleneck.



Source: Bottleneck Ranking Tool from the Probe Data Analytics Suite. University of Maryland CATT Laboratory.

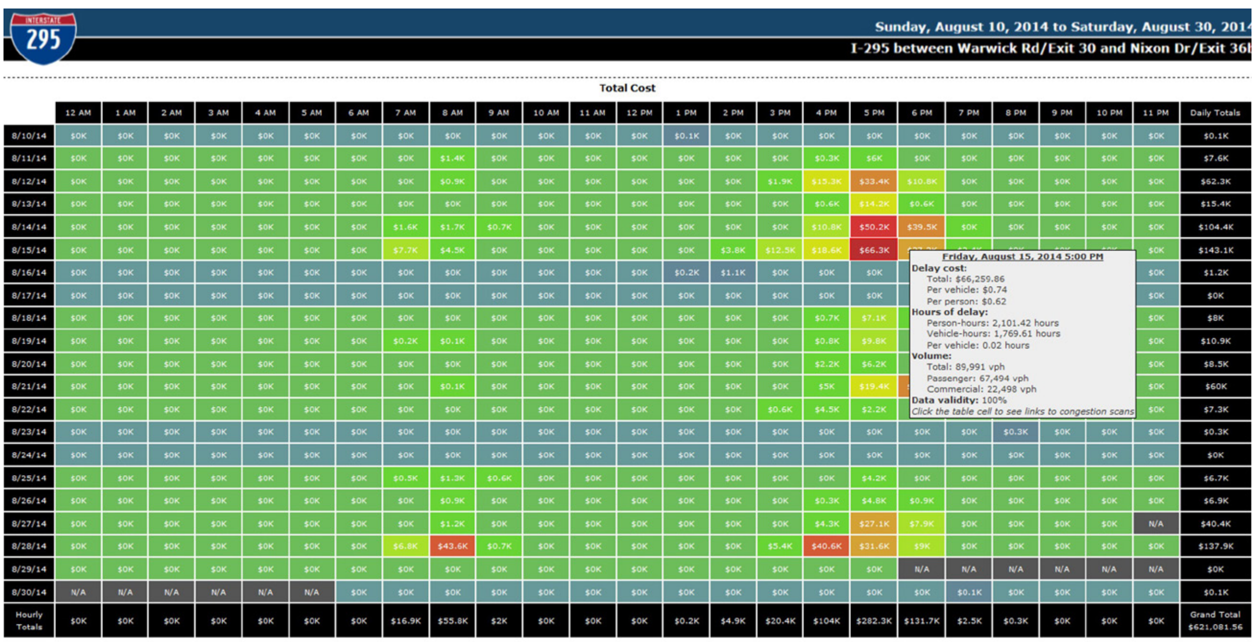
Figure 20. Chart. Time spiral showing when a bottleneck occurred and length during that time.

The graph in Figure 21 is another way that identifies congestion issues based on archived data. The graphic depicts a 20-mile section of I-95 in the State of Maryland in which the right side represents northbound traffic while the left side represents southbound traffic. Green shading denotes uncongested conditions, while red and orange shadings represent more congested conditions. Daily occurrences of incidents and events are overlayed. Figure 21 shows a large number of events that contributed to abnormal (non-recurring) congestion in the northbound direction of travel during the morning.



Source: Bottleneck Ranking Tool from the Probe Data Analytics Suite. University of Maryland CATT Laboratory.

Figure 21. Graph. Congestion scan graphic for I-95.



295

Sunday, August 10, 2014 to Saturday, August 30, 2014

I-295 between Warwick Rd/Exit 30 and Nixon Dr/Exit 36

Total Cost

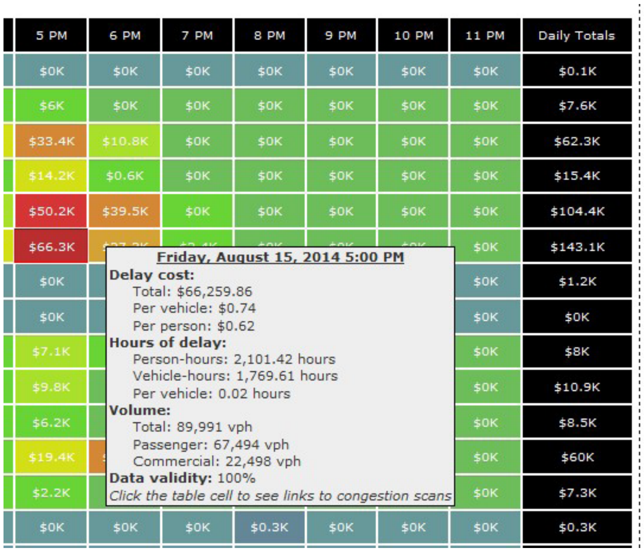
	12 AM	1 AM	2 AM	3 AM	4 AM	5 AM	6 AM	7 AM	8 AM	9 AM	10 AM	11 AM	12 PM	1 PM	2 PM	3 PM	4 PM	5 PM	6 PM	7 PM	8 PM	9 PM	10 PM	11 PM	Daily Totals
8/10/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.1K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.1K
8/11/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$1.4K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.3K	$6K	$0K	$0K	$0K	$0K	$0K	$0K	$7.6K
8/12/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.9K	$0K	$0K	$0K	$0K	$0K	$0K	$1.9K	$15.3K	$33.4K	$10.8K	$0K	$0K	$0K	$0K	$0K	$62.3K
8/13/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.6K	$14.2K	$0.6K	$0K	$0K	$0K	$0K	$0K	$15.4K
8/14/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$1.6K	$1.7K	$0.7K	$0K	$0K	$0K	$0K	$0K	$0K	$10.8K	$50.2K	$39.5K	$0K	$0K	$0K	$0K	$0K	$104.4K
8/15/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$7.7K	$4.5K	$0K	$0K	$0K	$0K	$0K	$3.8K	$12.5K	$18.6K	$66.3K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0K	$143.1K
8/16/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.2K	$1.1K	$0K	$0K	$0K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0K	$1.2K
8/17/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0K	$0K
8/18/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.7K	$7.1K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0K	$8K
8/19/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.2K	$0.1K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.8K	$9.8K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0K	$10.9K
8/20/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$2.2K	$6.2K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0K	$8.5K
8/21/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.1K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$5K	$19.4K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0K	$60K
8/22/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.6K	$4.5K	$2.2K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0K	$7.3K
8/23/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.3K	$0K	$0K	$0K	$0.3K
8/24/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K
8/25/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.5K	$1.3K	$0.6K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$4.2K	$0K	$0K	$0K	$0K	$0K	$0K	$6.7K
8/26/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.9K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.3K	$4.8K	$0.9K	$0K	$0K	$0K	$0K	$0K	$6.9K
8/27/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$1.2K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$4.3K	$27.1K	$7.9K	$0K	$0K	$0K	$0K	N/A	$40.4K
8/28/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$6.8K	$43.6K	$0.7K	$0K	$0K	$0K	$0K	$0K	$5.4K	$40.6K	$31.6K	$9K	$0K	$0K	$0K	$0K	$0K	$137.9K
8/29/14	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	N/A	N/A	N/A	N/A	N/A	N/A	$0K
8/30/14	N/A	N/A	N/A	N/A	N/A	N/A	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$0.1K	$0K	$0K	$0K	$0K	$0.1K
Hourly Totals	$0K	$0K	$0K	$0K	$0K	$0K	$0K	$16.9K	$55.8K	$2K	$0K	$0K	$0K	$0.2K	$4.9K	$20.4K	$104K	$282.3K	$131.7K	$2.5K	$0.3K	$0K	$0K	$0K	Grand Total $621,081.56

Friday, August 15, 2014 5:00 PM
Delay cost:
Total: $66,259.86
Per vehicle: $0.74
Per person: $0.62
Hours of delay:
Person-hours: 2,101.42 hours
Vehicle-hours: 1,769.61 hours
Per vehicle: 0.02 hours
Volume:
Total: 89,991 vph
Passenger: 67,494 vph
Commercial: 22,498 vph
Data validity: 100%
Click the table cell to see links to congestion scans

Source: User Delay Cost Table from the Probe Data Analytics Suite. University of Maryland CATT Laboratory.

Figure 22. Screenshot. User delay costs for a 17-mile stretch of I-295 in New Jersey (both directions of travel).

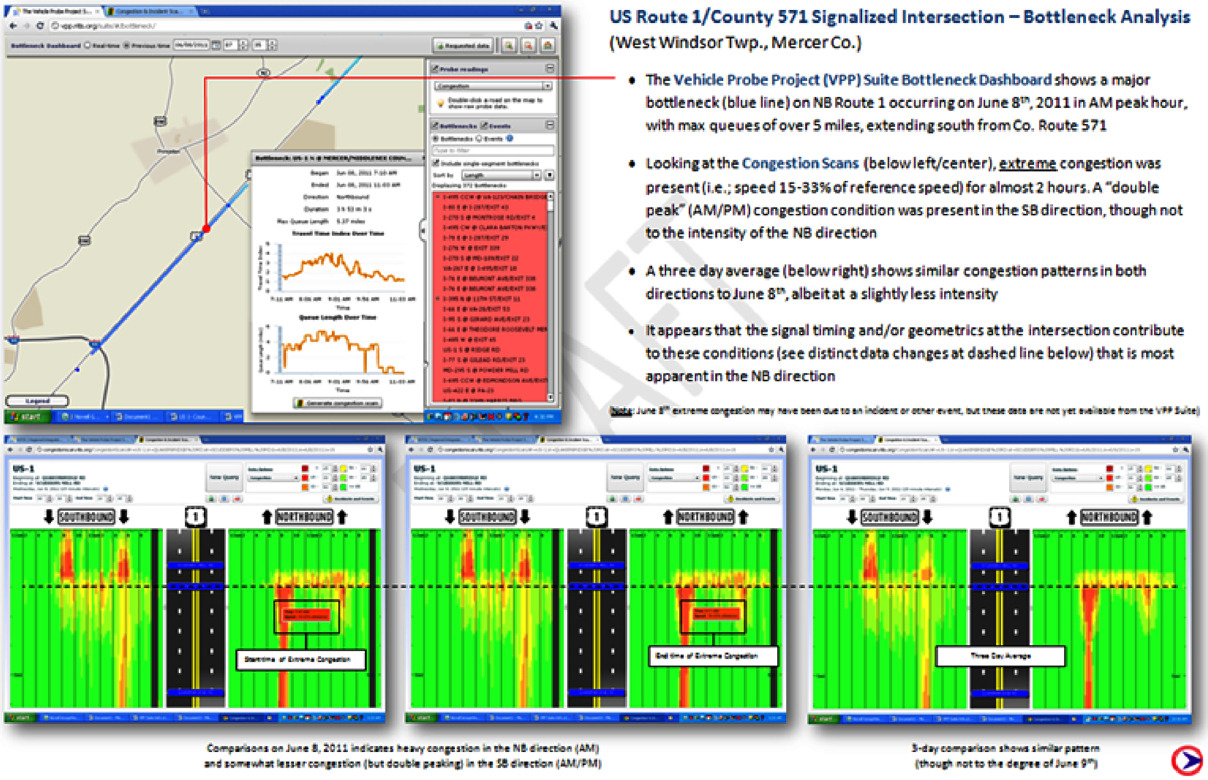
Figure 22 displays the user delay costs for I-295 in New Jersey that contribute to an economic impact study of congestion. Each cell represents 1 hour of the day on a particular day of the week. The cells are color coded to represent hourly user delay in thousands of dollars. Figure 23 provides a closer look at a summary report of delay costs for a Friday at 5 PM.



Source: Probe Data Analytics Suite. University of Maryland CATT Laboratory.

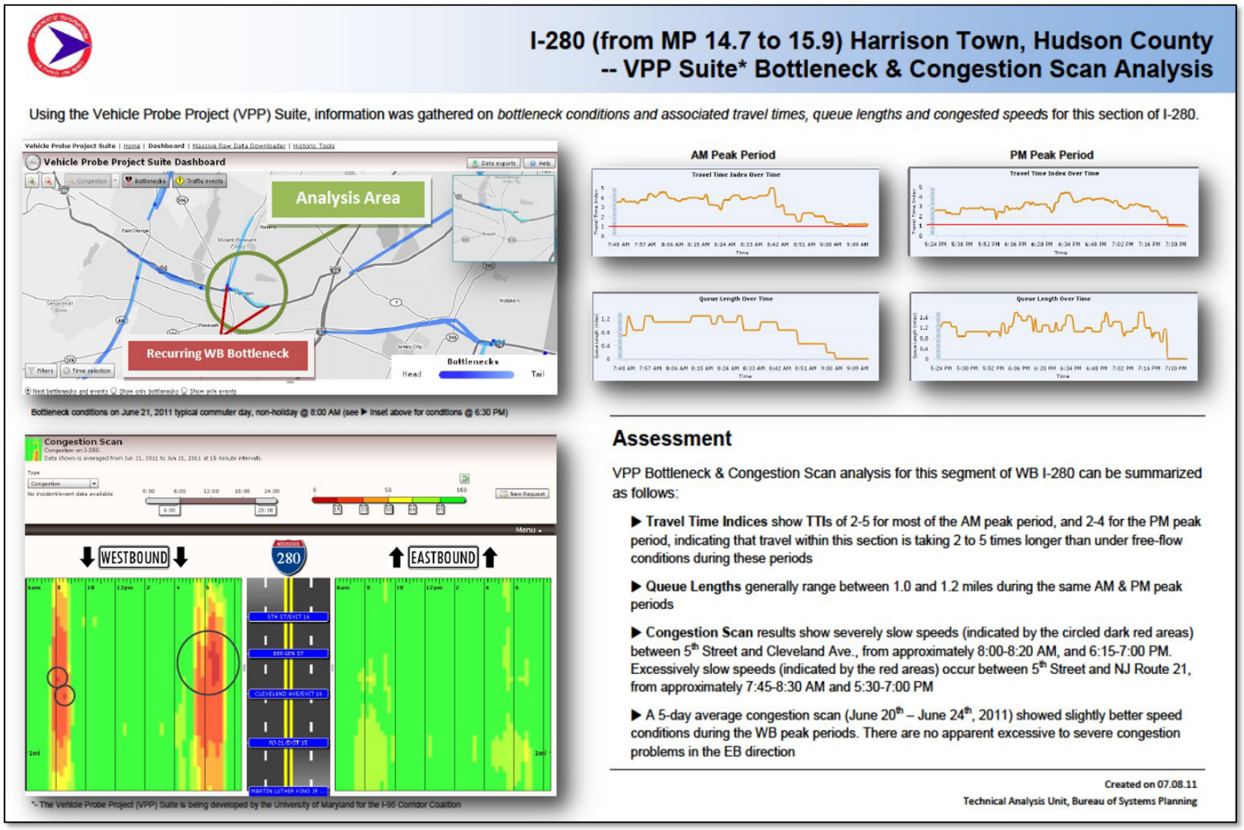
Figure 23. Screenshot. View of I-295 showing detailed statistics in a mouse tooltip for a particular hour of the day.

The following figures are examples of reports and facts sheets that NJDOT developed based on archived operations data to communicate to decision makers and the public about the problems on the system and the need for improvements.



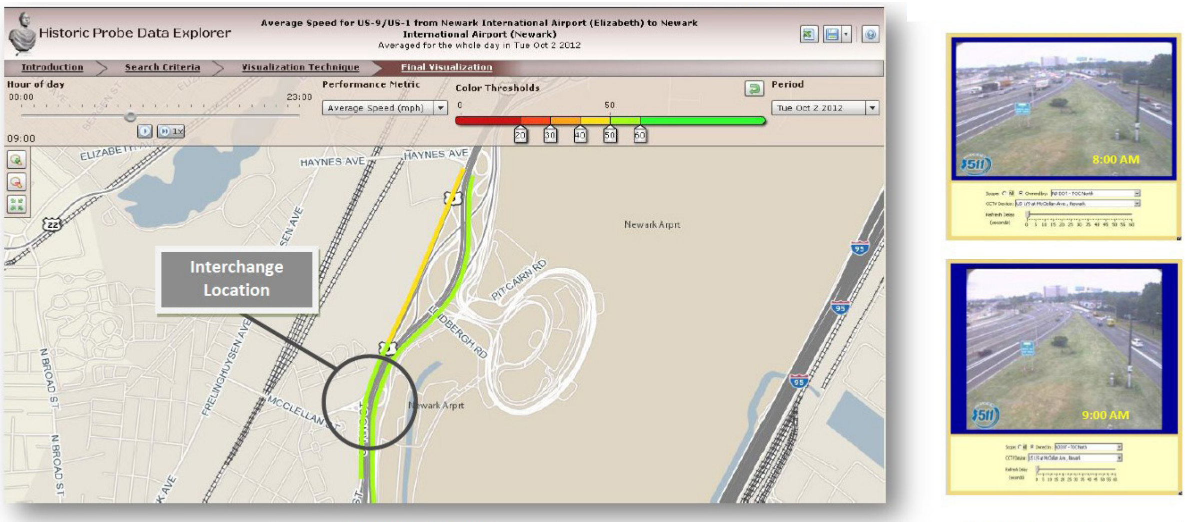
Source: New Jersey Department of Transportation.

Figure 24. Screenshot. Project confirmation presentation graphic to confirm a "high-need" signalized intersection in New Jersey.



Source: New Jersey Department of Transportation.

Figure 25. Screenshot. Example of problem identification and project confirmation at New Jersey Department of Transportation.



NAME	MILES	AM Peak Period (Southbound) (Average Speed per time period) 05:00	06:00	07:00	08:00	09:00
Newark International Airport (Newark)	0.098492	48	51	47	48	47
Newark International Airport (Newark)	0.60717	46	50	50	50	49
McClellan St	0.145905	51	53	52	50	51
McClellan St	0.106259	51	52	51	50	50
Newark International Airport (Elizabeth)	0.339471	53	54	54	52	53

NAME	MILES	AM Peak Period (Northbound) (Average Speed per time period) 05:00	06:00	07:00	08:00	09:00
Newark International Airport (Elizabeth)	0.092029	51	53	51	50	50
McClellan St	0.376631	54	56	55	54	57
McClellan St	0.170077	55	55	55	54	56
Newark International Airport (Newark)	0.176664	54	54	54	53	56
Newark International Airport (Newark)	0.146899	52	52	53	53	55

Source: New Jersey Department of Transportation.

Figure 26. Screenshot. Example analysis from New Jersey Department of Transportation using the Vehicle Probe Project Suite, 511 New Jersey cameras, and the New Jersey Consortium for Middle Schools.

The evaluation in Figure 26 of the McClellan Street Interchange is part of the Port Authority of New York and New Jersey's Southern Access Roadway Project. This location is not considered a high priority for the NJDOT from a congestion relief perspective.

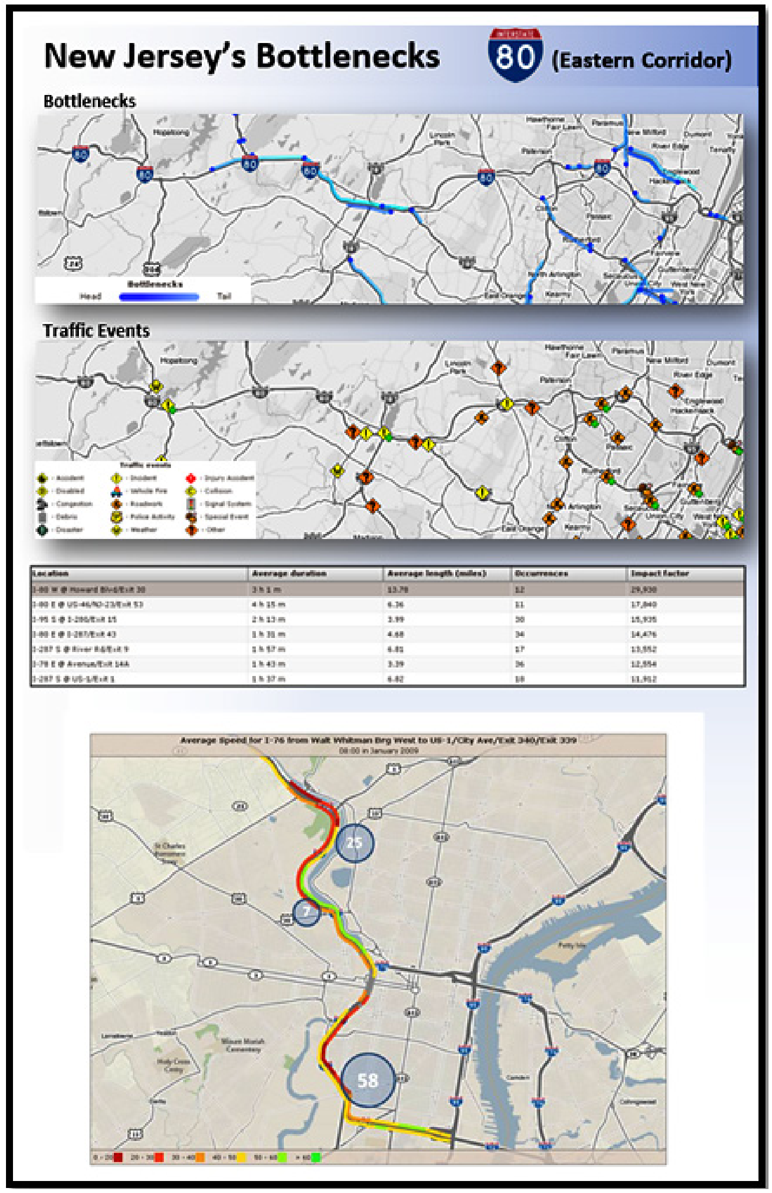
Tool

The VPP suite allows the user to generate a congestion impact factor, which is a simple multiplication of the average duration, average maximum length, and the number of occurrences. The impact factor helps the user filter out high occurrence—but short duration and short length—bottlenecks that are of lower significance (e.g., those that might be a result of normal stoppages as a result of traffic signals on arterials).

Once NJDOT identifies a candidate location for congestion reduction, it can conduct an economic analysis within the VPP suite. NJDOT can select a study region and use existing statistics on the regional value of time for both commercial and passenger vehicles. The suite produces a series of interactive tables that show hourly, daily, and total:

- User delay cost:
    - Total.
    - Per vehicle.
    - Per person.
- Hours of delay:
    - Person-hours.
    - Vehicle-hours.
    - Per vehicle.
- Volume:
    - Total.
    - Passenger.
    - Commercial.
- Data validity metrics.

After conducting their analysis within the VPP suite, NJDOT planners integrate some of the supporting graphics into slide shows, pamphlets, and posters that are used to help communicate with decision makers and the public. The embedded graphics are supported by explanatory text, call-out boxes, and other materials that combine to display all the information about the nature of the identified problem. Figures 24-27 are examples of some of these problem identification slides and pamphlets. It should be noted that many of these examples



Source: New Jersey Department of Transportation.

Figure 27. Screenshot. Concept graphics explaining where high accident locations contribute to bottlenecks.

are taken from larger slide decks and other materials that, when combined, provide further background and backup materials to help explain the nature of a problem, and in some cases, recommended countermeasures.

Calculations Behind the Visualizations

The calculations behind all of the performance measures within the VPP suite are posted online.[23] It is important to note that, as knowledge improves and technical capabilities mature, these calculations will frequently be updated and revamped within the suite to reflect the state of the art. One of the many advantages of using a web-based analytics tool is that the enhancements made to these calculations propagate instantaneously to every user, thus NJDOT and the relevant metropolitan planning organizations (MPOs) benefit from the new knowledge, and analysts are not burdened with having to re-implement complex functions multiple times within an organization. Another advantage is that these calculations are instantaneously and effectively standardized across multiple agencies and multiple departments, which makes it much easier to compare performance and ensure consistent and reliable analysis from State to State.

Impact and Lessons Learned

NJDOT has, within only a few short years, significantly increased the abilities of their own planning staff, their consultant partners, researchers, and MPOs through the adoption of these archived data analytics tools. They are saving time and money while providing much-needed insights into mobility and the problem of congestion.

With the advancements of these archived data tools, NJDOT can now answer difficult questions that were either previously unanswerable, or could have taken up to 1 year to analyze and report. As a result of this tool, NJDOT is able to operate more efficiently, respond to media and public inquiries rapidly, use a data-driven approach to problem solving, and focus more time and effort in identifying solutions to complex transportation problems, rather than chasing down other data.

> "Archived operations data may be the most important advance in data for planning purposes."
>
> **John Allen, Formerly of the NJDOT Planning Office**

The improved analysis tools provide NJDOT with a stronger basis for the recommendations, and enable staff to bring a fresh perspective to previous studies, as appropriate. The VPP suite also is being used to evaluate projects in concept development to help decide whether projects should continue to advance through NJDOT's Project Delivery Process. This type of evaluation is used primarily for projects that were initiated outside of the normal problem statement or management systems screening processes and can be particularly helpful in ranking the priority of projects throughout a region so that political influences do not overly impact a project's likelihood of moving forward.

The greatest lesson learned from NJDOT's experience with archived operations data is that data alone will not solve its problems. It needs tools that can empower a wide range of planners and analysts by removing the stress and burden of managing massive data sets and relying on external processing power to make them more efficient and proactive analysts.

23 Go to https://vpp.ritis.org/suite/faq/#/how-are-bottleneck-conditions-tracked for information on bottlenecks, and https://vpp.ritis.org/suite/faq/#/calculations for information on user delay cost.

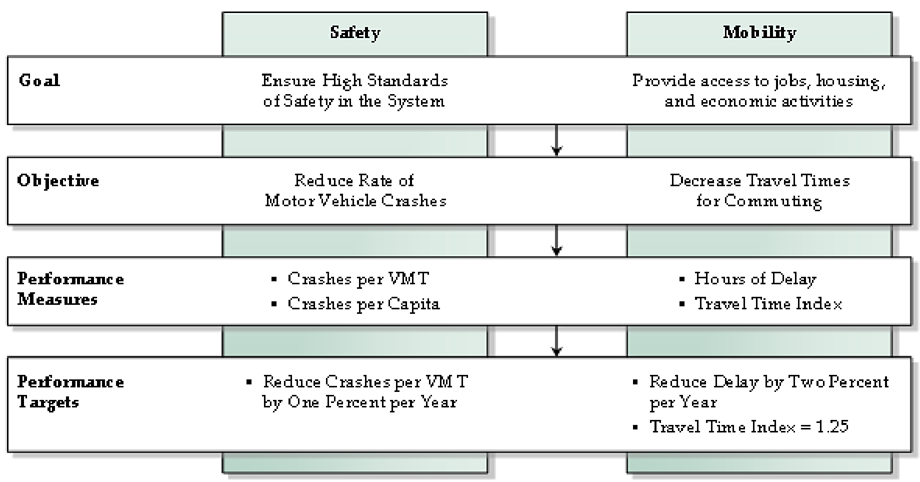
Development and Reporting of Mobility Performance Measures

Need

Performance measurement for mobility concerns is a key aspect of planning activities. It has been defined as a major step in the congestion management process (CMP) and is the basis for conducting sound performance management. In addition to fulfilling these requirements, undertaking performance measurement improves the quality of planning in several ways:

- **Provides accountability for investment decisions.** Performance measures assist transportation professionals in gauging what benefit was gained for the cost. As such, they enable the investment process to be transparent.
- **Demonstrates that planners understand where the problems are.** Instead of relying on partial information from models, planners now can identify exactly where problems exist. Therefore, performance measurement can be an effective tool for public relations; it is the first step in developing solutions to problems.
- **Helps to guide investment decisions and to understand the impacts of those decisions on travelers.** As with the private sector, performance measures help agencies and firms improve their products and labor by continuously monitoring impacts of policies and investments.

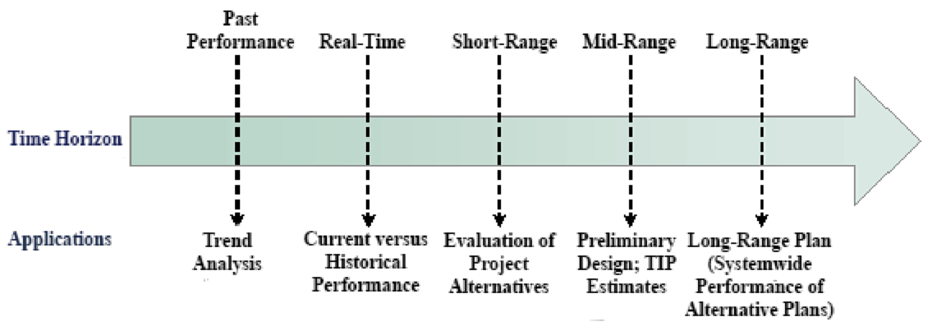
Performance measures should not be undertaken in isolation. Rather, they should be part of a process that relates to the strategic vision an agency has. That vision is embodied in goals and objectives and in performance measures that are used to monitor progress toward meeting them (Figure 28). Performance measures also should permeate all aspects of the planning process and planning products (Figure 29).[24] A high degree of commonality in measures should exist across planning stages, recognizing some measures will be unique to specific stages. To the degree possible, however, consistency in performance measures across the different stages should be maintained. Documentation of past performance, through the production of performance reports, is the focus of this section.



Source: Transportation Research Board, National Cooperative Highway Research Program.

Figure 28.Diagram. Performance measures provide a quantifiable means of implementing goals and objectives from the transportation planning process.

24 Cambridge Systematics et al., *Guide to Effective Freeway Performance Measurement: Final Report and Guidebook*, Web-Only Document 97, Transportation Research Board, August 2006, Available at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w97.pdf.



Source: Transportation Research Board, National Cooperative Highway Research Program.

Figure 29. Diagram. Performance measures should be carried across planning applications throughout the entire time horizon.

Archived data enables performance measures to track trends in system performance and to identify deficiencies and needs. Performance measures can be classified into different types:

- **Outcome measures** relate to how well the firm or agency is meeting its mission and stated goals. Outcome measures are sometimes called "efficiency" measures. In the private sector, outcome measures relate to the "bottom-line" (i.e., the financial viability of the firm: profit and revenue). For transportation agencies, outcomes are more related to the nature and extent of the services provided to transportation users.
- **Output measures** relate to the levels of effort expended, or the scale or scope of actions. Output measures are often called "activity measures" because they relate to specific activities undertaken by agencies.
- **Input measures** relate to the physical quantities of items used to undertake activities.

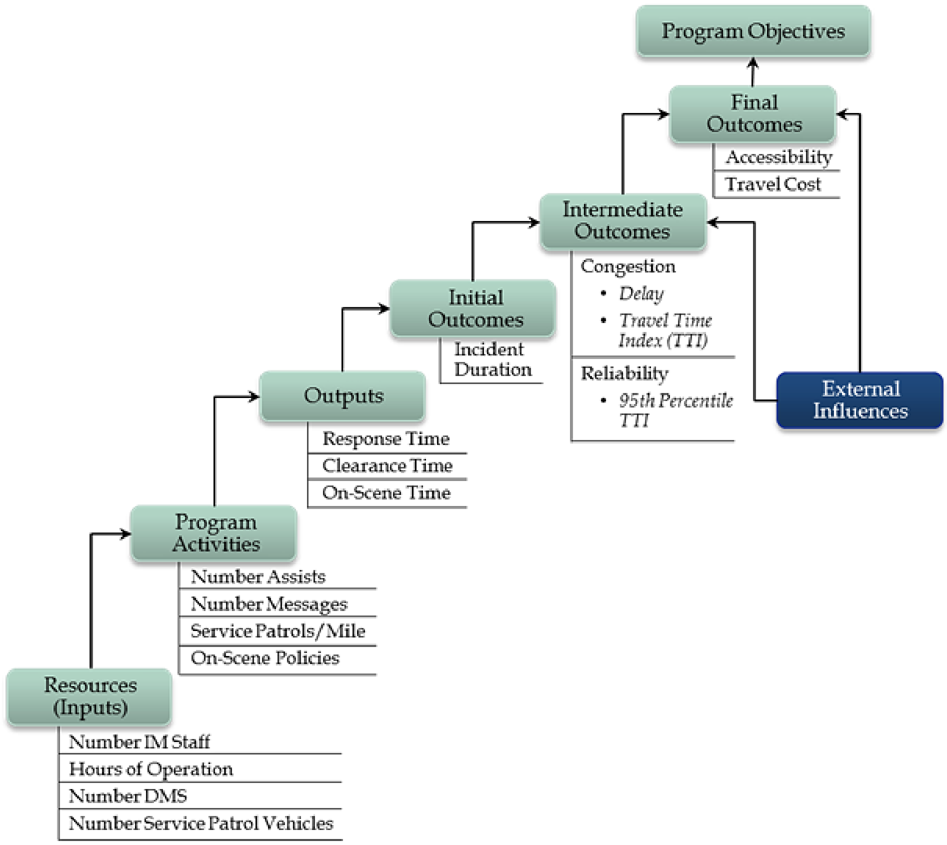
Figure 30 shows an example of a "program logic model" adapted from the general performance measurement literature for incident management. The model breaks down the three above categories (outcome, output, and input measures) into additional categories. "External influences" relate to the forces that create changes in demand that are outside the control of transportation agencies (i.e., changes in economic activity and demographic and migration trends). The main point of Figure 30 is that a comprehensive performance management program includes performance measurement at all levels and that a clear linkage exists between successively higher levels. That is, a change in a level leads to a change in the higher levels.

Performance measures for both congestion and reliability should be based on measurements of travel time. Travel times are easily understood by practitioners and the public, and are applicable to both the user and facility perspectives of performance.

Reliability is a key aspect of mobility that should be measured. There are two widely held ways that reliability can be defined. Each is valid and leads to a set of reliability performance measures that capture the nature of travel time reliability. Reliability can be defined as:

- The variability in travel times that occurs on a facility or for a trip over the course of time.
- The number of times (trips) that either "fail" or "succeed" in accordance with a predetermined performance standard or schedule.

In both cases, reliability (more appropriately, unreliability) is caused by the interaction of the factors that influence travel times: fluctuations in demand (which may be due to daily or seasonal variation, or by special events), traffic control devices, traffic incidents, inclement weather, work zones, and physical capacity (based on prevailing geometrics and traffic patterns). These factors produce travel times that are different from day to day for the same trip.



Source: Federal Highway Administration, Operations Performance Measurement Workshop Presentation.

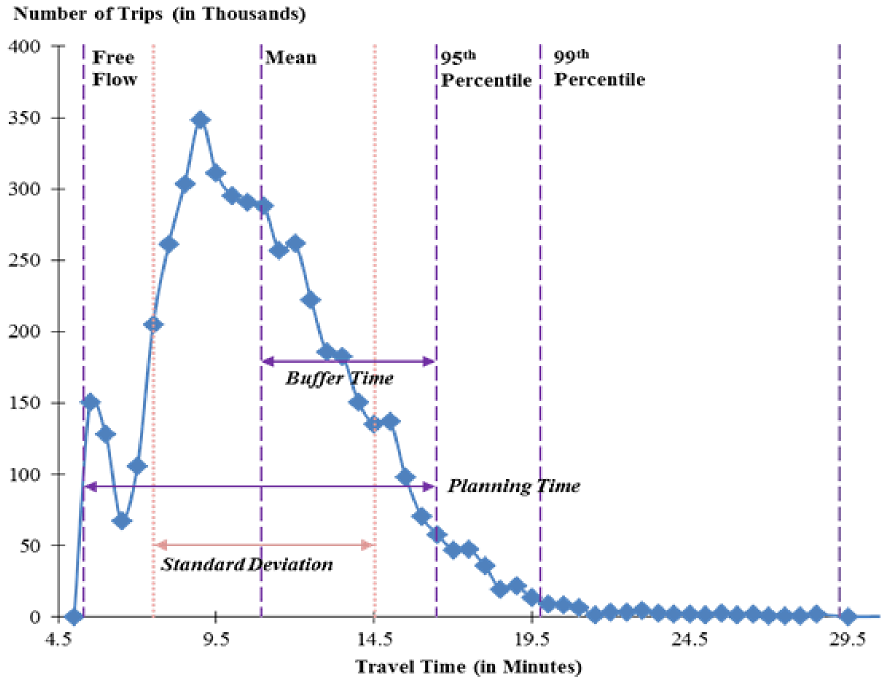
Figure 30. Diagram. Program logic model adapted for incident management.

From a measurement perspective, travel time reliability is quantified from the distribution of travel times, for a given facility/trip and time period (e.g., weekday peak period), that occurs over a significant span of time; 1 year is generally long enough to capture nearly all of the variability caused by disruptions. Figure 31 shows an actual travel time distribution derived from roadway detector data, and how it can be used to define reliability metrics.[25] The shape of the distribution in Figure 31 is typical of what is found on congested freeways; it is skewed toward higher travel times. The skew is reflective of the impacts of disruptions, such as incidents, weather, work zones, and high demand on traffic flow. Therefore, most of the useful metrics for reliability are focused on the right half of the distribution; this is the region of interest for reliability. Note that a number of metrics are expressed relative to the free-flow travel time, travel time under low traffic-flow conditions, which becomes the benchmark for any reliability analysis. The degree of unreliability then becomes a relative comparison to the free-flow travel time.[26] Table 7 shows a suite of reliability metrics. All of these are based on the travel time distribution.[27]

25 Kittelson & Associates et al., *Incorporation of Travel Time Reliability into the Highway Capacity Manual*, SHRP 2 Report S2-L08-RW-1,TRB, 2014. Available at: http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_S2-L08-RW-1.pdf.

26 Ibid.

27 Ibid.



Source: Transportation Research Board, Strategic Highway Research Program 2, Report S2-L08-RW-1.

Figure 31. Graph. Travel time distribution is the basis for defining reliability metrics.[28]

Table 7. Recommended set of reliability performance measures from Strategic Highway Research Program 2 Project L08.[29]

Reliability Performance Measure	Definition
Core Measures	
Planning time index (PTI)	95th percentile travel time index (TTI) (95th percentile travel time divided by the free-flow travel time).
80th percentile TTI	80th percentile TTI (80th percentile travel time divided by the free-flow travel time).
Semi-standard deviation	The standard deviation of travel time pegged to the free flow rather than the mean travel time.
Failure/On-time measures	Percent of trips with space mean speed less than 50 miles per hour (mph), 45 mph, and 30 mph.
Supplemental Measures	
Standard deviation	Usual statistical definition.
Misery index (modified)	The average of the highest 5 percent of travel times divided by the free-flow travel time.

Source: Transportation Research Board, Strategic Highway Research Program 2, Report S2-L08-RW-1.

28 From SHRP 2 Report SR-L08-RW-1: Incorporation of Travel Time Reliability into the Highway Capacity Manual, Figure ES.2, p. 3. Copyright, National Academy of Sciences, Washington, D.C., 2014. Reproduced with permission of the Transportation Research Board. Available at: http://www.trb.org/Main/Blurbs/169594.aspx.

29 From SHRP 2 Report SR-L08-RW-1: Incorporation of Travel Time Reliability into the Highway Capacity Manual, Table 2.1, p. 17. Copyright, National Academy of Sciences, Washington, D.C., 2014. Reproduced with permission of the Transportation Research Board.

Relevant Data

Prior to the appearance of archived operations data, congestion performance report production by planning agencies was rare. Where they existed, they were based on results from travel demand forecasting (TDF) models or small sample "floating car" travel time runs. The measurement of reliability was non-existent.

The data to accomplish performance reporting starts with field measurements of travel times or speeds from intelligent transportation system (ITS) roadway detectors or vehicle probes. At a basic level, this data can be used to produce mobility performance measures. However, more sophisticated and in-depth measures can be constructed with the integration of other data:

- Traffic volumes (volumes are paired with speed measurements when ITS detectors are used but there are no companion volume measurements from the current generation of vehicle probe data).
- Incident data from incident management logs (includes the nature, severity, duration, and blockage characteristics of incidents).
- Work zone data (similar to incident data).
- Weather data.

Analytical Procedures

Data preparation. Routine quality control (QC) procedures need to be applied to the data. These are documented elsewhere.[30] Travel time data is extremely voluminous and requires special software to be processed. Incident data is much more manageable and can be analyzed with spreadsheets.

Calculation methods. Detailed guidance on processing the raw travel time data into measurements, including QC, may be found in the literature.[31] In their simplest form, the steps are as follows:

1. Compute travel time for individual links (if not already provided). Compute these by the time periods of interest (e.g., PM peak period).
2. Compute volume-weighted moments of the travel time distribution for each link (e.g., mean and various percentiles). Each observation in the distribution should be an individual travel time record. This assumes that volumes are available at the same temporal and spatial level as the travel time measurements.[32] If volumes are not available, then a straight average is used. Note that an "observation" could in fact be an average value; for example, the average travel time for a 5-minute period. In such cases, some variability has been lost due the aggregation; if travel times from individual vehicles were used, the resulting reliability metrics would indicate higher variability.

30 Cambridge Systematics, et al., *Guide to Effective Freeway Performance Measurement: Final Report and Guidebook*, Web-Only Document 97, Transportation Research Board, August 2006. Available at: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w97.pdf.

31 For example: Cambridge Systematics et al., *Guide to Effective Freeway Performance Measurement: Final Report and Guidebook*, Web-Only Document 97, Transportation Research Board, August 2006, Available at http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w97.pdf. Also: Cambridge Systematics, Inc. et al., *Analytical Procedures for Determining the Impacts of Reliability Mitigation Strategies,* The Second Strategic Highway Research Program, Report S2-L03-RR-1, Transportation Research Board (Washington, DC: 2013).

32 The Texas Agricultural and Mechanical University's *Urban Mobility Report* demonstrates a method for deriving volumes to be used with vehicle probe data: http://mobility.tamu.edu/ums/report/.

3. Compute additional metrics, such as the travel time indices, at the link level.
4. Compute performance measures at aggregated spatial levels, such as facility and area wide, by taking the volume-weighted average of the metrics previously computed.

If results are desired at the facility level, a more statistically sound procedure is to compute the travel time over the entire facility by summing up the individual link travel times from Step 1 and using the aggregated travel times as the basis for travel time distribution. In this case, the weights become sum of the volumes for the links that comprise the facility. In both cases, the results achieved by volume-weighting will be different than not weighting. The difference will be more pronounced if both low volume and high volume periods are used in the calculation.

The calculation of certain performance measures that are defined by exposure requires special treatment. For example, delay is the number of vehicles or persons exposed to the delay time. The most accurate way of computing these measures is from paired volume-travel time measurements; i.e., measurements that were collected at the same point in time and space. The current generation of privately-supplied vehicle probe data does not contain volume estimates – only speed or travel time. Computing exposure-based measures from this data requires using volume data from another source, namely, average traffic counts. Users should be aware that mixing average traffic counts with vehicle probe data produces only crude estimates for exposure-based measures.

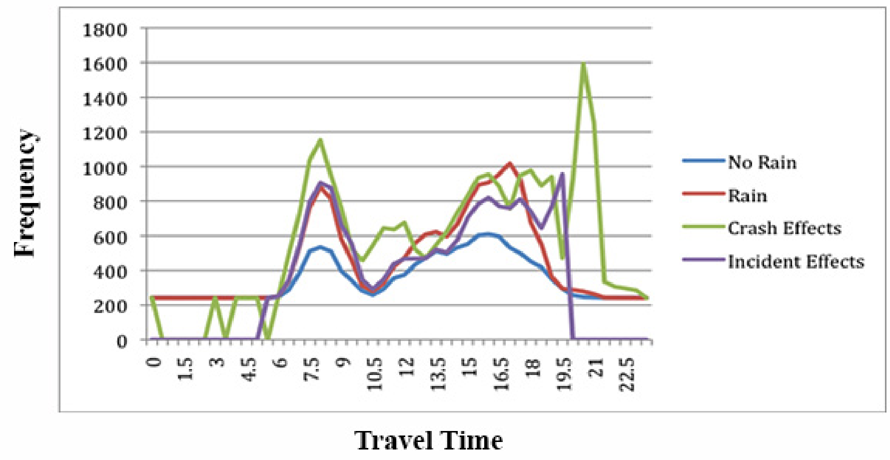
The analysis of incident, work zone, and weather characteristics data is much more straightforward; simple summaries usually suffice. Integrating incident and work zone data with travel time data is more problematic. First, the data must be matched temporally and spatially. The more difficult task is to assign congestion to a specific incident, work zone, or weather event. On facilities where recurring congestion rarely occurs, it is safe to assume that any congestion that occurs is due to an event. Where recurring congestion is routine, the assignment becomes much more difficult because some congestion would have occurred in the absence of the event.

Display of results. A variety of methods are used to display performance measures, including the display of trends:

- Figure 32 shows one example of how to combine travel time and event data into a single graphic for a corridor. This analysis shows the dramatically negative effect that disruptions can have on travel times when they occur.[33]
- Figure 33 shows how the Metropolitan Washington Council of Governments (MWCOG) depicts travel time reliability trends over time. This information is part of MWCOG's program to communicate problems to stakeholders and to be accountable for investment decisions. The data source for this analysis is vehicle probe data.[34]
- Figure 34 shows how the Georgia Department of Transportation (GDOT) displays trends in incident service times. This information is used to monitor the effectiveness of new incident management policies and practices. The data was collected by GDOT as part of their incident management system.

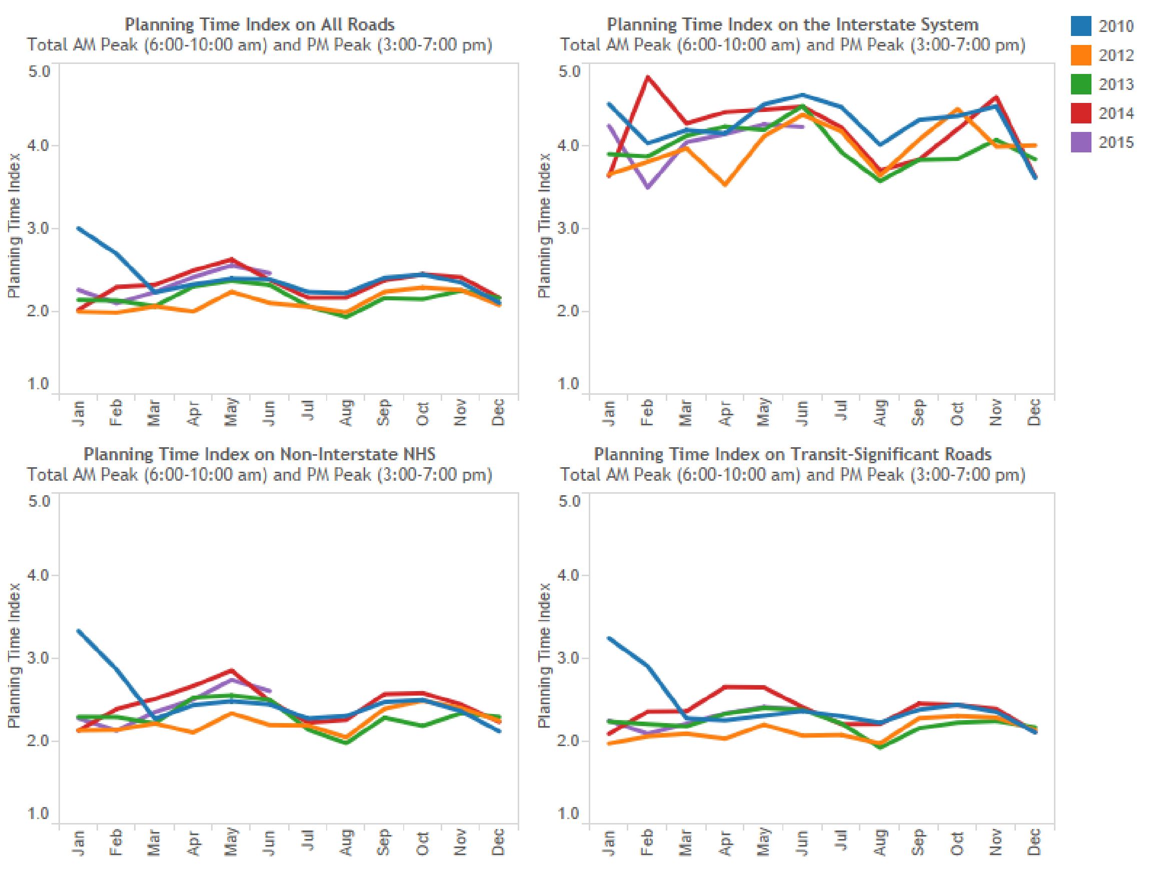
33 Cambridge Systematics et al., *Analytic Procedures for Determining the Impacts of Reliability Mitigation Strategies*, SHRP 2 Report S2-L03-RR-1, Transportation Research Board, 2013. Available at: http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_S2-L03-RR-1.pdf.

34 For more information, see the Metropolitan Washington Council of Governments (MWCOG) Transportation Congestion Dashboard at:https://www.mwcog.org/congestion/.



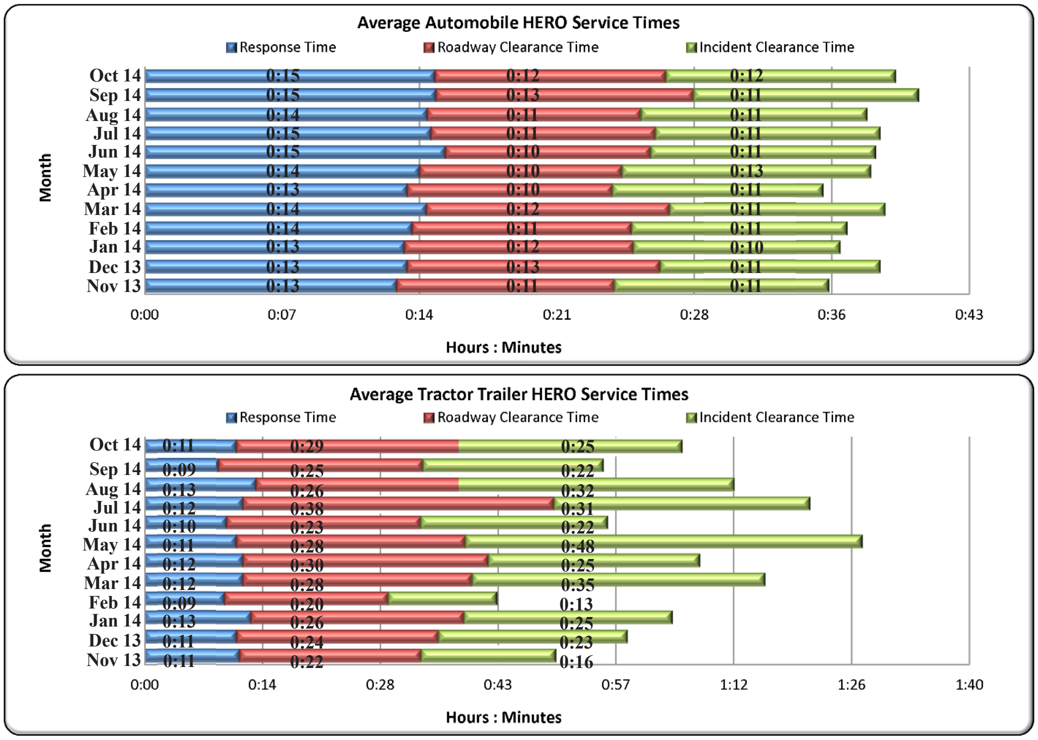
Source: Transportation Research Board, Strategic Highway Research Program 2, Report S2-L03-RR-1.

Figure 32. Graph. Mean travel times under rain, crash, or non-crash traffic incident conditions for I-5 southbound, North Seattle Corridor, Tuesdays through Thursdays, 2006.[35]



Source: Pu, Wenjing, Presentation at 5th International Transportation Systems Performance Measurement and Data Conference, June 1-2, 2015, Denver, Colorado.

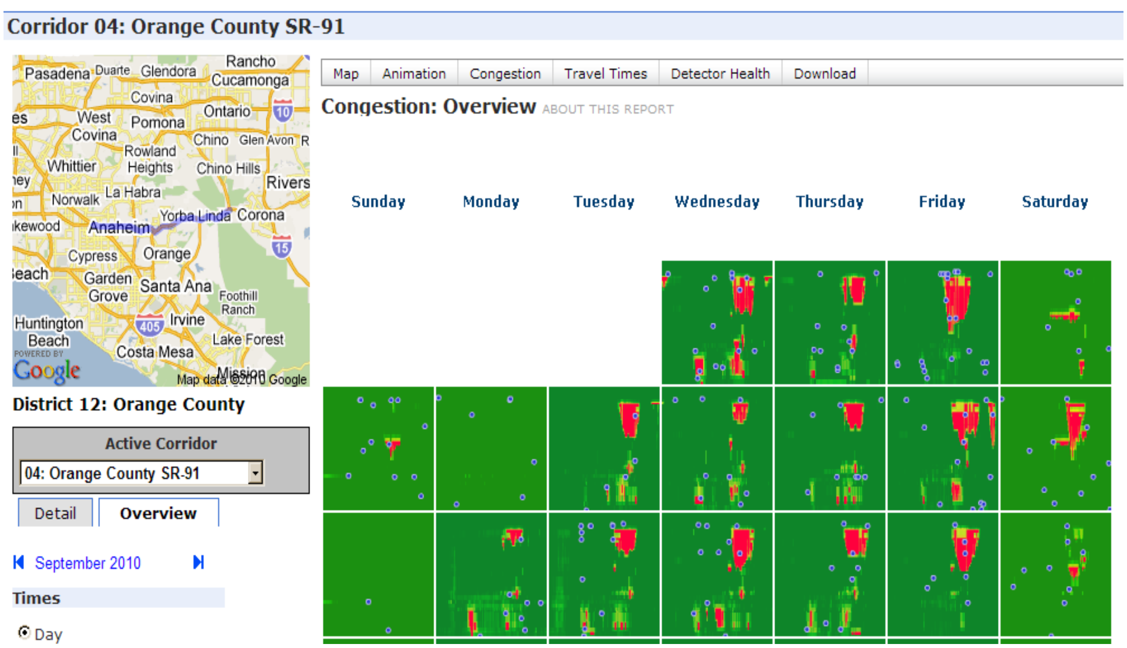
Figure 33. Graph. Reliability trends reported by the Metropolitan Washington Council of Governments.



Source: Georgia Department of Transportation, H.E.R.O Monthly Statistics Report, Available at:More information available at: http://www.dot.ga.gov/DS/Travel/HEROs.

Figure 34. Charts. Incident timeline trends reported by the Georgia Department of Transportation for the Atlanta region.

- Figure 35 shows how the Performance Measurement System (PeMS) displays congestion-by-source information for a corridor. PeMS is a data archival and diagnostic tool that is used by many agencies to identify the source of problems and to match countermeasures to them. The travel time data source for this particular analysis was ITS detectors.[35]



Source: California Department of Transportation

Figure 35. Screenshot. Heat map of congestion on a corridor using the Performance Measurement System.

35 California Department of Transportation, Performance Measurement System. Available at: http://pems.dot.ca.gov/.

Impact

In addition to demonstrating accountability, providing transparency, and promoting effective public relations, developing performance measures is the starting point for several other agency functions. These include:

- **Target setting –** Understanding the current state of the system is crucial to understanding what is achievable in the future. In addition, the same data used to measure current performance can be used as input to forecasting models.
- **Project and program evaluations –** Evaluating implemented projects is revealing in several ways. A library of impact factors can be developed for future project planning. Benefits can be highlighted as success stories, and lessons can be learned from evaluations where expected benefits did not materialize.

Case Study: Tracking Performance Measures at Delaware Valley Regional Planning Commission

Overview

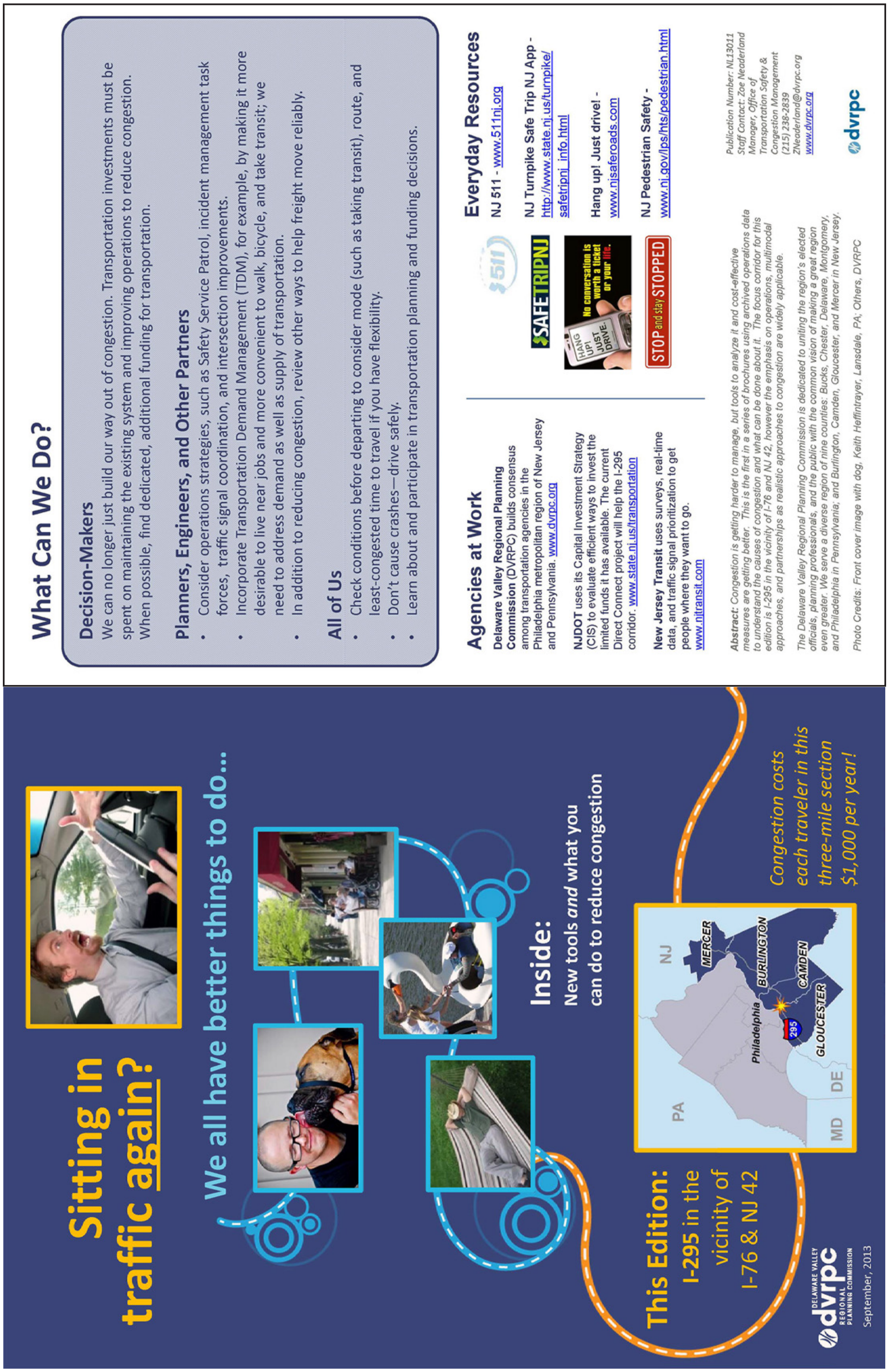
The Delaware Valley Regional Planning Commission (DVRPC) is the metropolitan planning organization for the greater Philadelphia region. The region includes five counties in Pennsylvania and four in New Jersey. The most congested freeway miles in the region are along major routes into and out of the City of Philadelphia. DVPRC uses archived operations data to inform a variety of planning efforts, including problem identification and confirmation, calibrating regional and corridor models, conducting before and after studies, and developing and reporting transportation performance measures for the regional congestion management process (CMP) and long-range transportation plan (LRTP).

Data

DVRPC primarily uses the VPP data purchased by the I-95 Corridor Coalition for developing performance measures based on archived operations data. It is currently exploring other vehicle probe-based data sources, including the National Performance Measures Research Data Set (NPMRDS), particularly for tracking and reporting progress toward the system performance measures. However, given DVRPC's concerns with the quality of the NPMRDS data, DVRPC is working with its partners at the Pennsylvania DOT and NJDOT to use the VPP data instead. Incident data from agency-collected traffic management systems also are analyzed.

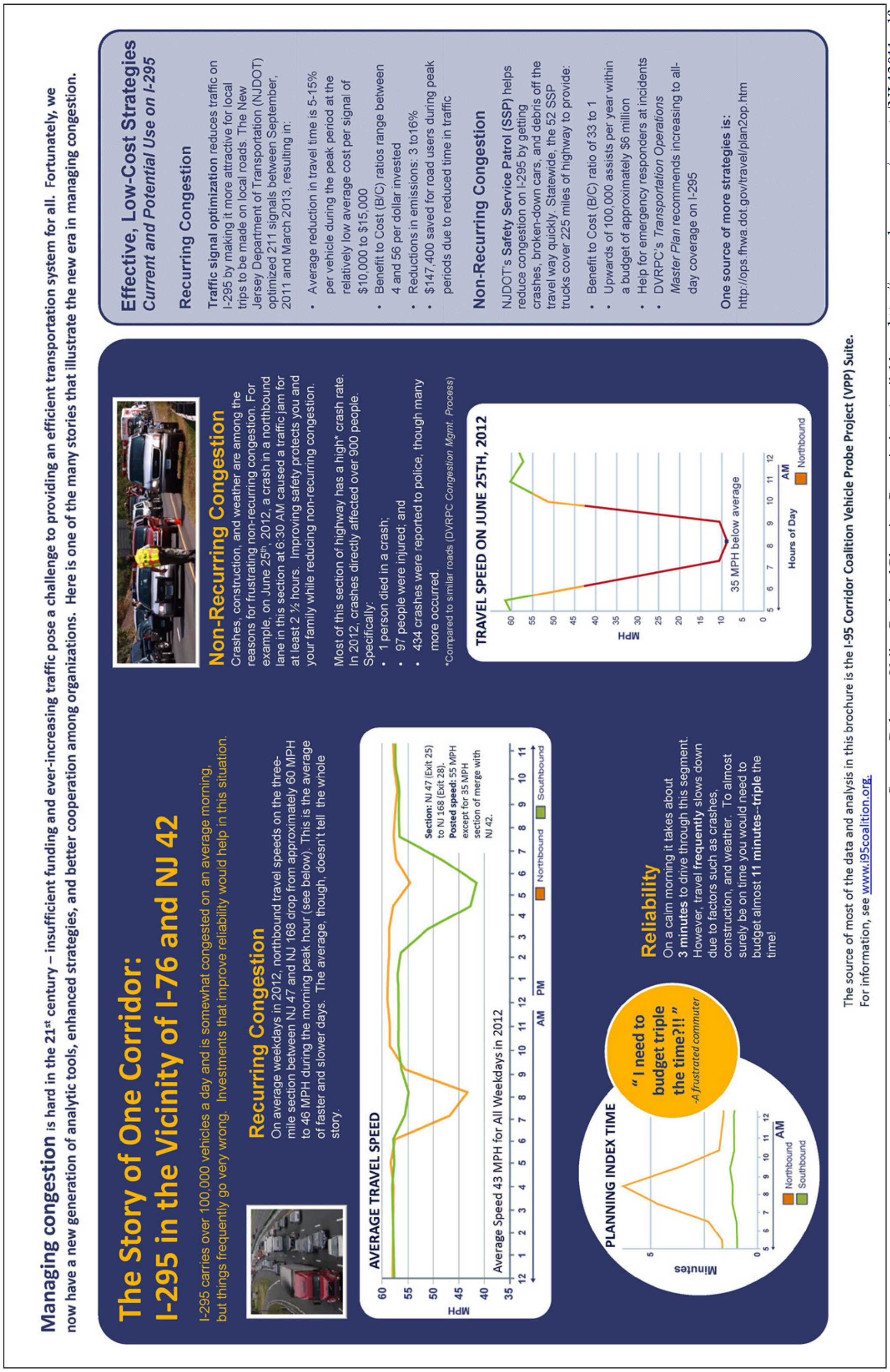
Agency Process

DVRPC uses archived operations data to track the performance of the regional transportation network, identify and prioritize congested locations, and justify funding for the appropriate projects. In the MPO's 2015 CMP, DVRPC used two measures based on archived operations data: travel time index (TTI) and planning time index (PTI). DVRPC selected these measures because it anticipated that they were likely to become part of the FAST Act reporting requirements at the time of the CMP update. DVRPC used VPP data to analyze the TTI and PTI for all limited-access highway and arterial road segments with data coverage at the time of the update. It then used the results of TTI and PTI analysis, along with other performance measures, to identify



Source: Delaware Valley Regional Planning Commission. Available at: http://www.dvrpc.org/reports/NL13011.pdf..

Figure 36. Image. Delaware Valley Regional Planning Commission's outreach document is based on archived data.



Source: Delaware Valley Regional Planning Commission. Available at: http://www.dvrpc.org/reports/NL13011.pdf..

Figure 37. Image. Inside view of Delaware Valley Regional Planning Commission's outreach document.

congested corridors throughout the region. These congested corridors were used as a basis for selecting corridors for investment in ITS infrastructure through the *Transportation Operations Master Plan*.[36] The designation of CMP corridor was also used as an evaluation criterion for projects to be programmed in the LRTP and TIP. In addition, DVRPC has developed several outreach documents based on archived operations data to communicate its approach to mitigating congestion with the general public and elected officials (See Figures 36 and 37 for an example).

In addition to developing and reporting performance measures, DVRPC participated in an evacuation planning effort for the City of Philadelphia. The project involved extracting a microsimulation model of the city center from DVRPC's travel demand model, and then calibrating the model with the VPP's archived travel time data on key arterial, collector, and local roadways in the city. The model was used to analyze how to best evacuate buildings, which routes all modes would likely take, which intersections would cause major bottlenecks, and other factors that may impact efficient evacuation of the city during a major event. DVRPC also was using archived operations data to spot-check calibration of the regional model and has begun an effort to use it more extensively in calibration.

Impact and Lessons Learned

Overall, access to archived operations data has been a significant benefit to planning at DVRPC, especially because the DVRPC staff have access to a user-friendly interface through the VPP Suite. The VPP Suite allows for both quick analysis at the corridor level and more complex regional analyses. The archived operations data enhances the agency's knowledge of key travel attributes, including peak hour travel times and reliability, but DVRPC has had challenges. One challenge has been working with the Traffic Message Channel network on which the archived speed and travel time data is reported and translating data from the Traffic Message Channel network to the base GIS line-work for analysis with other data sets.

DVRPC manages eight traffic incident management task forces (IMTFs) throughout the region. Incident data is reported at quarterly meetings of these corridor-based IMTFs. Using data from TMCs, incident durations help the task forces identify which incidents should be discussed in more detail in post incident debriefs. Occasionally, the VPP data is used to show congestion impacts of incidents.

DVRPC, a bi-state MPO, has several hurdles to overcome to report incident data on a regional basis. DVRPC would need to perform the non-standardized geospatial referencing of data, expend significant resources fusing data from two States, and clean up inconsistencies. The reporting of incident data would need to be standardized as well as geo-referenced across all corridors and agencies to support a complete regional picture. For example, incident duration, roadway clearance times, and other data may be entered differently depending on the agency, and NJDOT and Pennsylvania DOT—DVRPC's two primary resources for incident data—do not necessarily capture this data in the same way. DVRPC has considered cross-checking the incident data against police crash databases, although this data lags by 6 months or more and does not contain municipal crash reports. The integration of police CAD data with agency traffic incident data would help create regional incident data.

36 Delaware Valley Regional Planning Commission, Transportation Operations Master Plan (Philadelphia: PA, 2009). Available at: http://www.dvrpc.org/Transportation/TSMO/MasterPlan/.

DVRPC also faces challenges with inconsistent data reporting methods and formats with other types of archived operations data because it covers a bi-state area. Agencies in the region have various capabilities to collect, analyze, and share archived data, as well. The transit agencies in the region are an example of this discrepancy, which makes it difficult to perform an analysis for the entire region.

Another barrier to using archived operations data is the sheer amount of data that needs to be processed. The probe data for the DVRPC region amasses nearly 3.4 billion records per year. Working with this data is extremely difficult without powerful tools and hardware.

Use of Archived Operations Data for Input and Calibration of Reliability Prediction Tools

Need

Including travel time reliability as an aspect of mobility is becoming an important consideration for transportation agencies. Reliability is valued by users as an additional factor in how they experience the transportation system. Transportation improvements affect not only the average congestion experienced by users, but reliability as well. Therefore, adding reliability to the benefit stream emphasizes the positive impact of improvements.

Archived operations data enables the direct measurement of reliability, and many agencies have added reliability as a key performance measure. However, the ability to forecast reliability has lagged behind the ability to measure it. Being able to predict the reliability impacts of improvements, in addition to traditional notions of performance, provides a more complete perspective for planners. It also provides continuity in tracking the performance of programs because the tracking includes measures of reliability.

Fortunately, the Strategic Highway Research Program 2 (SHRP 2) produced several tools for predicting travel time reliability.[37, 38] These tools require a variety of inputs that can be developed from archived operations data and are used here to provide specific examples of how archived data can be used in conjunction with advanced modeling procedures. Several other modeling platforms—as well as those likely to emerge in the near future—will require similar levels of inputs.

Relevant Data

Data for input and calibration of reliability prediction tools include travel time (or speed) data, preferably continuously collected; volume data; and incident characteristics data.

Analytical Procedures

Procedures to predict travel time reliability are just now emerging in the profession (as of this writing). There are two basic approaches to predicting reliability: (1) statistical methods that develop predictive equations for reliability as a function of either congestion source factors (e.g., demand, capacity, incident characteristics, and weather conditions) or model-produced values for the average or typical condition, and (2) direct assessment of reliability by creating a travel time distribution that reflects varying conditions from day-to-day.

37 Economic Development Research Group, *Development of Improved Economic Analysis Tools Based on Recommendations from Project C03 (2)*, The Second Strategic Highway Research Program, Project C11, Transportation Research Board (Washington, DC: 2013). Available at: http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=2350.

38 Kittelson & Associates et al., *Incorporation of Travel Time Reliability into the Highway Capacity Manual,* The Second Strategic Highway Research Program, Report S2-L08-RW-1, Transportation Research Board (Washington, DC: 2014). Available at: http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_S2-L08-RW-1.pdf.

Impact

Reliability prediction tools are needed by analysts as a way to capture the impacts of operations strategies that are primarily aimed at improving travel time reliability. Many agencies have identified the need to improve reliability, yet have not had the means to predict it as a benefit of transportation improvements. Reliability tools provide a means for evaluating operations projects side-by-side with capacity expansion projects. When applied to planning activities, the tools will help transportation agencies better identify and implement strategies to reduce the variability and uncertainty of travel times for commuters and other travelers as well as the freight industry.

Case Study: Developing Data for Reliability Prediction Tools

Overview

A hypothetical MPO is embarking on an update to its LRTP. It has previously developed a scoring process for ranking potential projects in the plan, and it wants to add travel time reliability as one of the scoring factors. However, the MPO's TDF model only produces average speed estimates, and these estimates are artifacts of the traffic assignment process that are not reflective of actual speeds.

Data

The MPO has ITS detector data (volumes and speeds) on regional freeways available. It also has the NPMRDS travel time data for National Highway System (NHS) roadways in the region. Incident data on freeways is available via the State-run traffic management center.

Agency Process

The TDF model produces outputs by individual network link in a readable file. The time periods are a 2-hour AM peak and a 3-hour PM peak. Outputs include assigned traffic volumes, speeds, and delay for each of the periods.

Specific activities that can be undertaken with archived operations data to improve reliability prediction models include:

- Customize a volume-delay function for estimating recurring delay. Archived ITS detector data (volumes and speeds) can be used to develop a freeway function. This would involve removing observations that were influenced by incidents, weather, and work zones, then computing the volume to capacity (v/c) ratio for each speed measurement. Alternately, the volume-delay function could be developed using the speeds that include all sources of congestion.
- Customize the reliability prediction relationship. Archived ITS detector or vehicle probe data is used for this purpose. For a set of study links, the mean travel index is computed over the course of a year along with the reliability metrics of interest. Curve fitting is then performed to derive the relationships.
- Develop relationships between LRTP project types and the factors considered by the recurring and incident delay relationships, especially for operations strategies. This is done by searching the literature, not by using archived data, but it is a critical step. (Archived data may have been used by the researchers to develop these relationships, however.) The factors that can be influenced are free-flow speed, volume (demand), capacity, incident duration, and incident frequency.

The staff determined that operations strategies would best be handled by creating "bundles" of investments scenarios. These were combinations of strategies at different deployment intensity levels. Benefits and costs for multiple investment scenarios (bundles) were developed for further study. Then, the resulting system performance forecasts from the TDF model and the post-processor were used as part of a public engagement program for the LRTP, asking citizens to choose their preferred level of investment and corresponding performance outcome.

For example, travel time reliability was a performance measure in the operations bundles of projects. Investing in advanced traffic management systems does not produce a quantifiable benefit when using the TDF models alone, because it only considers demand, capacity, and recurring delay. Focusing instead on reliability produced a quantifiable benefit using the post processor. Those benefits were then paired with cost estimates for the ATMS and other operational improvements in the project scoring process. It also allowed the public and elected officials to make informed investment decisions when finalizing the project list.

▸ *Developing Capacity Values (Freeways)*

Use of localized capacity values is extremely significant in enabling forecasting models to replicate field conditions. Data from ITS detectors that produce simultaneous measurements of volume and speed can be used in two ways to determine the capacity of roadway segments, especially bottlenecks. First, classic speed-flow plots can be constructed. For this exercise, we assume that the volume and speed measurements have been archived in 5-minute time intervals:

- Aggregate the 5-minute measurements to 15-minute intervals. This is done for consistency with the definition of capacity used in the Highway Capacity Manual (HCM). Take only 15-minute time intervals with all three 5 minute intervals present.
- For each directional location, sum the volumes and compute the volume-weighted harmonic average speed.
- Convert 15-minute volumes to volume per hour per lane by first dividing by the number of lanes at the location and then multiplying by four.
- Plot the results.

Figures 38 through 43 show the results obtained by applying the above procedure to data from I-4 in Orlando, Florida. Several observations can be made based on these plots:

- The general shape of the plots conforms to those shown in the HCM and produced by other researchers. There is a small amount of "noise" in the data, as shown by points scattered away from the general trend lines. These could be errors in the measurements or indicative of transitional flow, which can vary in its traffic flow parameters.
- The flow values are in vehicles per hour per lane, not passenger cars per hour per lane (pcphpl), as specified in the HCM. That is, trucks are imbedded in the flow measurements, and an average value for percent of trucks must be used to expand vehicles to passenger cars.
- The large number of observations grouped around a horizontal line between 60 to 70 mph is indicative of the free-flow speed, although the next section will identify a more direct method for computing it.

- Capacity values are indicated by the right end "nose" of the curve. Capacity varies significantly throughout this highway section. A more formal analysis appears in Table 8. Selecting the maximum value as the capacity value is not recommended due to possible inaccuracies in the measurements, as can be seen by the spread between the 99.5 percentile and the maximum value. Therefore, a capacity value between the 99.5 percentile and the maximum value should be selected. For illustration, we have selected the midpoint.

Table 8. Upper end of speed-flow distribution, I-4, Orlando, Florida.

Location	Vehicles Per Hour Per Lane				
	95th percentile	99th percentile	99.5 percentile	Maximum	Selected Capacity
I-4 at EE Williamson_EB	1,687	1,889	1,944	2,101	2,023
I-4 at EE Williamson_WB	1,704	1,905	1,959	2,119	2,039
I-4 at Kaley Ave_WB	1,445	1,554	1,594	1,752	1,673
I-4 at Michigan Ave_EB	1,515	1,619	1,652	1,836	1,744
I-4 at Wymore_EB	1,999	2,116	2,144	2,247	2,195
I-4 at Wymore_WB	1,892	2,051	2,101	2,315	2,208
I-4 East of Wymore_EB	1,991	2,099	2,124	2,231	2,177
I-4 East of Wymore_WB	1,915	2,071	2,117	2,344	2,231
I-4 West of SR 408_EB	1,500	1,633	1,677	1,820	1,749

EB = eastbound, WB = westbound

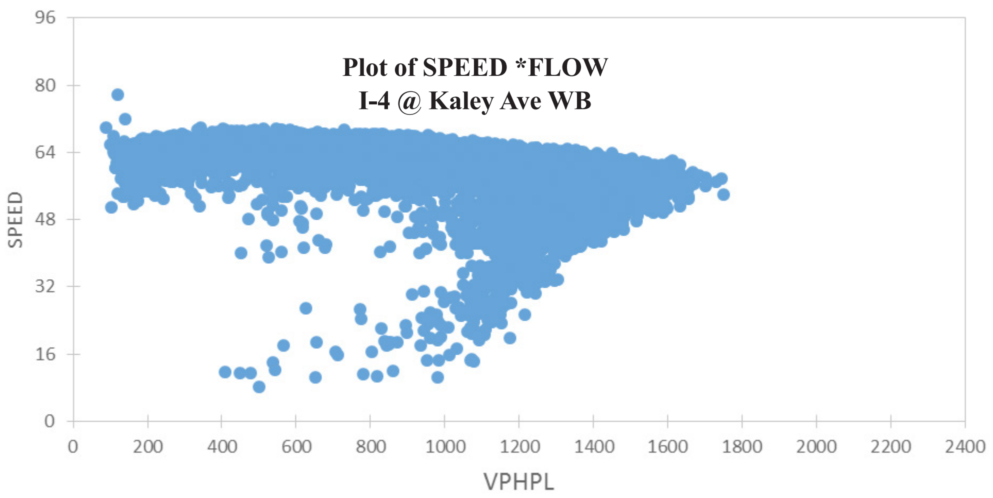


Figure 38. Graph. Plot of speed and vehicles per hour per lane on I-4 at Kaley Avenue, westbound.

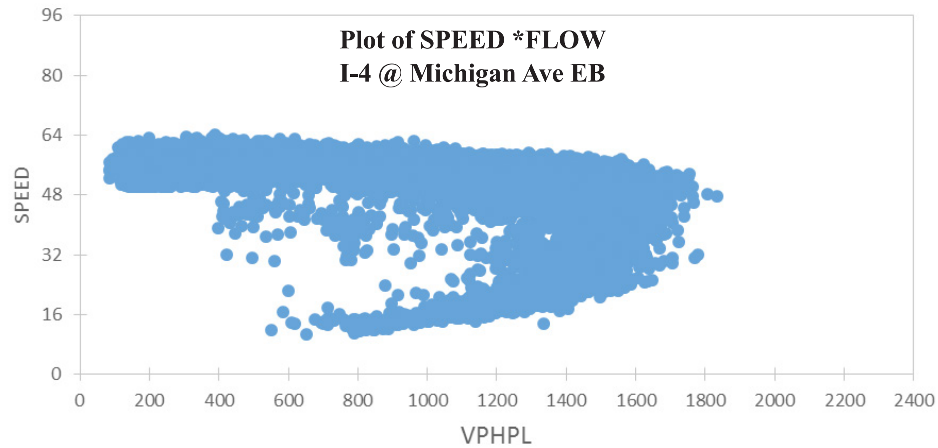


Figure 39. Graph. Plot of speed and vehicles per hour per lane on I-4 at Michigan Avenue, westbound.

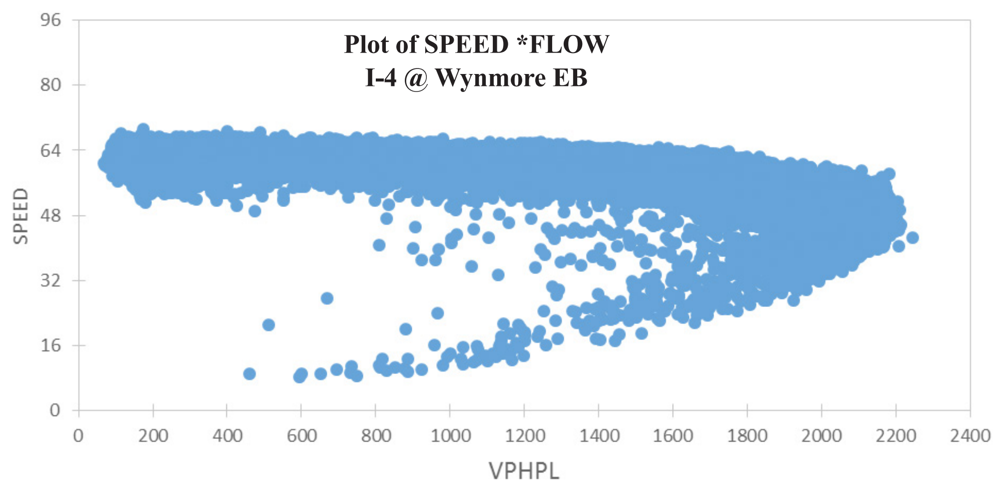


Figure 40. Graph. Plot of speed and vehicles per hour per lane on I-4 at Wynmore, eastbound.

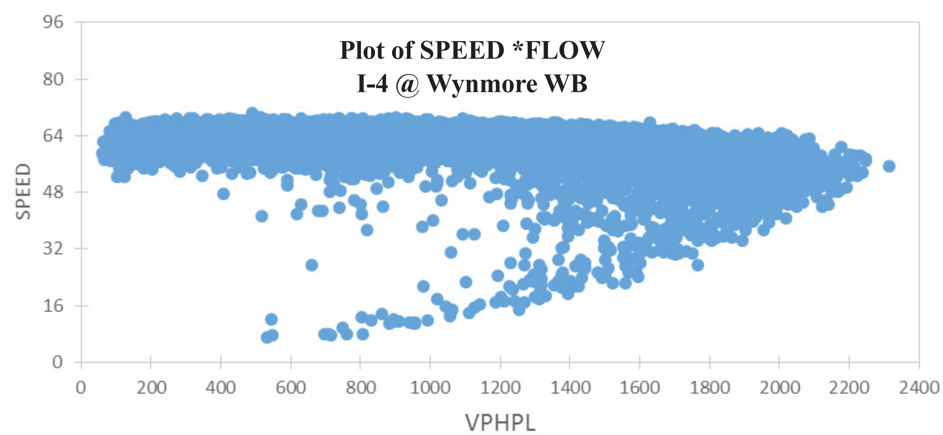


Figure 41. Graph. Plot of speed and vehicles per hour per lane on I-4 at Wynmore, westbound.

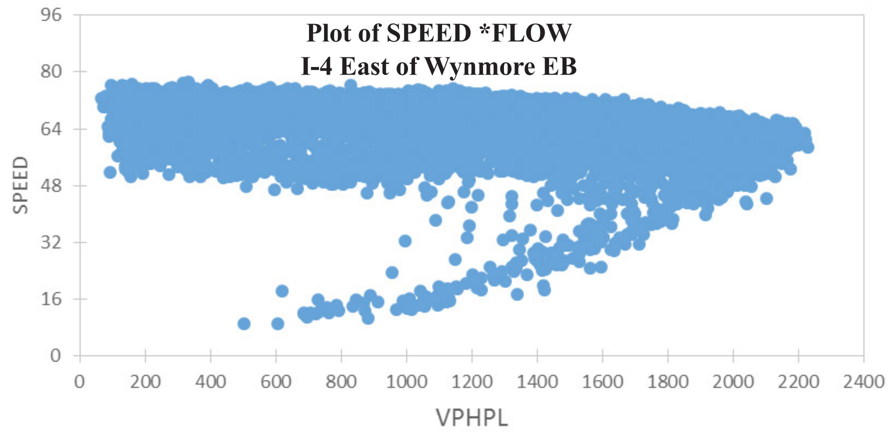


Figure 42. Graph. Plot of speed and vehicles per hour per lane on I-4 east of Wynmore, eastbound.

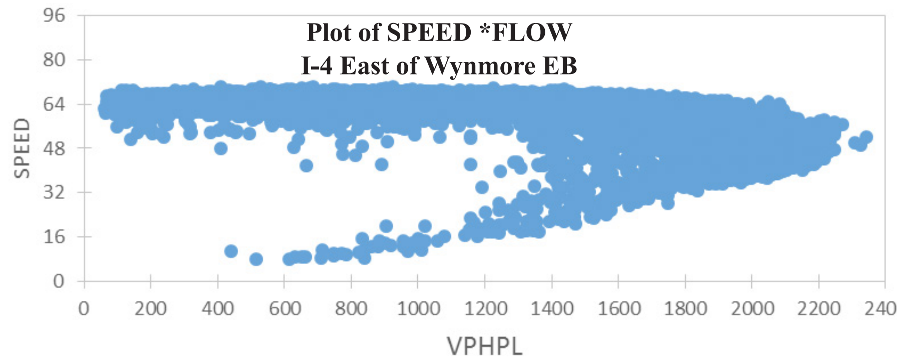
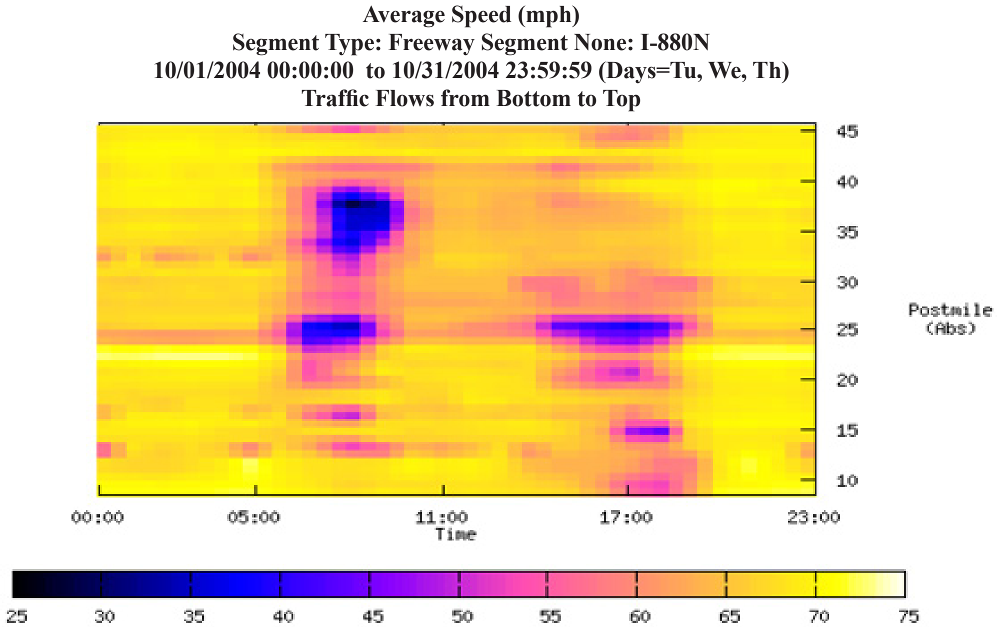


Figure 43. Graph. Plot of speed and vehicles per hour per lane on I-4 east of Wynmore, westbound.

A second approach to determining capacity is heuristic: adjust (calibrate) the capacity value in the model to match local traffic conditions. Here, either speed data from ITS detectors or vehicle probes can be used; volumes are not used in the calibration but must be input in the models separately:

- From the speed data in the corridor, identify existing bottlenecks. These will be locations (detectors or links) where queues are building upstream while downstream speeds are relatively high, unless influenced by another bottleneck (Figure 44).
- Set the bottleneck segment to a capacity according to the HCM capacity estimate based on facility free-flow speed (e.g., 2,400 pcphpl for 70 mph free-flow speed (FFS)). Run the model.
- If a queue does not result, lower the capacity value in increments of 50 pcphpl until a queue results and speeds in the queue roughly match the field data.



Source: System Metrics Group and Cambridge Systematics, unpublished presentation to California Department of Transportation, August 26, 2005.

Figure 44. Graph. Speed contours from archived intelligent transportation system detector data are useful for identifying bottleneck locations.

▸ *Determining Free Flow Speed (FFS)*

Both ITS detector and vehicle probe data can be used to determine the FFS of a facility. A variety of approaches can be used to determine the FFS. The first method uses only speed data and can be used on both freeways and signalized highways. It is based on selecting the 85th percentile speed under very light traffic conditions, defined as occurring morning hours on weekend days (e.g., from 6:00 AM to 10:00 AM). The second method uses freeway ITS detector volume and speeds. It determines the 85th percentile speed when volumes are low and speeds are relatively high (to weed out the low volumes that occur in extreme queuing conditions).

Table 9 compares these approaches for the freeways in the Atlanta, Georgia, region. The values in this table are computed for the entire facility, which includes multiple segments. To calculate this value, the space mean speed (SMS) over the entire facility is first calculated by the smallest time interval in the data (in this case, 5 minutes). With detector data, because companion volumes are available, SMS is the facility vehicle-miles of travel (VMT) divided by the facility vehicle-hours of travel. With probe data, SMS is the total distance divided by travel time in hours. Then, a distribution is developed from those facility space mean speeds.

Table 9. Facility free-flow speed calculations, Atlanta, Georgia, area freeways.

Section	Weekend FFS Calculation Method (mph)	Low Volume FFS Calculation Method (mph)
I-75 Northbound from I-285 to Roswell Road	72	70
I-75 Southbound from I-285 to Roswell Road	69	68
I-75 Northbound from I-20 to Brookwood	69	69
I-75 Southbound from I-20 to Brookwood	68	67
I-285 Eastbound from GA-400 to I-75	71	69
I-285 Westbound from GA-400 to I-75	69	68
I-285 Eastbound from GA-400 to I-85	68	66
I-285 Westbound from GA-400 to I-85	68	68
I-75 Northbound from Roswell Road to Barrett Parkway	71	70
I-75 Southbound from Roswell Road to Barrett Parkway	69	69

FFS=free flow speed

▸ *Identifying Demand (Volumes)*

Data from ITS freeway detectors can be used to develop input volumes at the hourly or finer level. The main issue with using detectors on already-congested freeways is that, under queuing, the volume they measure is lower than the actual demand. If this is the case, then the following procedure can be used. It is based on moving upstream from where queues typically occur to locate a detector that is relatively free of queuing influence, but its traffic patterns are similar to the facility being studied. If a major interchange is present between the study detector and the facility, traffic patterns are likely to be different. If the study detector meets these criteria, it is used to develop hourly factors for extended peak periods that can be applied to volumes on the freeway facility. The extended period should cover times on either side of the peak where queuing is not normally a problem (i.e., 1 hour before the peak start and 1 to 2 hours after the peak end). The factors are computed as the hourly volume divided by the extended peak volume. These are then applied to the extended peak period volumes on the facility.

Once demands have been created from the congested volumes, accurate temporal distributions can be created for the study section, such as the percent of daily traffic occurring in each hour. Also, the variability in demand can be established since some reliability prediction methods require knowledge about demand variability.

Determining Incident Characteristics

Advanced modeling methods require basic data about incidents. For example, annual number and average duration of the following incident categories:

- Crashes – property damage only.
- Crashes – minor injury.
- Crashes – major injury and fatal.
- Non-crashes – disabled vehicle/lane blocking.
- Non-crashes – disabled vehicle/non-lane blocking.

Two sources of data are used to develop these inputs: (1) traditional crash data, and (2) archived incident management data. Crash data has details on crash severity, whereas incident data may or may not. Conversely, crash data does not include duration information. One approach is to use state- or area-wide values for crash severity and then to rely on incident management data. The first step is to estimate incidents by type. Typical categories in incident management data are crashes, disabled vehicles, and debris. Table 10 shows data taken for I-66 in Virginia, from the I-495 Capital Beltway west for 30 miles.

Table 10. Incident characteristics for I-66, Virginia.

CRASH						
			Duration			
Obs	Lateral Location	Frequency	Average	Standard Deviation	Median	95th Percentile
1	1 Lane Blocked	366	59.2505	75.359	43.5	1.29
2	2 Lanes Blocked	158	67.7142	55.357	55.5	1.54
3	3+ Lanes Blocked	67	77.5970	60.302	57.0	1.82
4	No Blockage	212	52.7137	111.603	32.0	1.39
5	Shoulder Blocked	531	69.5431	101.793	51.0	1.53

Obs = observation

Table 10. Incident characteristics for I-66, Virginia. (Continued)

DEBRIS						
			Duration			
Obs	Lateral Location	Frequency	Average	Standard Deviation	Median	95th Percentile
1	1 Lane Blocked	68	53.0466	65.727	31.5	242
2	2 Lanes Blocked	5	83.6000	103.667	38.0	258
3	3+ Lanes Blocked	2	26.5000	20.506	26.5	41
4	No Blockage	54	33.6296	38.253	27.0	94
5	Shoulder Blocked	27	50.5926	50.400	45.0	165

DISABLED						
			Duration			
Obs	Lateral Location	Frequency	Average	Standard Deviation	Median	95th Percentile
1	1 Lane Blocked	264	58.126	91.48	32.5000	169
2	2 Lanes Blocked	12	49.833	21.02	53.0000	79
3	3+ Lanes Blocked	7	263.143	469.49	38.0000	1297
4	No Blockage	638	577.226	9908.04	8.5833	85
5	Shoulder Blocked	1552	54.147	118.88	29.0000	182

Obs = observation

Source: Margiotta, Richard, unpublished internal memo for SHRP 2 Project L08, August 26, 2005.

Impact and Lessons Learned

To explore the various investment options preferred by the public, several financial scenarios were prepared for the MPO board's consideration. The scenarios varied by spending pattern and by whether they included new funding sources for transportation, such as mobility fees on new development or increases in the gas tax or sales tax. The final outcome of the process was that operations projects were successfully evaluated on an equal basis with traditional projects and are now part of the LRTP.

Performance of Before and After Studies to Assess Projects and Program Impacts

Need

A key component of performance management is the ability to conduct evaluations of completed projects and to use the results to make better-informed investment decisions in the future. State departments of transportation (DOTs) and MPOs routinely invest great effort in forecasting the impact of capital, operating, and regulatory improvements on general and truck traffic trips, but practitioners seldom have the opportunity to evaluate the outcomes of their investments, except in isolated cases where special studies are conducted.

However, no standardized method for conducting before and after evaluations with these data sources exists, especially for operations projects. Although a rich history of evaluations has been accumulated, the vast majority of studies lack a consistent method; common performance measures; and, most problematic, controls for dealing with factors that can influence the observed performance other than the project treatment.[39] A thorough treatment of an evaluation methodology is beyond the scope of this primer. Regardless of what a comprehensive methodology entails, at its core, it will be examining travel time and demand data. The remainder of this section focuses on the use of these data types.

Relevant Data

A comprehensive and statistically valid evaluation methodology will have to draw on a variety of data types to establish controls. Many factors can influence travel time besides an improvement project, including incidents, demand (e.g., day-to-day and seasonal variations and special events), weather, work zones, traffic controls, and general operations policies. Ideally, the influence of these factors is stable (or nearly so) in the before and after periods of project implementation, which allows for observed changes in travel times to be untainted. In theory, all of these data types can be obtained from archived operations data, although in practice, some operations may choose not to archive certain types of data.

Analytical Procedures

Define the Geographic Scope of the Analysis

The geographic scope is driven by the target area of the project. For example, if an analysis is performed to pinpoint the safety benefits from the implementation of an incident management strategy over a stretch of a freeway, then the scope would be limited to that stretch. On the other hand, if the objective is to quantify safety improvements for an entire downtown area, then the scope would cover all arterials and roads in the downtown area.

In operations evaluations, a "test location" typically means a roadway section that has implemented certain operations strategies or has seen the impacts of such strategies. For analysis, roadway sections are typically 5 to 10 miles in length. If a larger area is covered, it is advisable to break it up into multiple sections. Study sections should be relatively homogenous in terms of traffic patterns, roadway geometry, and operating characteristics.

Define Analysis Period

The temporal aspects that should be considered in before and after evaluations are:

- **Time period.** Because of the need to compute reliability, a year's worth of data in each of the before and after periods is preferred. Six months is an absolute minimum, but it is likely that seasonal effects will come into play unless the same months of different years are used.
- **User-defined peak periods (AM and PM), mid-day, and off peak periods.** The peak periods should be defined for weekday non-holidays. At a minimum, the peak periods should be analyzed; other periods may be added at the analyst's discretion. Note that, in some cases, the time period of interest may be different from those mentioned (e.g., weekends in rural recreational areas may be the focus of an operational treatment).

39 United States Department of Transportation, Intelligent Transportation Systems (ITS) benefits database. Available at: http://www.itsbenefits.its.dot.gov/.

Establish Travel Time Performance Measures

Many different types of performance measures are available. Below is a non-exhaustive list that can be considered:

- **Travel time (minutes) –** Because the project length or coverage is usually the same in the before and after periods, straight travel time may be used as a performance measure.
- **Mean travel time index (TTI) (unitless) –** The ratio of average travel time to the ideal or free-flow travel time. For evaluating an individual project, the ideal or free-flow travel time should be the same in the before and after periods. Although straight travel time is recommended also, the TTI is useful because it is normalized and the project may be compared to others with different project lengths.
- **Delay (vehicle-hours and person-hours)** – The actual vehicle- or person-hours of travel that occur minus those that occur under free-flow/ideal conditions. Delay is a useful measure because economic analyses have a long history of applying monetary value to delay.
- **PTI –** The 95th percentile travel time divided by the free-flow travel time.
- **80th percentile TTI.**
- **Incident Performance Measures:**
  - Total incidents by type: crashes, stalls, and debris.
  - Incident rate (incidents per 100 million VMT).
  - Incident duration: mean and standard deviation.
  - Lane-hours lost due to incidents (number of lanes closed multiplied by the number of hours they are closed).
  - Shoulder-hours lost due to incidents.

Develop Performance Measures from the Data

The following example illustrates the initial steps in conducting a before and after evaluation, recognizing that more sophisticated analysis using statistical controls should be performed next.

An operational improvement is made on a freeway section. It was implemented quickly, with minimal disruption to traffic. Travel time and volume data were available for 1 year prior to the completion of the improvement and 1 year after. VMT, mean TTI, and PTI were computed for each week and plotted (Figures 45 and 46). Also, incident lane-hours and shoulder-hours were computed (Table 11).

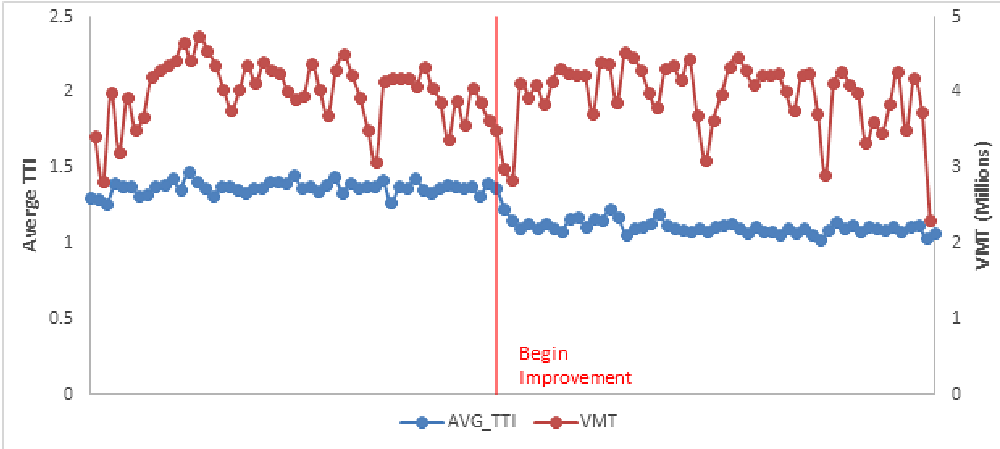


Figure 45. Graph. Hypothetical before and after case: mean travel time index and vehicle-miles of travel.

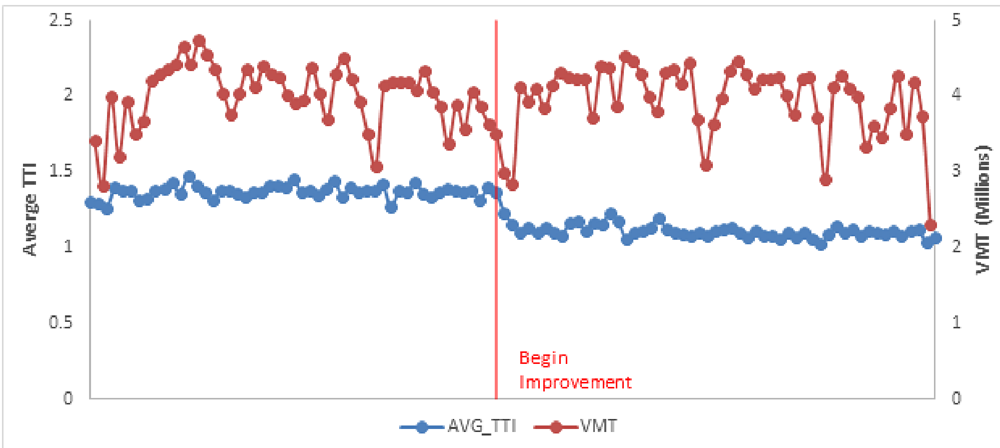


Figure 46. Graph. Hypothetical before and after case: planning time index and vehicle miles of travel.

Table 11. Incident blockage characteristics for case study.

Period	Incident Lane-Hours Lost	Incident Shoulder-Hours Lost
Before	35.3	51.3
After	26.7	66.1

The results show that the improvement reduced both the mean TTI and PTI and that VMT was roughly similar in the before and after periods. As expected, the PTI shows more volatility than the mean TTI because it measures extreme cases, which can vary greatly from week to week. Incident lane-hours lost is higher in the before period, while incident shoulder-hours lost are higher in the after period. A more thorough examination of incidents, as well as other factors, should be pursued to ensure that the congestion and reliability improvement are due to the project and not to other factors.

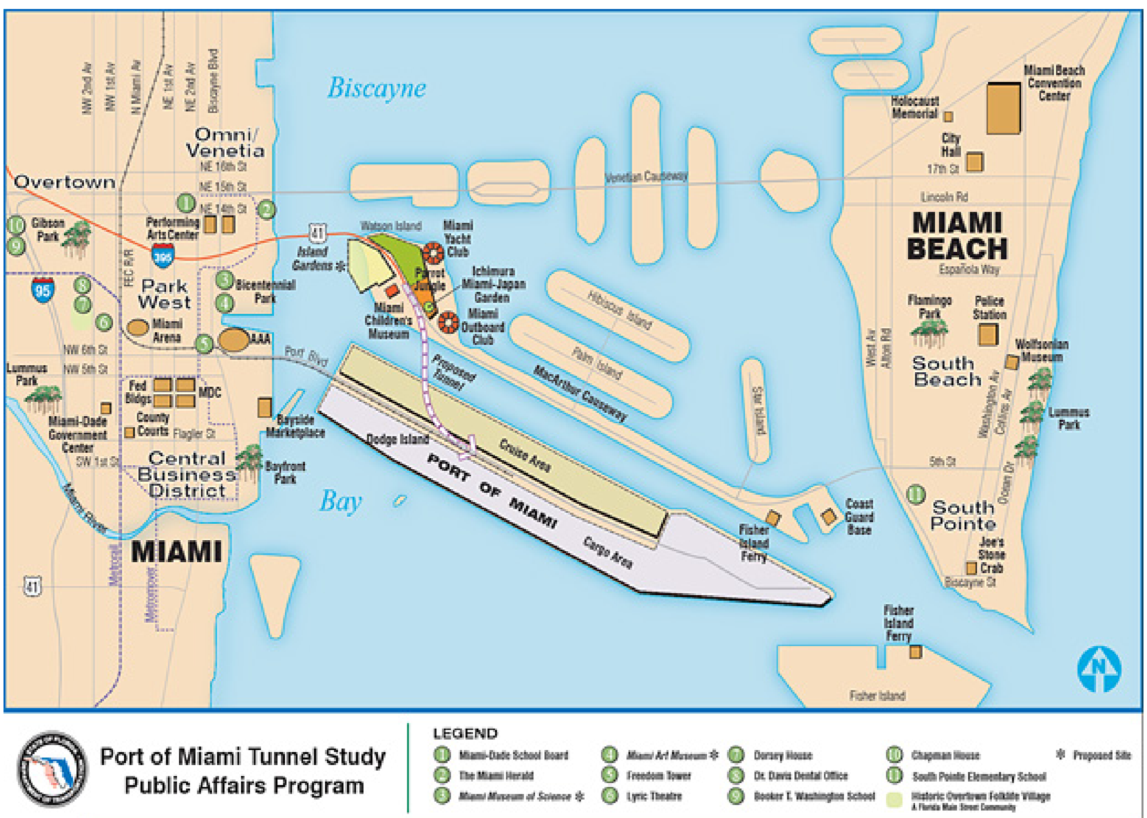
▸ *Impact*

Development of an ongoing performance evaluation program enables planning agencies to make more informed decisions about investments. The results of evaluations can be used to highlight agency activities to the public and to demonstrate the value of transportation investments. The knowledge gained from evaluations can lead to development of impact factors for future studies, as well as an understanding of where and when certain types of strategies succeed or fail. These impact factors, tailored to an agency's particular setting, are useful for future planning activities when alternatives are being developed. For example, evaluations of the conditions under which ramp metering is effective can avoid future deployments in situations where they will not be successful.

▸ *Case Study: Florida Department of Transportation Evaluation of the Port of Miami Tunnel Project*

Overview

The Florida DOT (FDOT) used archived data to evaluate the impact that the Port of Miami Tunnel project had on performance. The Port of Miami Tunnel project was built by a private company in partnership with FDOT, Miami-Dade, and the City of Miami. By connecting State Route A1A/ MacArthur Causeway to Dodge Island, the project provided direct access between the seaport and highways I-395 and I-95, thus creating another entry to Port of Miami Tunnel (Figure 47). Additionally, the Port of Miami Tunnel was designed to improve traffic flow in downtown Miami by reducing the number of cargo trucks and cruise-related vehicles on congested downtown streets and aiding ongoing and future development in and around downtown Miami.



Source: Florida Department of Transportation, Port of Miami Tunnel, Project Maps. Available at: http://www.portofmiamitunnel.com/press-room/project-maps/

Figure 47. Map. The Port of Miami tunnel project.

Data

FDOT used travel time data from NPMRDS and traffic volume data from their own system to conduct the evaluation.

Agency Process

The purpose of this study is to examine the potential impacts to the transportation system resulting from the opening of the Port of Miami Tunnel. The study utilizes tabular and geospatial data from FDOT and the NPMRDS. Using a geographic information system or GIS, these data sources are merged together spatially to create value-added information that would otherwise be unavailable. The resulting information provides FDOT with a better understanding of what impacts there are to the transportation system, the magnitude of these impacts, and where these impacts are occurring. Impacts considered in the study include:

- Differences in truck flows using truck count data from FDOT.
- Congestion – travel time data from HERE to compare and analyze average travel speeds (pre- and post- tunnel) within the tunnel vicinity.

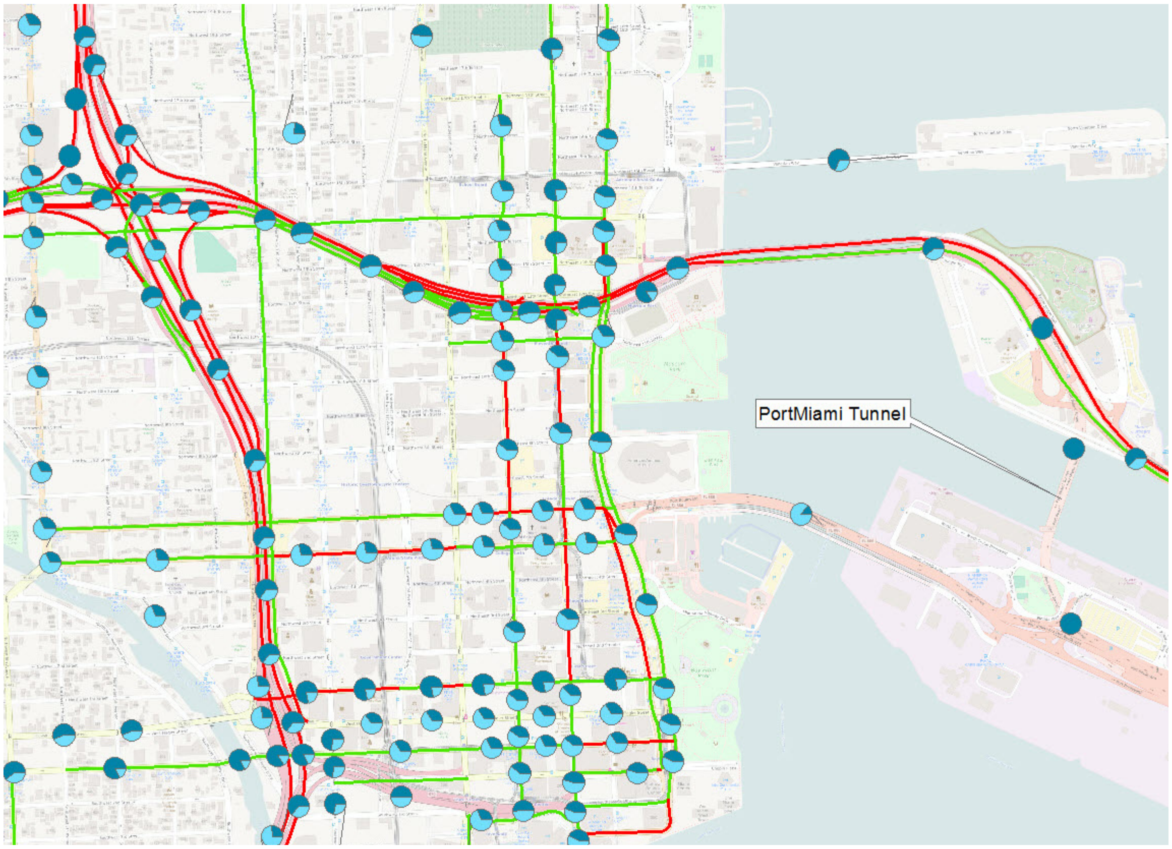
This study compares the average weekday travel speeds during peak hours in August 2013 and August 2015. The speed data for August 2013 was collected prior to the opening of the Port of Miami Tunnel. Data for the corresponding month in 2015 was collected after the opening of the Port of Miami Tunnel (Figure 48). In Figure 48, the pie symbols represent differences in truck counts, with 2015 data shown in dark blue and 2013 data shown in light blue. The red and green lines represent a decrease in average weekday travel speed and increase in average weekday travel speed, respectively. Results from the evaluation showed:

- Improved or similar travel speeds through downtown Miami.
- Improved or similar travel speeds along MacArthur Causeway eastbound, portions of I-395 East, and portions of Biscayne Boulevard.
- Slower speeds along portions of I-395 West, I-95, and Biscayne Boulevard westbound.

Impact and Lessons Learned

Analysis of the transportation system resulting from the opening of the Port of Miami Tunnel is just an example of how powerful value-added tabular and spatial data can be. The long-term vision is to empower planners, freight coordinators, consultants, and others throughout FDOT with value-added data and information that can assist with decision making and the achievement of Federal guidelines.

Although working with very large and complex data sets presents a number of challenges, including human resource, information technology (IT), and financial availability, the long-term benefits are expected to outweigh the challenges. Because much value-added data is generated via spatial analysis, disparate geometric (geographic information system) networks should be edited and aligned with each other in an ongoing fashion. With the various networks merged together, the attributes associated with these networks can work together via spatial coincidence. The value-added tabular and spatial information generated through a variety of complex mechanisms will be made available throughout FDOT, including the Central Office in Tallahassee, Florida, and the District offices throughout the State.



Source: O'Rourke, Paul, Florida Department of Transportation, 2016, Unpublished.

Figure 48. Map. Truck flow and speed impacts of the Port of Miami tunnel project.

Identification of Causes of Congestion

Need

The prior case study identifies how patterns in congestion can be identified, ranked, and prioritized. However, simply identifying locations with problems is not enough; analysts need to know "why" a problem is occurring regardless of whether it is a recurring problem or non-recurring problem. Perhaps the congestion is related to a geometric issue with the road, or maybe it is incident or event related. Maybe it is simply a capacity issue.

The analysis also can be reversed: if a large incident has just occurred, how can the impacts (environmental or otherwise) of the related congestion and delays be measured? In this case study, the following will be considered:

- Determine the cause of the recurring congestion.
- Determine the cause of the non-recurring congestion.
- Examine the effect of events, incidents, or policy decisions on congestion.

Relevant Data

The data required to identify patterns and investigate cause and effect is not that different from the FDOT Port of Miami Tunnel case study above. An agency will need a data source that first can identify the presence of congestion. This could be a dense network of point-sensor data (e.g., inductive loops, side-fired radar/microwave, video detection) or it could be probe-based speed and travel time data (from commercial third parties, such as HERE, INRIX, and TomTom, or from Bluetooth or WiFi detection systems). The above data helps to identify the presence of congestion.

To identify the cause of the congestion, the following data sources may be of value: traffic management center ATMS event logs, police accident records, weather radar, road weather measurements, aerial photography, and volume data profiles.

Analytical Procedures

Operations data from probes or other high-density sensors can be used to identify congested areas or bottlenecks. When these data are comingled with ATMS event data, it becomes possible to identify if the congestion is recurring or non-recurring. Each agency may define congestion slightly differently, but in its most basic sense, congestion occurs any time speeds and travel times drop below some desired threshold. For some, that threshold is the speed limit. For others, a segment of a road becomes congested when the space mean speed on it falls below some percentage of the speed limit. Queues can be measured on the roadway by simply adding adjacent road segments that meet the same condition.

Flow data from probes and dense sensor networks can sometimes be challenging to work with because of the sheer volume of data, complex linear referencing systems that might not match up with agency centerline road databases, and agency technical and resource capacity. A relatively small state, Maryland, has nearly 12,000 road segments. With each segment providing data every minute, that comes to well over 6.3 billion data points every year.

To conduct a causality analysis, an agency must be able to associate the non-speed/flow data elements to the roadway. ATMS incident, event, or construction data and police accident data records will likely need to be assigned to a specific point on the roadway or a roadway segment. This could require some basic geospatial queries, but in certain cases may be more complex. Some speed data from third parties is provided in a proprietary geospatial standard referred to as traffic message channel segments. This georeferencing system will likely be different than most agency centerline road networks. Furthermore, some police accident records only store their data in log-miles or other georeferencing systems. Aligning all of these different linear referencing systems can be time consuming.

As with police incident and ATMS data, weather data will need to be merged with the congestion information so that the two can be correlated. Road weather data directly from road weather information system (RWIS) stations can usually be linked in the same manner as ATMS records; however, RWIS stations are few and far between and do not offer great coverage. Weather data from the National Weather Service is fairly ubiquitous, but can be more complex to work with. Radar reflectivity data that represents precipitation rates or snowfall can be associated with road segments using geospatial queries. There are commercial third-party road weather data products for sale that can be used in a similar way. They are usually pre-processed to reflect conditions directly on the roadway, and will be associated with various road segments in the same manner as probe-based speed data.

Once all of the disparate operations data are joined both geospatially and temporally, real analytics can begin. The analysis can range from big-data style correlation of coefficients—searching for any possible pair or set of variables to see if there might be a correlation between them. However, the analysis can be less complex: beginning with identifying a problem location, and then considering additional questions (what percentage of the time was the congestion at that location related to weather and/or incidents, and was the congestion greater when weather was a factor?). The opposite can occur; first identify major weather events, and then determine if congestion was worse in various locations during those events compared to normal. These types of analyses are discussed in the real-world case studies below.

Agencies with strong, integrated IT departments, dedicated data analysts, sufficient data storage hardware, and access to statistical software such as SAS can handle this type of analysis given time and proper funding. However, many agencies simply subcontract their analysis to consultants or university transportation departments, or they purchase off-the-shelf congestion analytics software and data integration services tailored for transportation planners.

▸ *Impact*

With the advancement of tools and technologies, roadway problems are more easily identified. However, significant effort is needed to effectively understand why a problem exists at a particular location—ultimately leading to the identification of solutions to the various problems. Impact and causality analysis are the end goal of archived operations data analytics. Impact analysis affords the opportunity to change a simple observation (e.g., last week's snowstorm caused significant delays in the region) into a more detailed observation (e.g., last week's snowstorm impacted 3 million travelers, costing some $45 million in user delay and excess fuel consumption, and resulting in 60 serious collisions that claimed the lives of three individuals in two separate families).

Progressive agencies that use archived operations data can further justify their investment (e.g., while the impacts of last week's snowstorm were significant and deadly per the prior statement, the damage could have been worse. The roads were pre-treated with 100,000 tons of sodium chloride brine, which reduced the plowing cleanup time by nearly 9 hours. Furthermore, safety service patrols reduced the clearance time of the 60 serious collisions by over 20 hours. These actions saved area commuters $32.8 million in excess delay, wasted fuel, and the reduction of secondary accidents).

These last two statements are only possible with a robust data archive. The results of such an analysis allow agencies to better plan decisions for future storms, better justify operating budgets, and respond to questions from the media and public officials in a more complete and timely manner.

▸ *Case Study 1: Maryland State Highway Administration Examines Causes of Recurring and Non-Recurring Congestion*

Overview

This case study determines the cause of both recurring and non-recurring congestion, by discussing existing tools, hiring a consultant, and doing the study in-house.

The Maryland State Highway Administration (MD SHA) routinely evaluates the impacts of congestion on its roadways with a focus on understanding why congestion is occurring, how significant the congestion is in terms of measurable impacts (e.g., queues, user delay costs), determining potential solutions to the congestion (e.g., projects, operational strategies), and then measuring the effectiveness of the chosen solution.

Data

Five archived operations data sets are used by MD SHA to perform the analysis: probe speed, volume, ATMS event, police accident, and weather. The probe data is derived from a private sector data provider. This probe data is archived at 1-minute intervals from real-time. The probe data represents the average speed across the length of a road segment for each 1-minute period.

The volume data is derived from two sources: HPMS-derived volume profiles and side-fired remote traffic microwave sensors (RTMS) detectors.

The event data is derived from two sources: the MD SHA Coordinated Highways Action Response Team (CHART) ATMS platform, and police accident reports collected and digitized by the Maryland State Police in coordination with MD SHA. The ATMS data is entered by operations personnel into their CHART platform in real-time. The type of information recorded includes the type of event (e.g., construction, collision, special event, disabled vehicle, police activity), location, lane closure information, and number of responders on scene.

The police accident reports contain additional details about the cause and nature of the incident, including the types of individuals involved and their conditions.

Weather data is a mixture of both weather radar data (coming from freely available National Weather Service data feeds) and RWIS station data. The RWIS stations are installed and operated by a DOT and include measurements such as wind speed and direction, precipitation rates, visibility, salinity content of the moisture on the roadway, and barometric pressure.

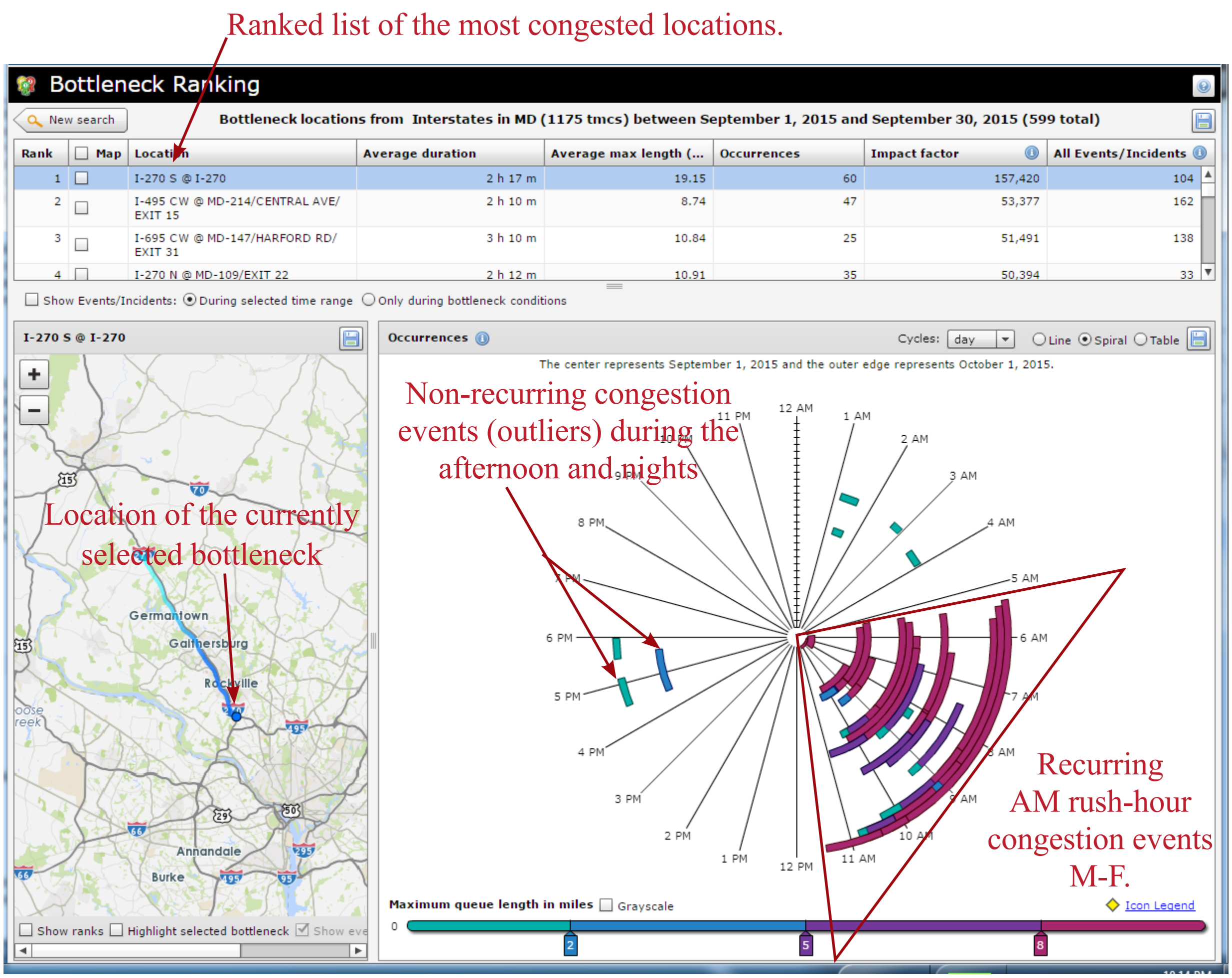
Agency Process

As with the case study on problem identification and confirmation with NJDOT, MD SHA uses a suite of third-party data analytics tools to collect, fuse, archive, and visualize their archived operations data. To evaluate the cause of congestion, MD SHA first determines where congestion is occurring. This can be accomplished using the probe data analytics suite and by following the same general process as described for NJDOT above.

Figure 49 shows an example of a bottleneck ranking query for September 2015 on all Maryland Interstates. By selecting the top-ranked congested location (I-270 southbound), the following can be determined:

- The congestion usually lasts 2 hours and 17 minutes.

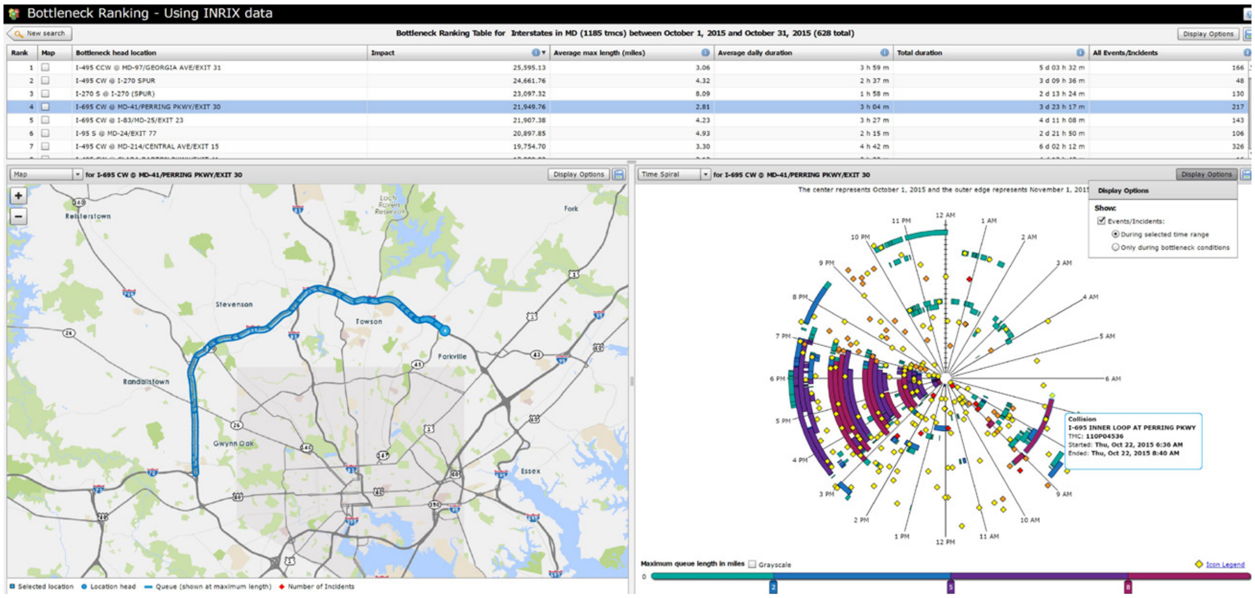
- The queue usually backs up to just over 19 miles.
- There were 104 events that fell within the geography of this congested location.
- The bulk of the congestion at this location occurs between 5:30 AM and 10 AM Monday through Friday, confirming this is primarily a recurring congestion event.
- Significant congestion can occur outside of the morning rush hour, as can be seen by the outlier color bands in the afternoon and the nights.



Source: Probe Data Analytics Suite. University of Maryland CATT Laboratory.

Figure 49. Screenshot. The bottleneck ranking tool.

Once a user determines when and where congestion is occurring, he or she can select additional tool overlays (e.g., construction, incident, special event) to map out data on top of the time spiral to help identify event causality. These event icons are interactive. Figure 49 shows a couple of afternoon outliers being the result of significant events. The one highlighted below happens to be an injury accident that occurred on a Saturday afternoon.



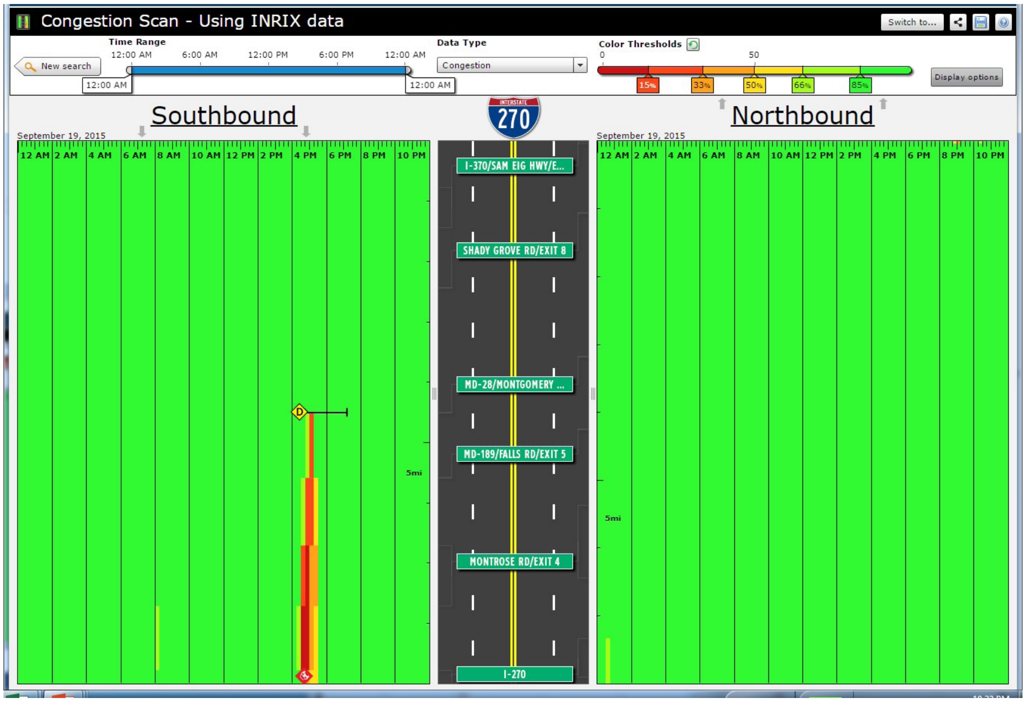
Source: Probe Data Analytics Suite. University of Maryland CATT Laboratory.

Figure 50. Screenshot. Congestion time spiral graphic with incidents and events overlaid.

In Figure 50, yellow diamonds denote incidents or disabled vehicles, red diamonds denote injury or fatality incidents, and orange diamonds denote construction or work zone events. Selecting the congestion lines then displays a congestion scan graphic of the speeds and travel times for that particular day and this particular corridor (as seen in Figure 57).

For further study and use, the user can select the congestion event in the time spiral, which will create other graphics depicting additional elements.

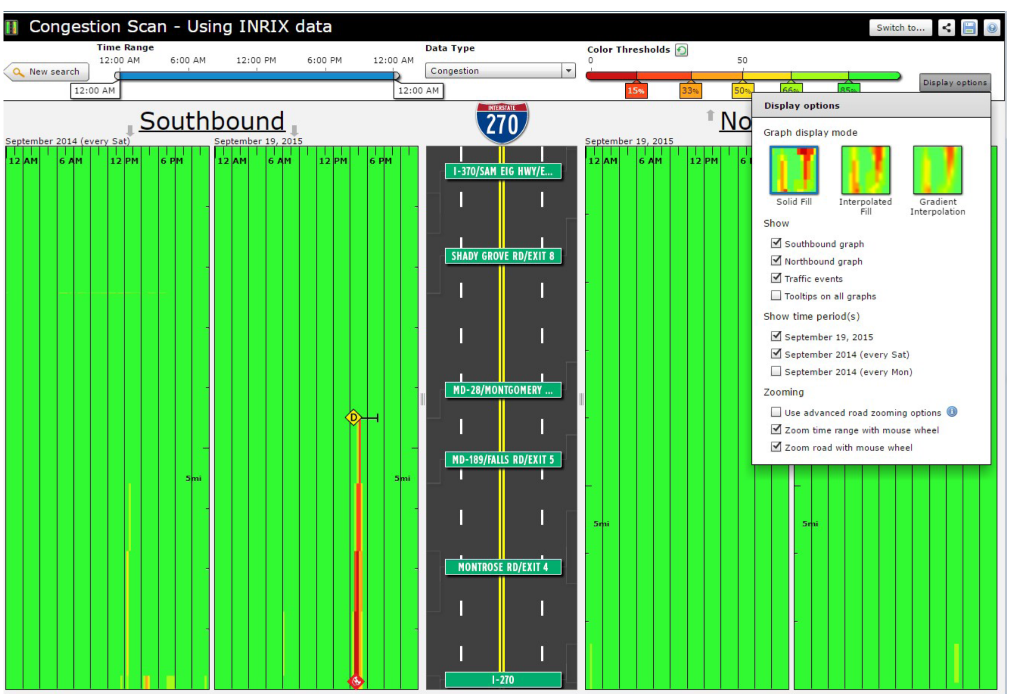
- The congestion on that stretch of road for that specific day of the week (Figure 51).
- The congestion on that stretch of road on that specific day compared to similar days that same month (Figure 52).
- The estimated user delay cost on that road and that direction of travel by hour of day.



Source: Probe Data Analytics Suite. University of Maryland CATT Laboratory.

Figure 51. Screenshot. Congestion scan graphic of outlier congestion showing both north- and southbound traffic on I-270.

There is unusual congestion on a Saturday evening in the southbound direction. Incident overlays show that this was caused by an injury incident (bottom of the screen), and that a vehicle also became disabled towards the back of the queue.



Source: Probe Data Analytics Suite. University of Maryland CATT Laboratory.

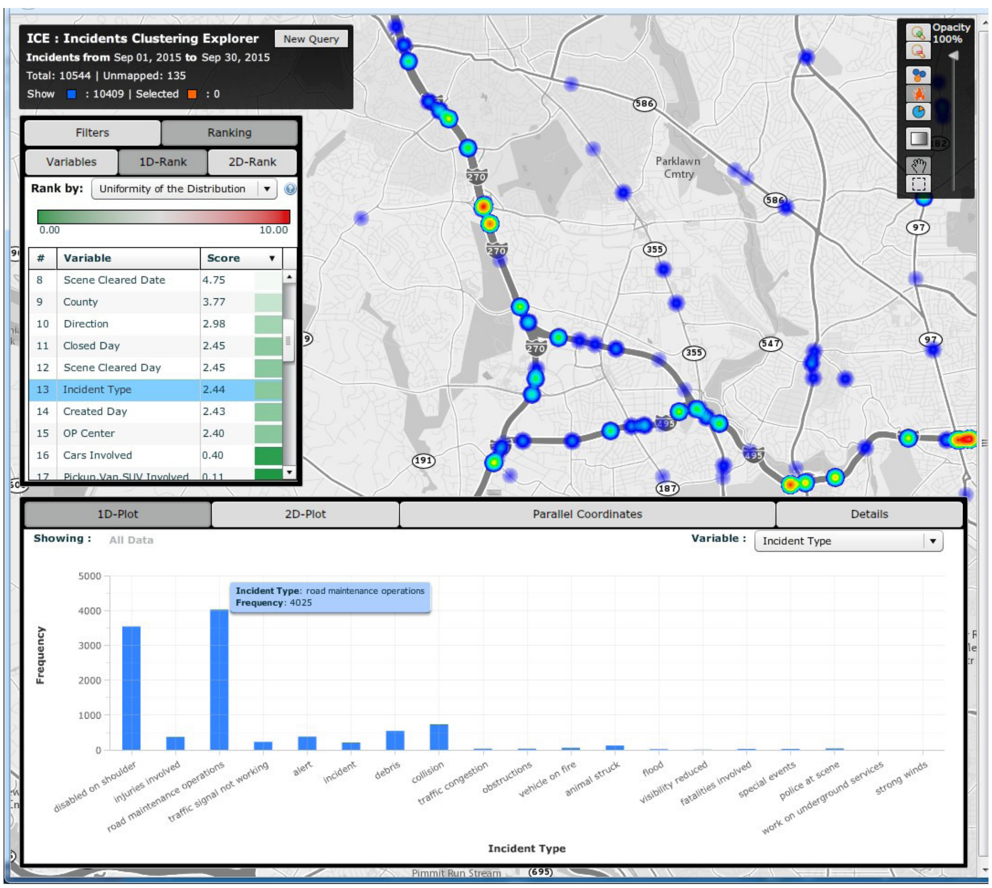
Figure 52. Screenshot. Congestion scan graphic including Saturdays.

Editing the display options, the analyst has now chosen to add an additional date range to the graphic. All Saturdays during the month of September are also drawn on the screen in the outside columns of data. Note that there is little or no congestion "normally" on Saturdays in the southbound direction. This analysis confirms that the congestion and incidents on September 19 (innermost columns) are non-recurring.

The agency can use other analytic tools found within the RITIS platform to review safety data from police accident reports or ATMS event/incident reports to determine if incident hot spots exist that may be globally affecting congestion on the roads in question. Some of these tools allow extremely deep querying and causality analytics. Figure 53 shows how heat maps can be drawn overtop of the roadway to show when and where collisions, weather events, or construction projects are most likely to occur for any month, day of week, or other time range during the year.

Figure 54 shows how the tool also can be used for deep analysis (e.g., correlation of coefficients, standard deviation of frequencies, number of outliers, maximum frequency, number of unique pairs, uniformity of the distribution) to further identify patterns, trends, and outliers within the data. The correlation of coefficients ranking function can be used to verify if any pair of variables within the data set might be affecting each other (e.g., does a wet road highly correlate with fatal collisions, or does the day of the week in which an incident occurs correlate with hazmat incidents or the number of vehicles involved?).

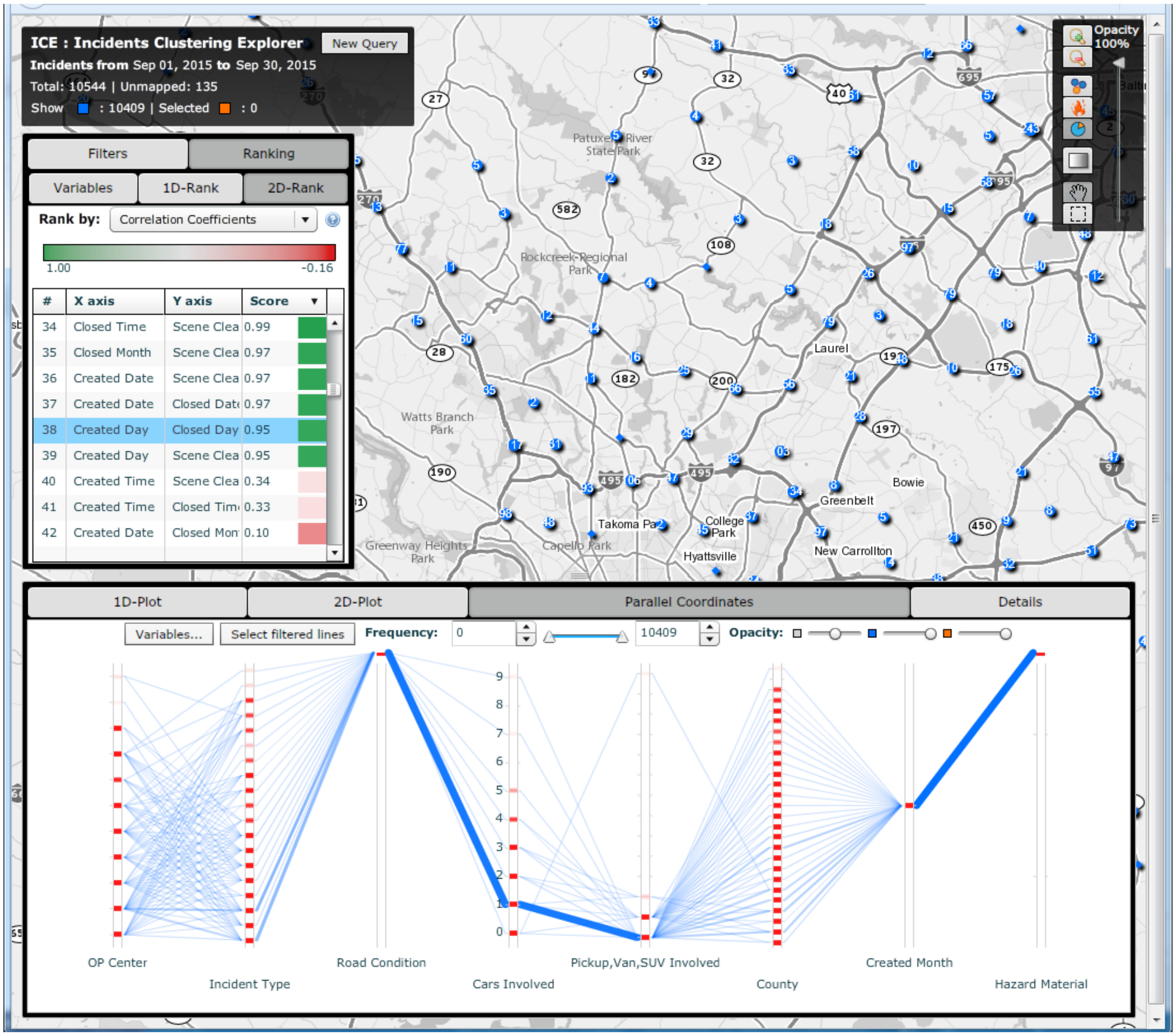
In Figure 53, the heat maps in the upper right show where incidents have occurred over a given date range. A list of variables in the upper left is clickable, filterable, and interactive. Selecting a particular variable generates a histogram (bottom) that shows the breakdown of possible values for that variable. Here, the user has clicked on "incident type,"



Source: Incident Cluster Explorer, University of Maryland CATT Laboratory.

Figure 53. Screenshot. Heat maps function within incidents clustering explorer.

which results in an interactive histogram depicting the number of each incident type occurring during the particular date range. Clicking on each bar of the histogram then highlights those incidents on the map above.



Source: Incident Cluster Explorer, University of Maryland CATT Laboratory.

Figure 54. Screenshot. Correlation of coefficients ranking function in incidents clustering explorer.

In Figure 54, the correlation of coefficients ranking function in the upper left allows a user to compare variable pairs to determine if any correlation exists. Clicking on any of these pairs of variables then brings up detailed two-dimensional histograms (not shown) that make it easy for users to do a deeper dive into the various relationships, trends, and statistics related to those variables.

Impacts and Lessons Learned

This type of analysis has historically been difficult for MD SHA to conduct, as their data sets have been unavailable. Creating a central data repository and developing tools that make it faster and more intuitive to ask questions has dramatically improved productivity and understanding within the organization. Historically, problems were identified through laborious studies that often involved dedicated data collection. These studies may have only been initiated if someone within the community complained enough about the congestion. Presently, with more ubiquitous data from probes and the integration of incident and event data from both law enforcement and the ATMS platforms, the agencies are able to identify problems, causality, and overall impact before the public raises concerns, and better respond when they do.

In addition to answering these questions more quickly, NJDOT is more confident in the results of their studies. The ease of access to the data, coupled with the compelling visualizations, result in analysts being able to effectively communicate with different audiences.

▸ *Case Study 2: Metropolitan Area Transportation Operations Coordination Examines Effects of Events, Incidents, and Policy Decisions on Congestion*

Overview

Unlike the case study above, which first identified heavily congested locations and then tried to determine causality, this case study reviews how to examine the effect of events, incidents, or policy decisions on congestion.

The Metropolitan Area Transportation Operations Coordination (MATOC) program is a coordinated partnership between transit, highway, and other related transportation agencies in the District of Columbia, Maryland, and Virginia, that aims to improve safety and mobility in the region through information sharing, planning, and coordination. MATOC is frequently called upon to help agencies perform after action reviews of major incidents, produce impact statements post-event for the media, or help agencies measure the impact of certain operations decisions (e.g., closing the Federal government due to winter weather, closing certain roads, responding in a certain way to an event).

This case study shows how MATOC used the archived operations data from the agencies that they support to help measure the impacts of delaying the opening of schools and the Federal government by 2 hours.

Data

The data for this case study are the same as that used for the case study discussion above; however, there is a greater emphasis on weather data—both RWIS and weather radar from the National Weather Service.

Agency Process

MATOC uses a series of tools found within RITIS for analyzing weather impacts, user delay cost, and travel times.

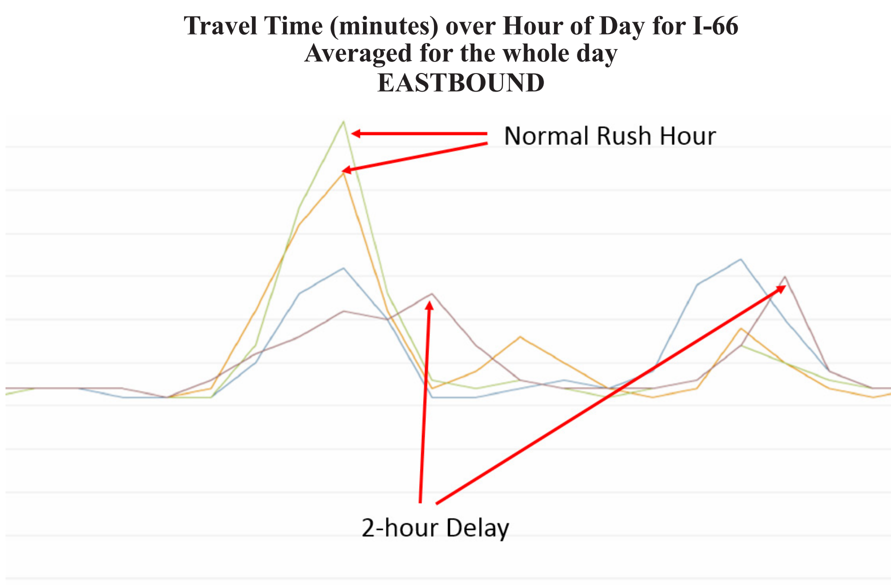
After a winter weather event ends, MATOC is usually contacted by various groups (e.g., the media, transit agencies, DOTs, MPOs, legislators) to help answer questions about the impacts of a particular event on the region. MATOC follows this process to develop graphics, statistics, and reports:

1. Choose a corridor of significance that will help tell a compelling story. The corridor must be of regional significance and a general pain point for commuters.
2. Use the probe data analytics suite or the detector analytics tools to run travel time queries for a common commute route on the corridor of significance.
3. Use the RWIS explorer and/or the other historic weather tools within RITIS to determine which dates did or did not have winter weather, other precipitation, construction, or other incidents.
4. Run a user delay cost query to determine the financial impacts of the winter weather event on the corridor (or the entire region, if needed).
5. Develop an animated trend map that shows traffic conditions during the winter weather event compared to normal days without events.
6. Compile all of the outputs listed above into a report or embed the results online using embedding and exporting support tools within the RITIS analytics.

Below are the outputs of each of the steps listed above.

Steps 1 and 2

In Step 1, the MATOC analysts determined that the I-66 corridor leading into Washington D.C. would be a well-known corridor of significance. A 20-mile section of the corridor was selected, and a travel time query was run for the day of the 2-hour delayed opening. Simultaneously, a query was run for two similar days of the week during the same season. Figure 55 is the result of this query. Note that the travel time during normal rush hours is nearly double that of the travel times during the 2-hour delayed opening. Two hours later, a slightly larger rush hour occurred. Note that, in the evening, the rush hour travel times are nearly the same as the normal rush hour; however, this rush hour is only 1 hour later than normal instead of 2 hours later. Therefore, it appears that most individuals went to work 2 hours later than normal but only stayed 1 hour later.



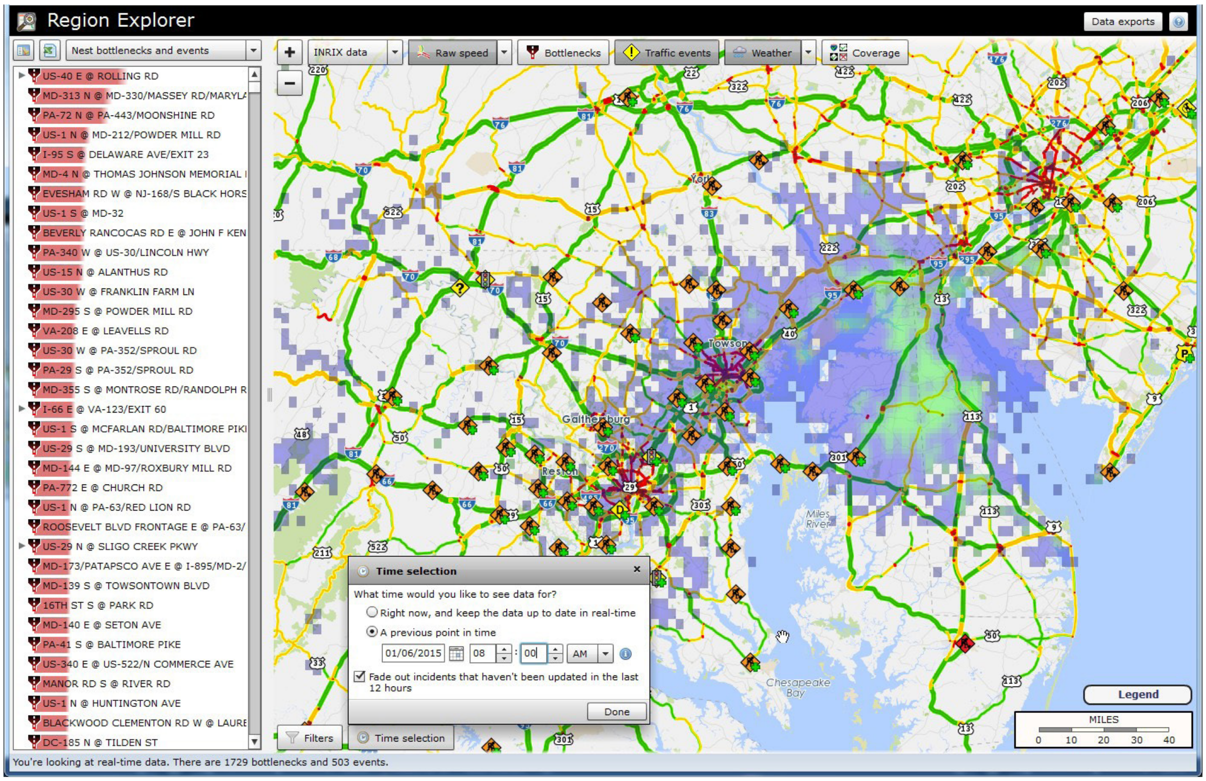
Source: Metropolitan Area Transportation Operations Coordination (MATOC) Program Facilitator, Taran Hutchinson.

Figure 55. Graph. Plot of travel time by hour of day for I-66, created via Regional Integrated Transportation Information Systems Vehicle Probe Project Suite.

Figure 56 depicts travel times along I-66 during normal rush hours compared to travel times when there was a winter weather event that resulted in a 2-hour delayed opening for most schools and the Federal Government.

Step 3

To ensure that the analysis is based on accurate data, MATOC should verify that the other date ranges for which they are comparing the winter weather event are clear of other events (weather or incidents). This can be done via the probe data analytics suite by overlaying historic National Weather Service radar data and historic incident/event data on top of any date/time in the past (Figure 56). This visually confirms that a date or date range is safe for inclusion in the analysis.

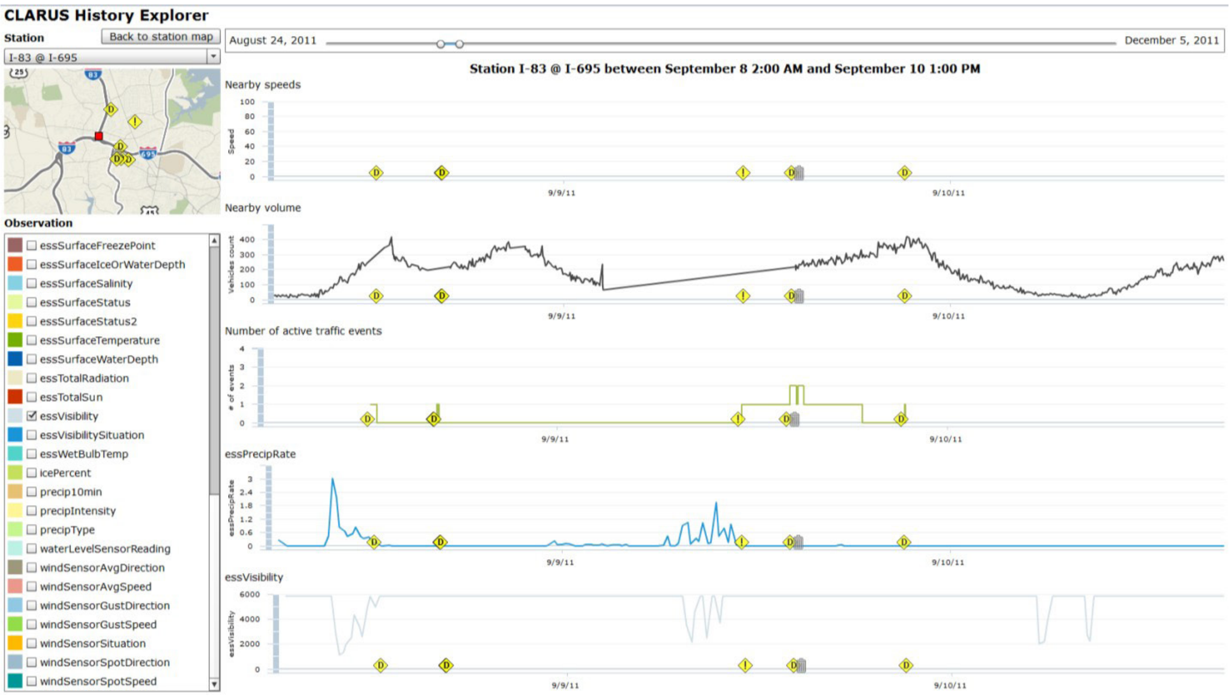


Source: University of Maryland, Probe Data Analytics Suite.

Figure 56. Screenshot. Historical dates and times to visualize weather and incident details during that date and time.

In the tool shown in Figure 56, after the date is entered, the map redraws itself to show what was happening at that particular time including: weather radar, incidents and events, bottlenecks, and speeds. The left side of the screen shows a ranked list of all of the bottlenecks that were occurring at that time.

Alternatively, MATOC can look at RWIS station data (if there is enough coverage) to see if road weather may have impacted travel conditions. This particular analytics tool from RITIS (seen in Figure 57) allows MATOC to select any date range and view weather conditions, such as precipitation rates, visibility, and surface conditions, while simultaneously viewing incident/event data, and flow data. This tool helps evaluate the effects of all forms of weather on congestion and incidents. This allows users to compare data from road weather information systems stations (e.g., precipitation rates, visibility, wind speeds, surface conditions) to other archived data (e.g., speeds, volumes, flow, incidents, events).

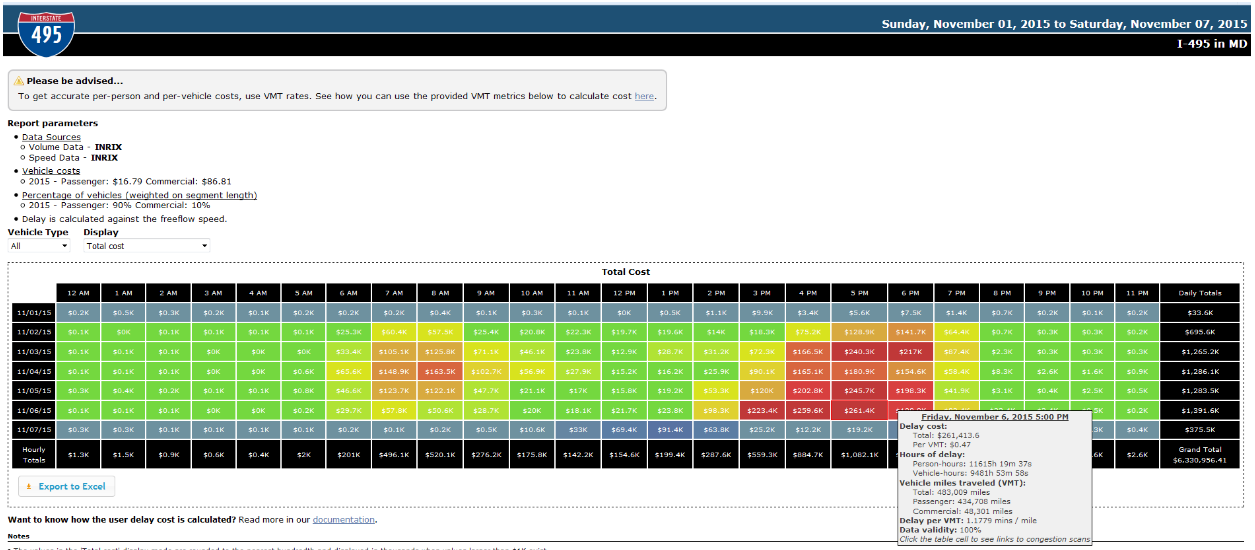


Source: Clarus History Explorer: University of Maryland CATT Laboratory.

Figure 57. Screenshot. The road weather information systems history explorer.

Step 4

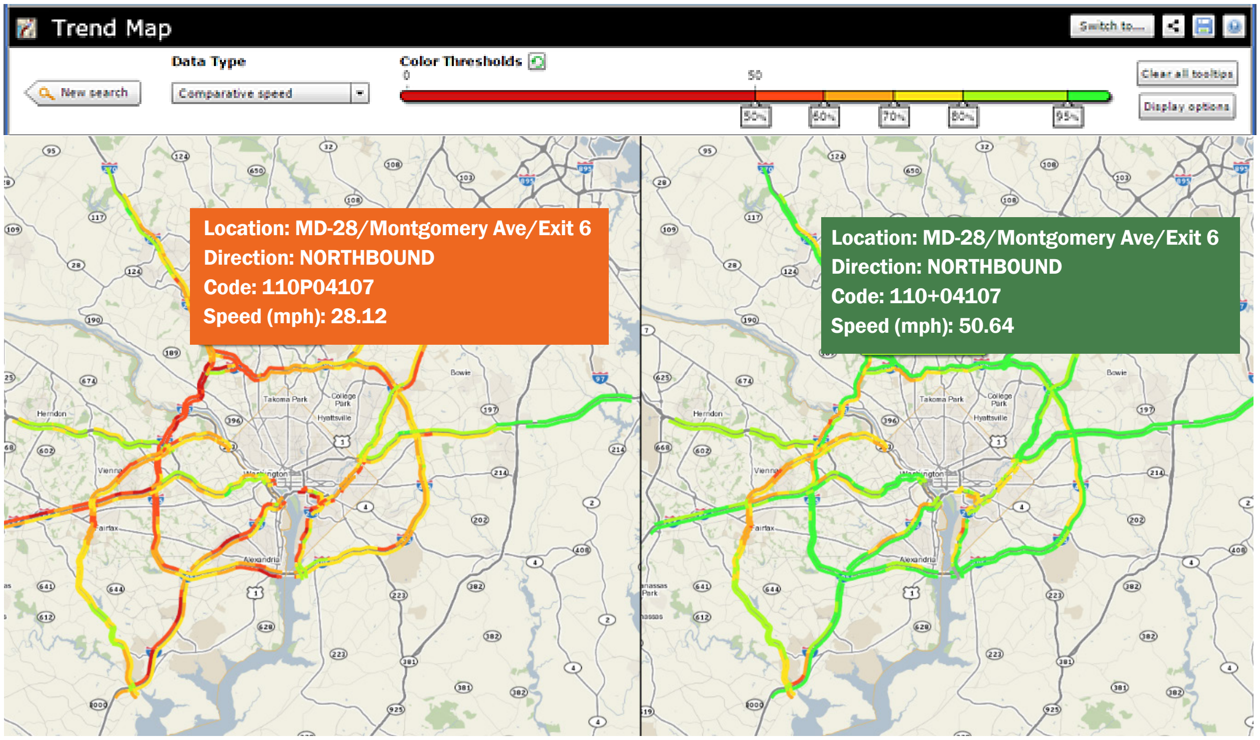
In Step 4, MATOC runs a user delay cost query on the I-66 corridor to determine the financial impacts of the winter weather compared to normal user delay on the same corridor. Figure 58 shows the resulting user delay for the morning snow event discussed in this case study. Note how much higher the user delay is during the morning of the snow event compared to normal.



	12 AM	1 AM	2 AM	3 AM	4 AM	5 AM	6 AM	7 AM	8 AM	9 AM	10 AM	11 AM	12 PM	1 PM	2 PM	3 PM	4 PM	5 PM	6 PM	7 PM	8 PM	9 PM	10 PM	11 PM	Daily Totals
11/01/15	$0.2K	$0.5K	$0.3K	$0.2K	$0.1K	$0.2K	$0.2K	$0.2K	$0.4K	$0.1K	$0.3K	$0.1K	$0K	$0.5K	$1.1K	$9.9K	$3.4K	$5.6K	$7.5K	$1.4K	$0.7K	$0.2K	$0.1K	$0.2K	$33.6K
11/02/15	$0.1K	$0K	$0.1K	$0.1K	$0.1K	$0.1K	$25.3K	$60.4K	$57.5K	$25.4K	$20.8K	$22.3K	$19.7K	$19.6K	$14K	$18.3K	$75.2K	$128.9K	$141.7K	$64.4K	$0.7K	$0.3K	$0.3K	$0.2K	$695.6K
11/03/15	$0.1K	$0.1K	$0.1K	$0K	$0K	$0K	$33.4K	$105.1K	$125.8K	$71.1K	$46.1K	$23.8K	$12.9K	$28.7K	$31.2K	$72.3K	$166.5K	$240.3K	$217K	$87.4K	$2.3K	$0.3K	$0.3K	$0.3K	$1,265.2K
11/04/15	$0.1K	$0.1K	$0.1K	$0K	$0K	$0.6K	$65.6K	$148.9K	$163.5K	$102.7K	$56.9K	$27.9K	$15.2K	$16.2K	$25.9K	$90.1K	$165.1K	$180.9K	$154.6K	$58.4K	$8.3K	$2.6K	$1.6K	$0.9K	$1,286.1K
11/05/15	$0.3K	$0.4K	$0.2K	$0.1K	$0.1K	$0.8K	$46.6K	$123.7K	$122.1K	$47.7K	$21.1K	$17K	$15.8K	$19.2K	$53.3K	$120K	$202.8K	$245.7K	$198.3K	$41.9K	$3.1K	$0.4K	$2.5K	$0.5K	$1,283.5K
11/06/15	$0.1K	$0.1K	$0.1K	$0K	$0K	$0.2K	$29.7K	$57.8K	$50.6K	$28.7K	$20K	$18.1K	$21.7K	$23.8K	$98.3K	$223.4K	$259.6K	$261.4K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0.2K	$1,391.6K
11/07/15	$0.3K	$0.3K	$0.1K	$0.1K	$0.1K	$0.1K	$0.2K	$0.1K	$0.2K	$0.5K	$10.6K	$33K	$69.4K	$91.4K	$63.8K	$25.2K	$12.2K	$19.2K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$0.4K	$375.5K
Hourly Totals	$1.3K	$1.5K	$0.9K	$0.6K	$0.4K	$2K	$201K	$496.1K	$520.1K	$276.2K	$175.8K	$142.2K	$154.6K	$199.4K	$287.6K	$559.3K	$884.7K	$1,082.1K	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	$2.6K	Grand Total $6,330,956.41

Source: University of Maryland Probe Data Analytics Suite.

Figure 58. Screenshot. Visual representation of the cost of delay for both passenger and commercial vehicles.



Source: University of Maryland CATT Laboratory.

Figure 59. Screenshot. An interactive animated map showing conditions during a winter weather event (left) compared to conditions during normal days of the week (right).

Step 5

Here MATOC runs a similar query to that in Steps 1 and 2; however, the output is an animated trend map instead of a simple line chart. The animated trend map shows travel conditions on additional corridors and shows a side-by-side comparison of travel speeds during the event compared to normal (Figure 59).

Step 6

Finally, the various statistics and visualizations previously mentioned are consolidated into a correspondence with a narrative description of the impacts of the event surrounding it. This narrative details the associated impacts of this weather event.

Impacts and Lessons Learned

MATOC's ability to efficiently produce these types of reports has allowed them to gain priority ranking for after action analysis and media requests, which helps secure their annual operating budget. MATOC now trains other agencies on how to perform their own analysis using these tools.

6. GETTING STARTED

Any transportation organization can and should begin to make use of their data. Even with extremely limited data, staff can begin to make better planning and programming decisions. This chapter provides concrete steps and information resources that transportation planners and their planning partners can use to begin making better use of existing archived operations data.

Transportation planners and their partners should review the following checklist when they decide to begin using archived operations data:

Table 12. Checklist for getting started.

Step #	Action	Explanation
1	Define your needs/ questions.	What is it that you're trying to do? What questions do you need to answer?
2	Inventory your data.	This can be the hardest step. Ask each department in your agency if they have any of the data sources listed in Table 1. Create a list of all of the data that you believe you can get, from whom the data will come, and any other relevant attributes (metadata) related to what is available from within the dataset.
3	Did you find the data you needed?	If yes, then skip to step #6, if not, then continue through this list.
4	If not, look elsewhere.	If your own agency does not seem to have data (or there are political reasons why it cannot be accessed), consider reaching out to partner agencies (local department of transportations (DOTs), metropolitan planning organizations (MPOs), and university partners). There are also a number of free or third party data sets that are available as shown in Table 15, 16, and 17. Consider exploring your regional intelligent transportation system (ITS) architecture to identify opportunities for obtaining data in the region from other systems or agencies. This is one of the benefits of a maintained ITS architecture.
5	Purchase data.	If free data sets are not available, you may need to procure data from a third party. Consult Chapter 4, section "Obtaining Third-party Data and Tools" for important considerations.

Table 12. Checklist for getting started. (Continued)

Step #	Action	Explanation
6	Is the data usable?	Just because data exists, does not mean that it is usable. Is it a scan of a paper document or is it truly electronic, machine readable data? Make sure it is of the right spatial and temporal resolution and in a format that you, your team, and your systems can use.
7	Do you trust the data?	While no dataset is perfect, it is important to understand where the data came from, how it may have been massaged or filtered, and if there are any data stewards that can attest to the overall quality of the dataset.
8	Determine your analysis process, and get to work!	Will you analyze the data with in-house personnel? Will you purchase data analysis tools to help with the analysis? Will you use consultant support? Later sections of this chapter give guidance on how to work with third party data analysts. Another section below also discusses what types of skills you might need internally for successful data analysis.

My Agency Does Not Have Any Data

If you followed the checklist above and have realized that your agency is not producing any of its own data (or there are political or technical impediments that make it too difficult to access that data), there are still other resources at your disposal. Some of these are free, while others are not. Some data are considered "raw," while other data is accompanied by readily available tools for analysis.

Examples of datasets include:

- National Performance Measures Research Data Set (NPMRDS).
- National Weather Service Data.
- Google/Waze Speed and Incident Data.
- Third Party Probe-based Speed Data.
- Third Party Origin-Destination Data.
- Partner Agency Data.
- Fatality Analysis Reporting System (FARS) Data.

These data are described in greater detail in the following tables.

Table 13. The National Performance Measurement Research Data Set (NPMRDS).

Overview			
The NPMRDS was purchased by Federal Highway Administration (FHWA) originally for internal purposes, but it has been made available to State and local transportation agencies at no cost. It will likely be used by agencies to develop Fixing America's Surface Transportation (FAST) Act congestion-related performance measures. However, there are other potential uses, including problem identification and after action review. The first iteration of NPMRDS is an unfiltered, unsmoothed data set that has gaps in temporal coverage and only covers the National Highway System (NHS). It is only available as an historic data set, made available at the end of the each month. Therefore, it has few real-time applications.			
What is it	**What it Costs**	**Access**	**Tool Availability**
Historic speed and travel time data for the NHS. Procured by the FHWA; officially referred to as the NPMRDS.	Free	Send an email to heretraffic.nhs.data@here.com to receive official instructions. More information at: http://ops.fhwa.dot.gov/freight/freight_analysis/performan_meas/vpds/npmrdsfaws.htm#2. Data can be accessed through several third-party tools, including Regional Integrated Transportation Information System [RITIS] (https://npmrds.ritis.org/analytics/) and Iteris (http://www.iteris.com). Other tools may be online.	Free tools are available at https://npmrds.ritis.org/analytics/.

Table 14. The National Weather Service Data.

Overview			
Historical weather data can be used to identify causes of congestion as well as area for improvement in road weather management activities.			
What is it	**What it Costs**	**Access**	**Tool Availability**
A host of weather products including radar, images, cloud cover, road weather information systems (RWIS), forecasts, alerts, and much more.	Free	http://www.nws.noaa.gov/tg	Yes. The National Weather Service has its own host of tools available for looking at real-time and historic data. Many third-party companies also provide tools to analyze weather data. Some free tools exist (e.g., https://weatherspark.com/). There are many other historic weather data tools available.

Table 15. Google/Waze speed deviation and incident data.

Overview			
In addition to real-time operations uses, the archived data can be used for performance measures generation, safety data analysis, various congestion studies, and problem identification. It is especially useful on roads where coverage by the agency may be limited (like on arterials).			
What is it	**What it Costs**	**Access**	**Tool Availability**
Crowd-sourced incident and event data plus speed variation data on segments of the roadway.	Free (with the Connected Citizens Partnership Program).	Google/Waze has a relatively new public sector data sharing partnership. At the time of publication, Google/Waze was brokering deals with agencies to share their crowd-sourced data with the public sector if the public sector was willing to provide their own operations data back to Google and/or the agency agrees to provide attribution to Waze or promote Waze on agency 511 websites or other materials.	University and private sector data fusion specialists are working on tools (similar to RITIS) that will allow agencies to look at trends in congestion and incidents from the Google data sets.

Table 16. Third party probe-based speed data.

Overview			
Probe-based speed data can be used to: develop system performance reports; identify problems; prioritize projects; perform after-action incident review; conduct before and after studies; make informed, real-time operations decisions; analyze travel time and reliability; monitor work zones; develop and publish press releases for public and media consumption; measure the economic and environmental impacts of passenger and commercial vehicle user delay.			
What is it	**What it Costs**	**Access**	**Tool Availability**
Speed and travel time data on Interstates and major and minor arterials from HERE, INRIX, TomTom and others.	Varies depending on the provider. Usually a per-mile cost; however, some companies charge only for major roads and include smaller arterials for free.	Directly from data providers, through application program interfaces (APIs), or through third-party data analytics like Iteris Performance Monitoring System (iPeMS) or Regional Integrated Transportation Information System (RITIS) Probe Data Analytics Suite.	Yes. Some of the probe data providers offer their own analytics tools. RITIS and iPeMS also offer analytics suites for probe data.

Table 17. Third party origin-destination data (O-D).

Overview			
This data usually comes from third-party probe data providers and includes everything from basic turning movements at intersections all the way to direct origins and destinations of vehicles moving throughout a network. O-D data can be used to track travel trends and calibrate models to support informed investment decisions.			
What is it	What it Costs	Access	Tool Availability
Information related to the movement of people through the trips they take, including their origin and destination for each trip.	Costs depends heavily on geography and date range.	From third-party data providers like INRIX, Airsage, and others.	Yes. The companies that sell these data usually offer analytics tools. Some universities are also building tools along with the RITIS platform.

Table 18. Partner agency data.

Overview			
Partner agency data is any and all data that you can get from outside of your own agency. For example, if you are a metropolitan planning organization (MPO) and do not have any of your own data, you may be able to request data directly from State or local department of transportations (DOTs), law enforcement, or other agencies in your region. Partner agency data could include any of the data mentioned previously in this document (speeds, volumes, incidents, weather, etc.) and can be used for all of the things that have been mentioned under any of the case studies that have been presented, plus any of the other relevant uses listed in the above data sets.			
What is it	What it Costs	Access	Tool Availability
Anything. Incident data, speeds and volumes, computer-aided dispatch, signal phasing and timing, transit schedules, and on-time performance, etc.	Usually free, but it depends on the agency/organization for which you are asking.	Work with other agencies in a region MPOs, State and local DOTs, law enforcement, etc.) to request copies of databases. In rare instances, these agencies may have already invested in tools that they may allow to be borrowed.	Maybe. It depends on each agency and the investments that they have already made in data and analytics.

Table 19. Fatality Analysis Reporting System (FARS) data.

Overview			
This dataset can be used to analyze fatality data from motor vehicle traffic crashes. It can be used to identify safety issues on roadways.			
What is it	What it Costs	Access	Tool Availability
Database of all fatal collisions in the United States.	Free.	http://www.nhtsa.dot.gov/Main/index.aspx	Yes. The FARS website has basic tools available for querying. Other universities and private-sector companies have developed additional web-based tools for analyzing FARS data. Some of these tools also incorporate non-FARS data.

Partnering with other organizations is also a great way to get access to additional data sets and tools that enable personnel to analyze the data. If an agency is able to leverage investments that others have already made in data collection and analysis, it is advisable to do so.

My Agency Has Limited Data or the Data is in Silos

Prior chapters of this report have provided examples of how various types of operations data are being used. Table 20 below gives an overview of what can be done with limited data sets (or data that exists in silos).

Table 20. Options for agencies with limited amounts and types of data.

Data	What You Can Do With It
Traffic volume, speed, class, and occupancy from point sensors	Travel time analysis, model calibration.
Event, work zone, and incident information from an ATMS	Safety analytics, support of before and after studies, identification of effective operations strategies.
Weather data	N/A when by itself.
Traffic speeds from probe-based data sources	• System performance reports. • Problem identification/project development. • Before and after studies of construction projects, signal retiming, etc. • Travel time analysis including reliability measurements.
Transit data	Schedule adherence and fare collection can be used for identifying transit accessibility in communities and for planning for new stops, schedules, etc.
Signal data	Intersection and ramp meter signal data (timing plans, maintenance, etc.) can be used for arterial systems performance measures, signal retiming studies, and more.

Table 20. Options for agencies with limited amounts and types of data. (Continued)

Data	What You Can Do With It
Police crash reports	Safety analytics and model calibration.
Freight movements	Model calibration.
Origin-Destination data	Useful for many planning activities that require an understanding of travel patterns for freight and passenger vehicles. Also useful for real-time operations when trying to measure the impact of various traveler information strategies and the impact of incidents on arterials and other secondary roads.

Advanced Transportation Management System (ATMS) Data Example

Most State DOTs collect some form of incident or event data through their ATMS or traveler information platforms (e.g., 511). These data usually include (at a minimum) the location of an event, the type of the event, and the lane status (e.g., the number of lanes closed). In other agencies, these data might include information about which responders are on the scene, how long they took to get to the scene, which variable message signs were used to provide information, and much more.

Minimally, these data can and should be leveraged to improve operations, but, more importantly, to justify budgets and help to effect change within the organization. For example, a quick geospatial and temporal query can help an agency to understand when and where certain types of incidents are most likely to occur. The agency can then use that data to make better decisions about how many safety service patrols are needed for each shift, where they should be positioned, and what types of equipment might be needed on the vehicles to better tend to the types of events they respond to on a daily basis.

Going a few steps further, planners and operations agencies can use these data to:

- Identify which responders are arriving on the scene quicker.
- Identify problem locations that might need countermeasures to reduce the occurrence of incidents and events.
- Justify the need for additional responders or operators during certain shifts.
- Justify budgets.

My Agency Has Lots of Data

For agencies that are fortunate enough to have many disparate data sources at their disposal, blending and fusion concepts can and should be explored to attempt to get more out of the archived operations data sets. Creating links between data sets so that they can function together as a much larger database can make possible the use of more sophisticated analytics—searching for cause and effect, correlation of weather to safety issues, and others.

If a vast array of data is indeed available, big data techniques can allow an agency to enter the world of multi-variate data analytics, correlations, what-if scenarios, prediction, and more. Big data analytics are less about the size of the data sets than they are about a process for searching through multi-variate data sets to search for context, correlation, and other aspects that the agency may not have previously thought to ask. The goal is to extract value from the data more quickly and easily than traditional data analysis.

Numerous business and university groups specialize in big data concepts and data science. One example of applying big data analytics is the problem of determining the contribution of underlying factors to total congestion. Congestion and unreliable travel are caused by the interaction of several factors: physical capacity, demand, incidents, weather, work zones, traffic control devices, and operating policies. Assigning how much each of the factors contributes to measures like total delay is problematic, but big data analytics can help.

Other examples of what can be derived from the fusion of multiple data sets are shown in Table 21 below. Ultimately, as the size, quality, and diversity of data continue to grow, the potential for new and innovative data uses will be discovered.

Table 21. Options for agencies with extensive amounts and types of data.

Data	What You Can Do With It
Volumes and probe-based speed data.	User delay cost, emissions, fuel consumption, model calibration.
Advanced transportation management systems (ATMS), crash reports, speeds and travel times, and weather.	Bottleneck analytics, causality, correlation of incident management practices with reductions in delay and incident severity, and the benefit/costs of various strategies.
Signal data (SPaT and sensor actuation), Bluetooth re-identification or origin-destination data, volumes.	True arterial performance measures including travel times between intersections, travel time reliability, capacity utilization at intersections, true integrated corridor management, safety applications, dynamic signal retiming, what-if scenario planning, and more.

Before starting your analysis, consider the capabilities and availability of your in-house resources. Is there a need to hire new staff? What is the expected lead time for them to accomplish your goal? If it is lengthy, consider the costs associated with delays and consider purchasing existing tools or services. Leverage the investments that others have already made. Whether you are looking to hire in-house staff or are looking to hire university/consultant support, the following skills and job descriptions should be considered:

- Database administrator.
- Data scientist/data analyst.
- Information visualization.

Ultimately, if resources are lacking and there is neither the time nor ability to hire new staff, consider outsourcing to meet these needs. Alternatively, you can talk to other agencies to determine what tools they are using—they might be willing to share or offer guidance. Try to leverage their investments, if possible.

Guidance on Procuring Data and Analytics Services

Many agencies will struggle to analyze complex, disparate data from operations groups. When this is the case, external support from universities or consultants, and/or the purchase of tools, may be necessary.

Whenever an agency procures either a system or a set of services, a certain amount of rigor is needed to ensure the agency's needs are communicated properly to bidders. When developing requirements for consultant support and/or tools, two approaches are typically used:

1. Option 1: A detailed set of system requirements are issued at the beginning of the contract to identify the functions and top-level design of the system. This approach requires a major effort on the part of the agency, and can become expensive and frustrating if new requirements are identified midway through the initiative or the requirements were not thought through entirely.
2. Option 2: The second approach is similar to the "design-build" methodology in construction. In this option, the requirements are developed as part of the contract prior to development beginning. This method can allow for greater dialogue between the bidders and the agency which ultimately can lead to a better end product.

Both of the options have their advantages; however, the design-build approach typically results in an end-state that is more favorable to the agency. Evaluate external service providers (consultants in the private sector or universities). Check their references and project portfolios.

If the decision is made to hire a consultant or a university, distinguish between a one-time study using available data and the development of a system for ongoing use by the agency. For one-time analysis projects, make sure to write-up a detailed scope of work that describes:

1. The data available (be as detailed as possible and be ready to provide samples for evaluation).
2. The questions that need to be answered.
3. Expected deliverables (broken down into phases that ensure success).
4. Desired skills.
5. Avoid specifying technologies or techniques. Allow the consultant team to suggest an approach to solving the problem.
6. Any other expectations.
7. If possible, allow the consultant to suggest alternative approaches and value-added services. This will ensure that the best that each consultant has to offer will be available.

Purchasing Services

For agencies that are not comfortable using analytic tools, and are not interested in performing in-house data analysis, hiring outside consultant support may prove to be a viable option. Purchasing support from universities is a similar option. Consultants and universities frequently have access to scientists, statisticians, database programmers, economists, and other analysts that would otherwise be difficult to hire at State and local agencies. When seeking out-of-agency services, it is wise to review product and project portfolios for examples of prior work to ensure an agency's needs match the skills of the consultant or university personnel being proposed on a project.

When hiring outside support (consultant or universities), consider a phased approach to projects. Start small, and ensure the consultant is able to perform basic analysis and fusion tasks with the data available. If the consultants are successful, then work can progress on bigger analysis tasks—adding layers of complexity and building off of prior work and available data sets. Initiating extremely large analysis tasks that are not easily broken down into smaller deliverables can be a recipe for confusion, cost overruns, disappointment, and waste.

Regardless of who does the work, it is advisable to avoid demanding that consultants use specific tools, technologies, or techniques to deliver a solution. New technologies, methodologies, and tools are developed quickly and often, and requiring outdated technologies can result in unnecessarily limiting the agency and the consultant in performing analytical tasks. Allow the consultants to drive these decisions based on what they perceive to be the most efficient and effective tools and methods.

U.S. Department of Transportation
Federal Highway Administration

U.S. Department of Transportation
Federal Highway Administration
Office of Operations
1200 New Jersey Avenue, SE
Washington, DC 20590
Toll-Free "Help Line" 866-367-7487
www.ops.fhwa.dot.gov

December 2016

FHWA-HOP-16-082