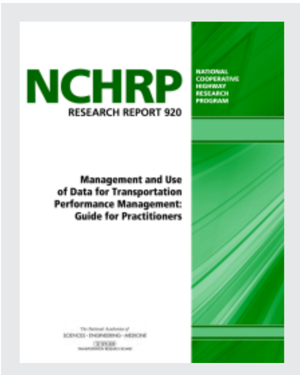


Management and Use of Data for Transportation Performance Management: Guide for Practitioners (2019)

DETAILS

148 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-49328-4 | DOI 10.17226/25462

CONTRIBUTORS

Frances D. Harrison, William Duke, Juliet Eldred, Michael Pack, Nikola Ivanov, Joe Crosset, and Larry Chan; National Cooperative Highway Research Program; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine

SUGGESTED CITATION

National Academies of Sciences, Engineering, and Medicine 2019. *Management and Use of Data for Transportation Performance Management: Guide for Practitioners*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25462.

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

NCHRP RESEARCH REPORT 920

Management and Use of Data for Transportation Performance Management: Guide for Practitioners

Frances D. Harrison
William Duke
Juliet Eldred
Spy Pond Partners, LLC
Arlington, MA

in association with

Michael Pack
Nikola Ivanov
MLP, LLC
College Park, MD

Joe Crosset
Larry Chan
High Street Consulting Group
Pittsburgh, PA

Subscriber Categories
Highways • Data and Information Technology

Research sponsored by the American Association of State Highway and Transportation Officials
in cooperation with the Federal Highway Administration

The National Academies of
SCIENCES • ENGINEERING • MEDICINE
TRB
TRANSPORTATION RESEARCH BOARD
2019

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

Systematic, well-designed, and implementable research is the most effective way to solve many problems facing state departments of transportation (DOTs) administrators and engineers. Often, highway problems are of local or regional interest and can best be studied by state DOTs individually or in cooperation with their state universities and others. However, the accelerating growth of highway transportation results in increasingly complex problems of wide interest to highway authorities. These problems are best studied through a coordinated program of cooperative research.

Recognizing this need, the leadership of the American Association of State Highway and Transportation Officials (AASHTO) in 1962 initiated an objective national highway research program using modern scientific techniques—the National Cooperative Highway Research Program (NCHRP). NCHRP is supported on a continuing basis by funds from participating member states of AASHTO and receives the full cooperation and support of the Federal Highway Administration, United States Department of Transportation.

The Transportation Research Board (TRB) of the National Academies of Sciences, Engineering, and Medicine was requested by AASHTO to administer the research program because of TRB's recognized objectivity and understanding of modern research practices. TRB is uniquely suited for this purpose for many reasons: TRB maintains an extensive committee structure from which authorities on any highway transportation subject may be drawn; TRB possesses avenues of communications and cooperation with federal, state, and local governmental agencies, universities, and industry; TRB's relationship to the National Academies is an insurance of objectivity; and TRB maintains a full-time staff of specialists in highway transportation matters to bring the findings of research directly to those in a position to use them.

The program is developed on the basis of research needs identified by chief administrators and other staff of the highway and transportation departments, by committees of AASHTO, and by the Federal Highway Administration. Topics of the highest merit are selected by the AASHTO Special Committee on Research and Innovation (R&I), and each year R&I's recommendations are proposed to the AASHTO Board of Directors and the National Academies. Research projects to address these topics are defined by NCHRP, and qualified research agencies are selected from submitted proposals. Administration and surveillance of research contracts are the responsibilities of the National Academies and TRB.

The needs for highway research are many, and NCHRP can make significant contributions to solving highway transportation problems of mutual concern to many responsible groups. The program, however, is intended to complement, rather than to substitute for or duplicate, other highway research programs.

NCHRP RESEARCH REPORT 920

Project 08-108
ISSN 2572-3766 (Print)
ISSN 2572-3774 (Online)
ISBN 978-0-309-48074-1
Library of Congress Control Number 2019948893

NOTICE

The research report was reviewed by the technical panel and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the National Academies of Sciences, Engineering, and Medicine.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; or the program sponsors.

The Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; and the sponsors of the National Cooperative Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

Published research reports of the

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

are available from

Transportation Research Board
Business Office
500 Fifth Street, NW
Washington, DC 20001

and can be ordered through the Internet by going to

http://www.national-academies.org

and then searching for TRB

Printed in the United States of America

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, non-governmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. John L. Anderson is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.national-academies.org**.

The **Transportation Research Board** is one of seven major programs of the National Academies of Sciences, Engineering, and Medicine. The mission of the Transportation Research Board is to increase the benefits that transportation contributes to society by providing leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied committees, task forces, and panels annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation.

Learn more about the Transportation Research Board at **www.TRB.org**.

COOPERATIVE RESEARCH PROGRAMS

CRP STAFF FOR NCHRP RESEARCH REPORT 920

Christopher J. Hedges, *Director, Cooperative Research Programs*
Lori L. Sundstrom, *Deputy Director, Cooperative Research Programs*
Dianne S. Schwager, *Senior Program Officer*
Jarrel McAfee, *Senior Program Assistant*
Eileen P. Delaney, *Director of Publications*
Natalie Barnes, *Associate Director of Publications*
Heidi Willis, *Editor*

NCHRP PROJECT 08-108 PANEL
Field of Transportation Planning—Area of Forecasting

Cindy L. Owings-Hutchison, *Maine DOT, Augusta, ME* (Chair)
Imad S. Aleithawe, *Waggoner Engineering, Inc., Jackson, MS*
James M. "Jim" Appleton, *Retired (formerly California DOT), Sacramento, CA*
Michael Bruff, *Capital Area MPO, Clayton, NC*
Laurie L. Goudy, *Maryland State Highway Administration, Baltimore, MD*
Jason Ray Junge, *Minnesota DOT, Roseville, MN*
Camille Kamga, *City College of New York, New York, NY*
Mario B. Rojas, IV, *A&P Consulting Transportation Engineers, Miami, FL*
Christopher J. Allen, *FHWA Liaison*
Matthew Hardy, *AASHTO Liaison*
Simon Randrianarivelo, *Bureau of Transportation Statistics Liaison*
Thomas Palmerlee, *TRB Liaison*

FOREWORD

By Dianne S. Schwager
Staff Officer
Transportation Research Board

Transportation Performance Management (TPM) is an established and maturing practice at state and local transportation agencies. *NCHRP Report 920* provides practical guidance to transportation agencies to improve their use of data for performance management. It will assist agencies in making visible progress in meeting their objectives—such as reducing fatalities, improving asset condition, reducing congestion, or speeding project delivery.

Recent federal legislation has established requirements for agencies to set performance targets and report on safety, pavement and bridge conditions, transit asset state of good repair, system performance, freight, and mobile source emissions. These requirements have resulted in increased visibility and attention to TPM and increased awareness of the importance of data within that process. Transportation agencies are recognizing that the value of performance management goes far beyond meeting federal requirements. *NCHRP Report 920* will assist agencies in making visible progress in meeting their objectives.

Many transportation agencies collect data but need to improve their capabilities to transform available data into useful information. This requires deliberate effort at all stages of the data life cycle, from specification through analysis, to make sure that data is of sufficient quality and that it can be integrated, visualized, and used to provide insights. Having people with the right skills and experience to carry out these activities is essential.

Under NCHRP Project 08-108, a research team led by Spy Pond Partners, LLC was asked to prepare guidance to improve data utilization in support of transportation performance management. The research team conducted a literature review and a series of interviews to identify current transportation agency practices for managing data-supporting TPM. Based on this practice review, they identified success factors and challenges related to efficient and effective data utilization within the TPM processes. They created guidance organized around six data life-cycle stages. The guidance includes a discussion of what is involved in implementing each step and some of the critical choices to be made; a synthesis of key points in the form of "Do's and Don'ts"; checklists that can be used to assess agency capabilities and identify opportunities for improvement; and illustrative examples.

While this guide draws many examples related to the federally defined TPM areas (safety, pavement, bridge and system performance), it does not provide official guidance for MAP-21/FAST Act target setting or reporting. It provides a framework for assessing current data management practices and a source of ideas for practice improvement. Its purpose is to promote practices that will enable agencies to go beyond meeting reporting requirements, to get valuable insights from data that can be used to boost agency results.

The *Guide for Practitioners* is accompanied by a downloadable report, *Developing National Performance Management Data Strategies to Address Data Gaps, Standards, and Quality: Final Research Report,* available on the TRB website (www.trb.org) by searching for "NCHRP Research Report 920."

Table of Contents

Introduction

Why This Guide?

Transportation Performance Management (TPM) is an established and maturing practice at state and local transportation agencies. Recent federal legislation has established requirements for agencies to set performance targets and report on safety, pavement and bridge conditions, transit asset state of good repair, system performance, freight, and air quality. These requirements have resulted in increased visibility and attention to TPM and increased awareness of the importance of data within that process.

Transportation agencies are recognizing that the value of performance management goes far beyond simply meeting federal requirements—it provides a way to make visible progress in meeting agency objectives such as reducing fatalities, improving asset condition, eliminating bottlenecks, or speeding project delivery.

Meanwhile, with the emergence of new collection technologies, data sources, and private-sector data product and service offerings, seismic shifts are occurring in the data landscape. These changes are causing transportation agencies to re-examine their approach to data management.

TPM Defined

This guide uses the term "TPM" to mean a process that involves

1. Measuring current transportation system performance,
2. Setting goals and targets for performance improvement,
3. Allocating resources and planning work to achieve the desired improvement, and
4. Monitoring results achieved to adjust plans and programs and update targets as needed.

Performance Measurement:

Understanding current performance and tracking progress toward meeting objectives.

Performance Management:

Using performance measurements and other information to drive decisions.

Data are essential throughout the TPM process. Good quality data allow agencies to understand current performance, assess risks, and plan and prioritize improvement actions. Reliable data can be used to pinpoint trouble spots, identify where to intervene, and illuminate where investment is most beneficial. It helps agencies to maximize the effectiveness of available funding and staff resources, and it provides the strong, credible foundation for decision making that is essential for building and maintaining public trust.

For many agencies, the problem is not a lack of data; it is a lack of capability to transform available data into useful information. This requires deliberate effort at all stages of the data life cycle, from specification through analysis, to ensure that data are of sufficient quality and that it can be integrated, visualized, and used to provide insights. Having people with the right skills and experience to carry out these activities is essential.

Careful attention to assessing data needs and implementation of proven data management practices can provide the foundation for a successful TPM process—helping agencies to make wise transportation investments that save taxpayer dollars, reduce congestion, encourage economic growth, improve the quality of life, and even save lives. Lack of skillful data management creates inefficiencies due to duplicative effort and manual, error-prone processes. It also creates risks ranging from failed data collection efforts to lack of progress on improving performance.

Most transportation agencies are not starting from scratch with TPM; they will have already established a set of transportation performance measures and data sources used for internal and external reporting. However, there are many scenarios that may cause an agency to seek ideas for improving performance measures and associated data programs:

- Current data are not sufficiently reliable in calculating performance measures that support funding and resourcing decisions.
- Current performance measures are not helpful in identifying how and where to make improvements.
- Agency performance measures were established based on what data were available at the time, and now there are better data sources available.
- New analysis or visualization tools are available that would provide value, but require new data or changes to data formats.

Target Audience

This guide is intended for transportation agencies—state departments of transportation (DOTs), Metropolitan Planning Organizations (MPOs), or local agencies—seeking to improve the efficiency and effectiveness of how they collect and use data for TPM. It can be used to review and assess existing data-related activities.

The target audience includes

- Staff with responsibility for TPM and/or individual performance areas (e.g., safety, pavement, mobility) who are seeking ideas for making effective use of data;
- TPM staff seeking to gain a broader understanding of data sources and management methods;
- Managers seeking to communicate the importance of data sharing and coordination to improving performance and strengthening accountability;
- Managers seeking materials to help change their agency culture so that people see data as an enterprise asset; and
- Data analysts seeking to understand the entire data life cycle and how what they do fits into a bigger picture.

- New data management practices are being implemented agency-wide that offer opportunities to re-think how performance data are collected and processed.
- Changes to organization or staffing of TPM functions have taken place, and the new manager is looking to make improvements.
- New or modified federal or state requirements for performance reporting are issued.

While this guide draws many examples related to the federally defined TPM areas (safety, pavement, bridge, and system performance), it does *not* provide official guidance for MAP-21/FAST Act target setting or reporting. It *is* intended to provide a framework for assessing current data management practices and a source of ideas for practice improvement. Its purpose is to promote practices that will enable agencies to go beyond meeting reporting requirements and to get valuable insights from data that can be used to boost agency results.

Research Report Available

This guide is a product of NCHRP Project 08-108, "Developing National Performance Management Data Strategies to Address Data Gaps, Standards, and Quality." A final research report for this project is available that includes a detailed literature review and practitioner interview results that were used to shape this guide.

Relationship to Other TPM Guidance

This guide provides an overview of data management practices for TPM. It is not a comprehensive reference on working with TPM data and does not go into depth on individual data sets or technical management processes. It is intended for use in conjunction with other available references.

To understand how the material in this guide fits with the bigger-picture TPM processes, the FHWA TPM Toolbox provides a useful framework, which is illustrated in Figure 1. This guide builds on two components of the FHWA TPM framework: C-Data Management and D-Data Usability and Analysis. It also acknowledges the essential role of organizational and cultural factors in successful TPM, illustrated by the outer circle (component A) that encompasses all other elements of TPM.

Institutional barriers are often much harder to overcome than technical barriers in making improvements to data management and utilization. It is important to have strong leadership and buy-in across the organization for any improvement initiative. The FHWA TPM Toolbox includes a maturity assessment that agencies can use to gauge the state of their organization and chart a future course.

See the end of this chapter for a link to the TPM Toolbox and additional reference material on TPM and transportation data assessment and management.

Scope of the Guide

This guide provides a high-level view of data management practices for TPM. Readers are encouraged to consult the references provided at the end of each section for further information on topics of interest.

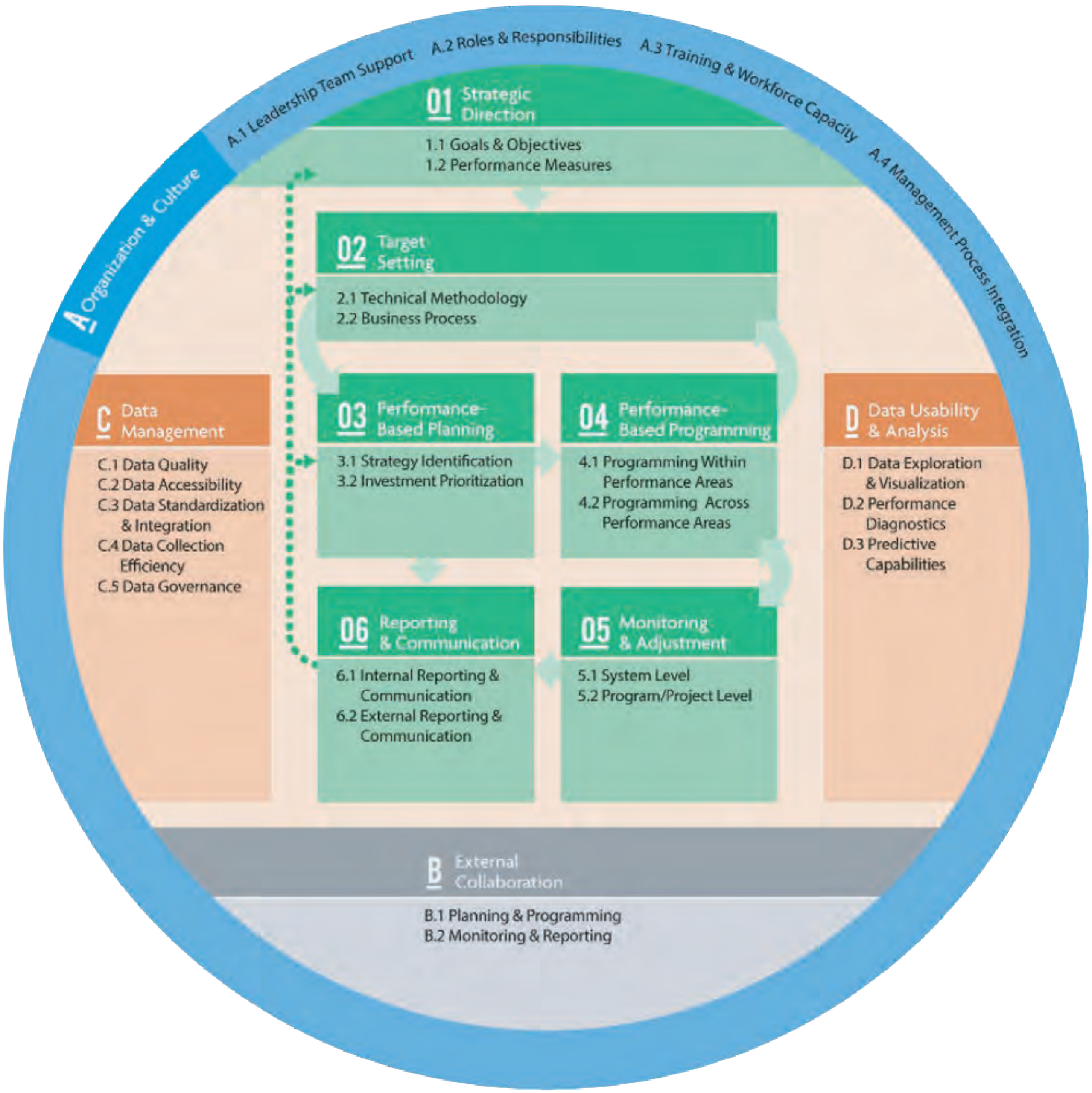


Figure 1. The FHWA TPM framework.

Framework for Improving Data Utilization in TPM

The framework for this guide is based on a cyclical model of a generic data management process, depicted in Figure 2.

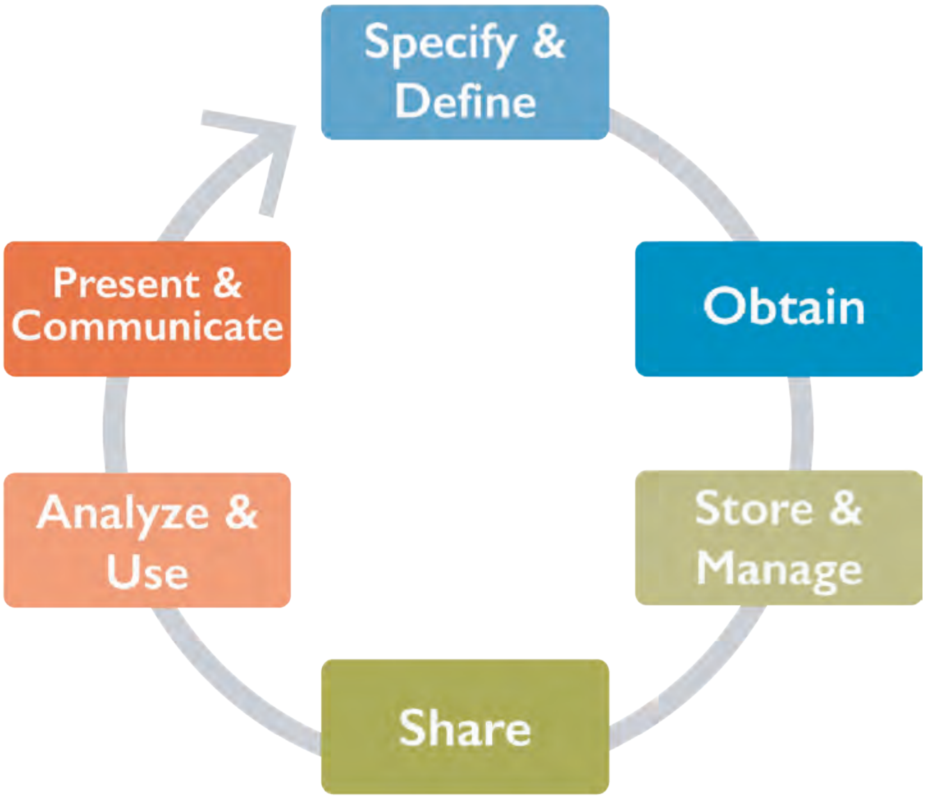


Figure 2. The data management cycle.

This cycle begins with specifying and defining data requirements, and then proceeds to obtaining the data, storing it in one or more repositories, and processing it as needed to support use. Then, data are shared in various forms, analyzed and used for decision making, and communicated to different audiences.

Figure 3 shows the organizing framework for this guide. It extends this cyclical model of data management to illustrate the process of utilizing data for TPM.

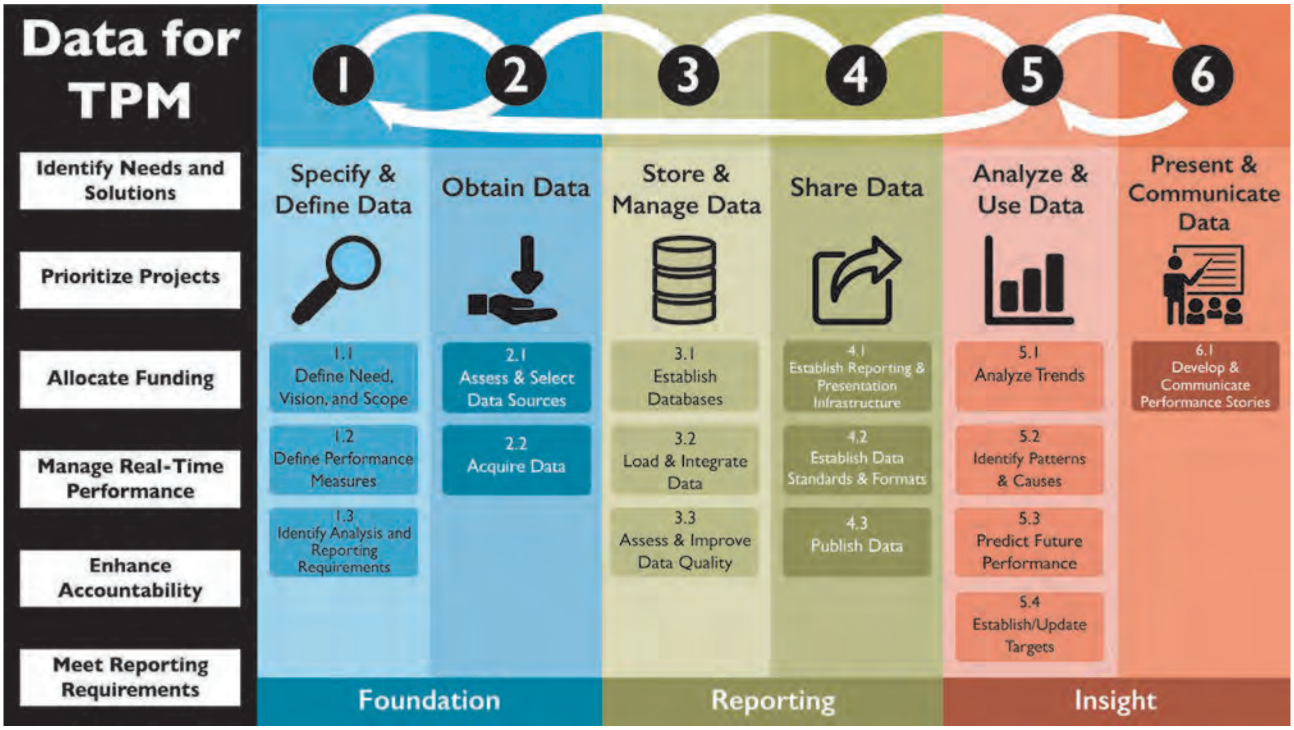


Figure 3. Framework for improving data utilization for TPM.

The left side of Figure 3 shows how data can be used within TPM. This provides the motivation and the requirements for the six data management processes shown on the right side of the diagram.

Every investment in data—whether it is collecting new data or improving reporting tools—should be evaluated based on how it helps the agency to better identify needs and solutions, prioritize projects, allocate funding, manage real-time performance, enhance accountability, and/or meet reporting requirements.

The bottom row of Figure 3 identifies three fundamental processes for making use of data for TPM: Foundation, Reporting, and Insight:

- **Establish a data *Foundation***—defining performance measures, identifying data requirements, selecting data sources, and obtaining the data.

- **Set up *Reporting* processes**—establishing databases and reporting systems and publishing data and reports.
- **Gain *Insights* for decision making**—analyzing performance data to understand trends and explanatory factors/root causes, developing predictive capabilities, setting performance targets, and communicating the agency's "performance story" to different audiences.

The numbered columns of Figure 3 show how each of these processes are broken down into more specific steps to specify, obtain, store and manage, share, analyze and use, and communicate data for TPM.

Establish a Data Foundation

Step 1: Specify & Define Data. Determine what types of data are needed, how data will be used within TPM business processes, and—based on this—specify attributes, scope, level of spatial and temporal granularity, and frequency of updates.

Step 2: Obtain Data. Acquire the data needed for calculating performance measures and for understanding trends and root causes of performance results.

Set Up Reporting Processes

Step 3: Store & Manage Data. Set up data repositories either within the agency or "in the cloud"; load, validate, clean, and integrate the data; document the data for both technical and business users; and manage access to the data—to both protect it from unauthorized use and to ensure that it is accessible to those who need it.

Step 4: Share Data. Put the infrastructure in place to produce data products (e.g., reports, maps, interactive portals) and share data across business units within the agency, with partner agencies, and/or with the general public.

Gain Insights for Decision Making

Step 5: Analyze & Use Data. Configure and use various data analysis tools to understand trends, predict future performance, and formulate performance targets.

Step 6: Present & Communicate Data. Translate data and analysis results into information that can be understood and used by different data customers.

It is important to note that getting data right for TPM can take time and experimentation. While these steps are numbered and generally follow a linear order, they may be carried out in parallel, in a different sequence, or in an iterative fashion. For example, in the process of creating data visualizations (Step 6), one might identify needs for additional data (Step 2) or improved data cleansing (Step 3), or perhaps creation of an additional performance measure (Step 1).

Common feedback loops are illustrated in Figure 3 with arrows connecting different steps along the top:

- Steps 5 and 6 are often accomplished iteratively—analysis of data is conducted in preparation for presentation, but the process of presenting and communicating data may lead to further analysis to answer questions that arise.
- Data analysis in Step 5 may identify needs for changes in how data and performance measures are specified.
- Review of data after it is acquired may also uncover gaps or limitations that were not anticipated, and they may lead to the need for changes in data or performance measure specification.

Nevertheless, this simplified framework can be used to consider all of the steps needed to make effective use of data for TPM and identify areas for improvement.

The remainder of this guide provides guidance for each of the six steps and presents a set of case study vignettes that illustrate various aspects of data utilization for TPM. Note that this guide does not attempt to be comprehensive in its treatment of data management or TPM practices given the extensive available literature on these topics. It focuses on key points that agencies should keep in mind when considering practice improvements.

Guidance for each step includes

- A discussion of what is involved in implementing each step and some of the critical choices to be made
- A synthesis of key points in the form of “Do’s and Don’ts”
- Checklists that can be used to assess agency capabilities and identify opportunities for improvement

  - o For each step, checklists for "basic" and "advancing" capabilities are provided. The intent is that agencies focus first on getting the "basic" capabilities in place, and then seek to further advance their practices. A set of pull-out worksheets are provided at the end of this guide with these checklists that can be duplicated and used to assess agency capabilities.
- References that provide further information related to the step

For more information...

1. FHWA TPM Toolbox
   Guidebook: https://www.tpmtools.org/guidebook/
   Maturity Assessment: https://assessment.tpmtools.org/
2. NCHRP Report 866: Return on Investment in Asset Management Systems and Practices (TRB, 2018) https://www.nap.edu/catalog/25017/return-on-investment-in-transportation-asset-management-systems-and-practices
3. Applying Archived Operations Data in Transportation Planning: A Primer (FHWA, 2016) https://ops.fhwa.dot.gov/publications/fhwahop16082/
4. NCHRP Report 814: Data to Support Transportation Agency Business Needs: A Self-Assessment Guide (TRB, 2015) http://www.trb.org/Main/Blurbs/173470.aspx
5. Performance-Based Planning and Programming Guidebook (FHWA, 2013) https://www.fhwa.dot.gov/planning/performance_based_planning/pbpp_guidebook/
6. Benefit-Cost Analysis of Investing in Data Systems and Processes for Data-Driven Safety Programs (FHWA, 2012) http://safety.fhwa.dot.gov/rsdp/downloads/bcareport.pdf

Foundation

Step 1

Specify & Define Data

This step involves the up-front work to define data requirements for TPM.

You may be seeking to add new performance measures, modify your existing measures, or improve the efficiency of your data gathering processes. Whenever a change to performance measures or associated data is contemplated, the following guidance can help to plan and scope improvements.

Taking the time to work through each sub-step below will pay off in the form of a solid business case for change and reduced risk that re-work will be required in the future.

"We don't have time to do it right, but we always have time to do it over."
– Anon

Step 1.1

Define Need, Vision, & Scope

Define business needs for data. Begin with the business objective(s) and concern(s) in mind and consider how your performance measures will be used to support them. The goal is to clearly articulate the business case for new data to agency managers and stakeholders. To do this, you need to answer three questions:

- What will new data tell us?
- How will we act on it?
- Is the cost of obtaining the data worth the value that will be added?

Define and document what information is needed to meet both internal agency decision-making requirements as well as external reporting and information-sharing requirements.

- For *agency decision support*, document each decision: what is the decision, when is it made, who makes it, and who provides supporting information. Example decisions are which bridges to program for rehabilitation, which intersections to target for safety improvements, what percentage of available funding should be allocated to bridges versus pavements, and what strategies should be considered to address freight bottlenecks.
- For *external performance reporting requirements*, document what needs to be reported, when reports are due, and the required format for the information. Include references to any applicable regulations or guidance documents.
- For external performance information sharing, define what performance information the agency will share with the traveling public and with external partners. Document the intended uses of the information by each type of audience.

Specify the data requirements. For each of the above business needs, identify the following:

- What data attributes are essential for calculating performance measures, and what additional attributes might be helpful for providing context and interpreting performance results?
- What scope of data coverage is needed?

Case G

To support its TPM efforts and as part of a broader strategy for making effective use of data, the Mid-America Regional Council (MARC) data coordination committee compiled a top 10 list of high-priority data sets for automation. This top 10 created a road map for subsequent work activities to organize critical data sets at MARC. Priorities included pavement and bridge conditions, safety measures, and system performance.

Tip

Don't limit your scope to the data needed to calculate performance measures. Also consider the data needed to understand trends or patterns, formulate strategies, and identify appropriate actions to improve performance.

- What granularity of data is needed—what time period and spatial unit should each data observation represent?
- What level of precision is required?
- How current do data need to be—one year old? one day old? real time?

Current federal TPM rules establish many of these requirements for pavement, bridge, safety, and system performance measures. However, many agencies choose to go beyond these minimum requirements or maintain additional measures to meet their own internal needs. For example, state DOTs generally track pavement condition for all paved roads they maintain, not just on the National Highway System, and use locally defined condition indices that support project-level decision making.

Identify common data needs across business units. Rather than looking at the data requirements of individual business units in isolation, identify common needs. Socio-economic data, population forecasts, traffic data, roadway characteristics, and project status are examples of shared data needs. Engage each business unit in vetting of possible data sources and work toward a coordinated data gathering strategy.

Case J

PennDOT identified five specific business goals that guided collaborative establishment of a Statewide Operations Data Warehouse that broke down agency data silos.

Carefully specify location attributes. Involve data users and geographic information system (GIS) staff in your agency in specifying location accuracy requirements and measurement methods. Be sure to understand the reasons for gathering data based on a linear reference rather than using Global Positioning System (GPS) locations. Linear references such as route-milepoint can provide convenient ways for someone without technology to find a feature. If only GPS coordinates are available, it can sometimes be difficult to determine a specific route location—for example, for complex interchanges.

Case E

Maryland SHA found that linking disparate data sets enabled effective after action reviews to improve operational preparedness and response.

Consider data integration needs. Effective TPM rarely relies on a single isolated data set. Meaningful metrics often require integration of multiple data sets and types. For example, speed data alone may identify slowdowns and delay patterns, but it does little to provide insight as to *why* there are slowdowns. Combining speed data with incident and weather data can paint a more complete picture of system performance as it relates to recurring and non-recurring congestion.

In order to integrate data sets effectively, agencies must focus on linkages between data sets using time and geography information, as well as available contextual information. For example, the location and time of the start of delay should be cross-referenced with available weather conditions in that location and at that time, as well as understanding that the delay is happening during a busy holiday season in a retail area. Considering data integration needs early can save effort later that may be required to build integrated data sets for analysis.

Key questions to address are the following:

- What data sets need to be linked?
- What data elements can be used for temporally and spatially linking the data sets?
- What quality checks are required to make sure that these link elements are complete and accurate?
- What data standards should be followed for newly acquired data so that these linkages can be made?

Answering these questions will ensure that any new data collected are "integration-ready." It is much easier to address integration needs up front rather than after data are collected.

Tip

Anticipate that an iterative approach will be needed to fully understand your data requirements. There is no substitute for having actual data in hand and trying to use it. That will inevitably lead to adjustments in specifications.

For more information...

1. Guide for Prioritizing Assets for Inclusion in Transportation Asset Management (TAM) Programs (FHWA, 2019—forthcoming)
2. Priorities in Roadway Safety Data Guide (FHWA, 2017) https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa17032.pdf
3. Minnesota Department of Transportation Data Business Plan (Minnesota Department of Transportation, 2007) https://www.dot.state.mn.us/tda/databusinessplan.docx

Step 1.2

Define Performance Measures

Specify, test, and document performance measure calculations. For each performance measure, precisely document the data inputs and calculations needed. This documentation should have all of the information needed for a programmer/analyst to implement the calculations. Test the calculations with sample data and compare values and trends against other similar measures that may be available.

Describe performance measures in plain English. Performance measures involving multiple data inputs and complex calculation logic should be documented in a manner that end users can understand. For example, the measure "Buffer Time Index" can be described as "the amount of extra buffer time a commuter needs to allow to avoid being late to work more than one day per month."

For more information...

1. National Performance Measures for Congestion, Reliability, and Freight, and CMAQ Traffic Congestion (FHWA, 2018) https://www.fhwa.dot.gov/tpm/guidance/hif18040.pdf
2. Validation of Pavement Performance Measures Using LTPP Data: Final Report (FHWA, 2018) https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpp/17089/17089.pdf
3. Computation Procedure for the Bridge Performance Measures (FHWA, 2018) https://www.fhwa.dot.gov/tpm/guidance/hif18023.pdf
4. Freight Performance Measures Primer (FHWA, 2017) https://ops.fhwa.dot.gov/publications/fhwahop16089/fhwahop16089.pdf
5. Transportation Management Center—Data Capture for Performance and Mobility Measures Reference Manual (FHWA, 2013) https://rosap.ntl.bts.gov/view/dot/3373

Case I

Ohio DOT established "regain time" as a winter performance measure and defined it as the elapsed time from the end of the snow or ice event to the time at which speeds recover to typical levels. Regain time is the type of measure that is easily communicated to decision makers and the general public, yet it ties well with operational actions that directly influence it.

Step 1.3
Identify Analysis & Reporting Requirements

Understand data user needs. Consider the information needs of different audiences and identify how they want to consume this information. Conduct interviews or focus groups to learn about analysis and reporting needs and desired improvements.

Specify and document data extracts and report formats. For some audiences, standard, static reports will be sufficient; others may need more flexible views of the data with the ability to drill down into details from a summary view or to obtain direct access to data via an application programming interface (API). Others may want to load detailed performance data into specialized analysis tools such as pavement management systems, safety analysis systems, or traffic simulation models. Identifying these different needs early can help to avoid unexpected data requests that may be difficult to satisfy once data systems and processes are established.

For more information...

1. Road Safety Fundamentals—Unit 3: Measuring Safety (FHWA, 2017)
   https://rspcb.safety.fhwa.dot.gov/RSF/unit3.pdf
2. TRB Web Document 9: Meeting Critical Data Needs for Decision Making in State and Metropolitan Transportation Agencies—Summary of a Conference. Transportation Research Board Conference Proceedings on the Web (Transportation Research Board of the National Academies, 2013)
   http://onlinepubs.trb.org/onlinepubs/conf/CPW9.pdf

Case G

To ensure MARC's data developers understand how the data they manage is ultimately used, data personnel regularly participate in meetings with transportation planning staff, both to discuss high-level data needs and more focused detail-oriented breakout information.

Capabilities Checklist: Specify & Define Data

Basic

- ☐ The business need for data has been identified and documentation of this need is available for future reference.
- ☐ An inventory of existing agency data sources has been compiled.
- ☐ Managers of the units responsible for data collection can describe the primary users and uses of that data.
- ☐ Data requirements to meet internal and external performance reporting requirements are defined and documented, including attributes, scope, and granularity.
- ☐ Location referencing methods for performance data are established to enable linkages with other agency data sets.
- ☐ Update frequencies for new data are defined and documented.
- ☐ Authoritative data sources have been designated for performance measure calculations.

Advancing

- ☐ Discussions about data requirements are not constrained by the status quo; they reflect what is important to know about transportation performance in order to improve.
- ☐ Data needs are identified to support the entire TPM cycle (beyond performance reporting) including root cause analysis, identification and prioritization of improvements, and evaluation of impacts.
- ☐ Minimum data quality standards are established considering timeliness, accuracy, completeness, consistency, and accessibility.
- ☐ Data requirements are defined collaboratively across business units, including GIS and information technology.
- ☐ Data communities of interest (or equivalent) have been established to identify data improvements to support different business needs.

Do's & Don'ts

Do:

- ✔ Have a clear business case for new data that articulates how the data will be used to improve performance.
- ✔ Educate stakeholders on the benefits of performance measures.
- ✔ Put in the time and effort needed to nail down data requirements in a precise fashion—ambiguity can lead to downstream problems.
- ✔ Implement small-scale pilots to test out new types of data and assess their value.

Don't:

- ✕ Establish performance measures solely based on what data you have (instead, figure out what you would like to measure first and then assess options for getting the data you need).
- ✕ Approach data requirements from the perspective of a single business unit (instead, try to identify common needs).
- ✕ Go too lean on data requirements in order to save money in the short term and miss an opportunity to gain new insights that provide a bigger bang for the buck.
- ✕ Neglect to understand location accuracy requirements (instead, investigate how location should be measured and represented).

Step 2
Obtain Data

This involves activities related to obtaining the data needed to support the entire TPM process, including

- data needed to calculate performance measures,
- data needed to provide context necessary to understand performance trends,
- data needed to understand root causes and factors contributing to performance results,
- data needed to set realistic targets, and
- data needed for selecting strategies to improve performance.

These data may be obtained from existing internal, external, and commercial sources. New data may also be gathered using in-house resources and/or via contract.

"Increasingly, data is gathered by information-sensing mobile devices, remote sensing, software logs, cameras, microphones, and wireless sensor networks. Global technological information per-capita capacity has approximately doubled every 40 months since the 1980s."

Institute of Engineering and Technology

Step 2.1

Assess & Select Data Sources

Identify available data. Identify existing data sources that could be tapped to meet some or all of the requirements. Review sources within the agency, sources from external partners (federal agencies, state agencies, Metropolitan Planning Organizations (MPO), local agencies, universities), and commercial sources. Obtain detailed information about each source, including data elements and their definitions, scope, date of last update, frequency of updates, available formats, costs, and use restrictions.

Evaluate available data sources against requirements. For each available data source, identify gaps between what is required and what the source can deliver. Develop estimates of the costs to fill the gaps through new data gathering activities. Use the list of gaps and the cost information to discuss

- whether some of the requirements can be relaxed,
- whether proposed performance measures can be modified to better align with available data, and
- whether there are workarounds that can be applied (e.g., through sampling techniques or fusion of multiple data sources).

Also, identify any constraints associated with data sources (e.g., usage restrictions) that could limit the value of those sources to support TPM.

Case I

Ohio DOT wanted to track how quickly roads returned to normal speeds following a storm. They were able to leverage existing data sources, including their RWIS and commercial speed data.

For more information...

1. Institution of Engineering and Technology: IET Sector Insights: Big Data in Transport https://www.theiet.org/sectors/transport/topics/intelligent-mobility/articles/big-data.cfm
2. Using Truck GPS Data for Freight Performance Analysis in the Twin Cities Metro Area (Minnesota Department of Transportation, 2014) http://www.dot.state.mn.us/research/TS/2014/201414.pdf
3. NCFRP Report 25: Freight Data Sharing Guidebook (Transportation Research Board of the National Academies, 2013) https://www.nap.edu/catalog/22569/freight-data-sharing-guidebook
4. Asset Management Data Collection for Supporting Decision Processes (FHWA, 2006) http://www.fhwa.dot.gov/asset/dataintegration/if08018/assetmgmt_web.pdf

Step 2.2

Acquire Data

It may be appropriate to launch a data acquisition effort if

- current agency data sources will not meet the requirements,
- there are no suitable commercial sources that meet the requirements (for an acceptable price),
- there is a business case for new data collection, and
- resources are available—both for initial collection and ongoing upkeep of the data.

Once you have decided to collect new data, you must determine whether to collect the data with in-house personnel or outsource data collection to a vendor. This will depend on the scale of the effort and in-house staff capacity.

Regardless of who will be collecting the data, it is essential to have a documented plan describing how it will be collected.

Create a data collection and quality management plan. Prepare a detailed plan to guide both data collection and quality management activities (see Table 1 for suggested elements of such a plan). Data quality management takes additional time and effort but should be integral to the data collection process. Without sufficient resourcing for data quality, there is a risk that the data collected will not be usable.

Remember that the people who collect the data are in the best position to ensure quality. Build in training activities so that they understand not only how to collect the data, but why the data are being collected and what the intended uses are. Check in with them during the data collection and see if they have suggestions for improving the process.

Case E

Maryland SHA captures a rich set of data about highway incidents including the name of responders, the road surface conditions, lane closings/openings over the course of the incident, and operator notes. These data are then combined with data from ITS devices [Dynamic Message Signs (DMS), Closed Caption Television (CCTV) images, volume and speed detectors, signals] and probe-based speed data.

Table 1. Contents of a data collection and quality management plan.

Data Specifications	Data element descriptions and allowable values Measurement methods Accuracy and precision Data formats
Data Collection Procedures	Schedule Scope Roles and responsibilities Procedures Issue reporting and resolution protocols
Staffing	Training and certification of data collectors Supervision and management roles
Equipment	Specifications Calibration Certification
Quality Control	Quality checks before and during collection
Data Acceptance	Acceptance criteria Error resolution procedures
Data Review and Validation	Review procedures and responsibilities Accuracy checks for sample records Independent verification Aggregation—check totals Field-level validation Record-level validation Check against prior value Visualization
Reporting	Data quality metrics and targets Data quality reporting protocols Error reporting procedures for data users

For more information...

1. Guidelines for Development and Approval of State Data Quality Management Programs (FHWA, 2018) https://www.fhwa.dot.gov/pavement/management/pubs/dqmp.pdf
2. Southeast Michigan Council of Governments—Innovative Traffic Data Quality Assurance/Quality Control Procedures and Automating AADT Estimation (Case Study) (FHWA, 2015) https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa17035.pdf
3. Practical Guide for Quality Management of Pavement Condition Data Collection (FHWA, 2013) https://www.fhwa.dot.gov/pavement/management/qm/data_qm_guide.pdf
4. I-95 Corridor Coalition, Vehicle Probe Project Scope and Methodology (2009) http://i95coalition.org/wp-content/uploads/2015/02/Validation-Process-May-19-2009-distr-June-20092.pdf?x70560

Capabilities Checklist: Obtain Data

Basic

- ☐ Data collection procedures and protocols are defined and documented.
- ☐ Data collection and processing workflows are mapped to clearly assigned responsibilities and deadlines.
- ☐ Existing agency data sources are reviewed prior to collection of new data.
- ☐ Available external (public and private) data sources are reviewed prior to collection of new data.
- ☐ Quality management procedures are defined and documented—including training and certification for data collection personnel.
- ☐ Requirements are in place that ensure new data collection adheres to agency location referencing standards.
- ☐ Impacts of changes to existing data collection methods are assessed to minimize loss of consistent trend data and disruption to existing reports.
- ☐ Data sources are assessed to understand usage restrictions that may limit value.

Advancing

- ☐ The full cost of new data acquisition is estimated, considering initial collection, ongoing updates, and supporting staff and technology infrastructure.
- ☐ Funding for regular data updates (beyond the initial collection) is planned and committed.
- ☐ There is regular communication with partner agencies to identify opportunities for collaboration on data collection.
- ☐ Periodic scans are conducted to identify ways to improve data quality and collection efficiency.
- ☐ Agency guidance and/or coordination protocols have been established to assist business units wishing to purchase commercial data sources.
- ☐ Specialists with appropriate expertise (in-house or contractors) evaluate use of emerging private data sources.
- ☐ Data requirements are defined with consideration of opportunities to create valuable information through integration of multiple data sources.

Do's & Don'ts

Do:

- ✔ Coordinate across units within the agency prior to collecting new data to avoid duplication; plan for integration needs.
- ✔ Invest in data quality management to make sure that data collected are reliable.
- ✔ Communicate with partner agencies to identify areas for collaboration.
- ✔ Periodically evaluate if there is a better way to get the data you need.

Don't:

- ✕ Purchase private data without understanding its derivation and limitations.
- ✕ Enter into a data use agreement with terms that are overly restrictive for your agency.
- ✕ Change data collection methods without identifying how this may impact existing reports and the ability to understand trends.

Reporting

Step 3

Store & Manage Data

This step includes validating, cleaning, normalizing, aggregating, and integrating data; storing the data in one or more repositories—either within the agency or "in the cloud"; producing documentation needed for both technical and business users of the data; and managing access to the data—to both protect it from unauthorized use and to ensure that it is accessible to those who need it. This step also includes activities to design, develop, and manage databases and technical infrastructure for data storage and data integration.

"Data is just like crude. It's valuable, but if unrefined it cannot really be used."

Michael Palmer

The key decisions that agencies must make are

- where and how to store data,
- how to make sure data can be integrated across repositories as needed,
- which best practices should be implemented for QA and documentation, and
- how much data to keep.

Step 3.1

Establish Databases

Design databases to support analysis needs. Performance measures rely on a deep archive of data to develop an accurate baseline; understand multi-year, seasonal trends; and establish reasonable targets. Database design supporting performance measures should consider requirements for reporting, trend analysis, and root cause analysis. Design should also consider the possibility that requirements may change over time—for example, an agency may decide to calculate different metrics, drawing on the same raw data sources. Therefore, both raw and transformed data may need to be stored. When raw data is voluminous (for example, pavement images), processed data can be maintained in active storage and the raw data can be kept in lower-cost archive storage.

Determine data retention policies. If retention policies are not modernized to reflect changes in storage costs, or if they are set without full understanding of business needs, there is a danger of loss of valuable data and TPM capability. Ten or more years ago, data storage hardware was both physically large and expensive. Therefore, agencies implemented data retention policies to better manage budgets and constrained physical space in data centers by limiting the amount of storage and the duration of the storage. Both the size and cost of storage have dropped dramatically over the years. With the exponential cost savings and available storage options, agencies can re-examine their retention policies to make sure they align with business needs. For example:

- An agency may be required to report performance of the system to the federal or state government in 15-minute intervals. That agency may be tempted to aggregate raw data coming in 1-minute intervals and only retain the aggregate information to save space. Later, the agency may identify a need to track incident management performance metrics—requiring the original 1-minute data that tracks growing and shrinking queue lengths, user delay, arterial signal performance, and the effects of secondary incidents. If the 1-minute data are gone, the agency may be unable to track that metric accurately (if at all).
- A data set may have limited value by itself and would be considered unimportant for retention. However, when combined with other

Case G

MARC repurposed two open positions, including a "GIS specialist" and a "demographer," into "data developer" positions capable of creating and managing systematic workflows for data gathering and organization. The developers have since created automated processes to obtain data sets and import them into SQL databases. The databases have front-end interfaces that greatly simplify the process of querying them to extract the information MARC needs.

data sets, it may provide new and important performance measures insights that neither data set could in isolation.

Data storage architectures can be designed to consider both current and future TPM data requirements. For example, cleaned and processed data that support frequent and/or immediate TPM needs can be stored using faster and more accessible servers. Raw data used to generate those measures and calculations can be stored on lower-end servers or storage arrays locally or with a cloud-based back-up provider. Agency TPM personnel should work collaboratively with information technology (IT) departments and records managers within the agency to draft more modern retention policies.

Plan for both "big" and "small" data. As the concept of "big data" becomes more prevalent and hyped, there is a potential for data proponents and decision makers to focus too much on big data and overlook the value of "small data." For example, in an effort to utilize a big data set, agencies may invest in big data platforms that are not designed to handle smaller data sets. The push to keep on top of the latest technology trends can usurp resources from existing data and performance management programs and may not meet the agency's predominant requirements.

How does an agency determine when data is big enough to require a big data storage approach? The true sign of a need for a big data platform is when traditional storage and processing techniques become inadequate for the specific usage needs that have been identified. Agencies should strike a balance in their investment in big and small data platforms and work backwards from their use cases to the storage strategies. For example:

- It is appropriate to store some data sets in relational databases if they lend themselves well to normalization, indexing, joining, and database-supported statistical analysis. Calculating incident clearance metrics can be done very effectively using a traditional database management system.
- On the other hand, complex analytics that parse through billions or even trillions of data points for identifying problem locations, prioritizing projects, computing user delay, or understanding the complexities of signal retiming efforts are computationally expensive and are not cost effective in traditional relational databases. In order to fully leverage both types of data sets,

agencies can develop simple interfaces that operate across both the traditional relational databases and big data platforms.

Agencies should have strategies for both big and small data, including a way to integrate across both types of data. A well-designed storage architecture can be flexible enough to accommodate both data sets as well as incremental changes as technology and data continue to develop.

Identify storage options. In the past, agencies had to rely on in-house systems to store and process their data. As the concept of cloud computing becomes more mainstream, agencies are now presented with a choice between storing data in-house or in the cloud. There are also a variety of commercial and open source data storage options available for both cloud and on-premise. What is right for each agency is highly dependent on the TPM use case, agency policies (procurement, IT, etc.), IT staff support, the type of data, and how frequently the data will be accessed. There are several general considerations each agency should address as they evaluate their options.

Consider commercial cloud and specialized hosted solutions. Commercial cloud providers' services (like Amazon, Azure, etc.) appear to be very affordable when pricing is based solely on the amount of data to be stored. However, agencies need to consider more than just the size of their data when estimating costs. Pricing is also highly dependent on the number of transactions (or the number of times you access the data and process it) and how much bandwidth is used.

For example, it may be inexpensive to load raw data into the cloud, but very costly to extract it back when needed to calculate metrics to support TPM. The appropriate approach then may be to move data processing and metrics calculations to the cloud as well and avoid extraneous costs of downloading data each time it needs to be processed.

Alternatively, agencies can rely on specialized hosted solutions from their partners: universities, consultants, and sister agencies. Partners may have a better understanding of the transportation and TPM domain and provide more cost-effective approaches for storing transportation data.

While in-house solutions often provide more control, they also require a larger investment in workforce and physical infrastructure. However, storing data in the cloud may not be cost effective for highly transactive systems. It is important to consider all these aspects to avoid becoming a "hostage" of a vendor or service or investing in an internal system that becomes obsolete due to its inability to innovate and stay current.

Case H

New Jersey DOT uses an outsourced cloud-based hosting platform and subject matter expertise to collect, store, and analyze complex data sets to support agency TPM efforts.

Plan for data security. Implement a sound data back-up strategy that will allow you to restore data in the event of a hardware failure, cyber-attack, or inability to physically access facilities. If your data contains personally identifiable information (PII) or other sensitive elements, it should be clearly categorized as sensitive and managed to prevent unauthorized access. TPM-related data that may be sensitive include crash reports, travel survey data, and data from mobile devices. Some agencies have policies that do not allow sensitive data to be stored in the cloud. However, many cloud providers have a robust security policy to both prevent and recover from cybersecurity compromises. In contrast, agencies may have limited funds and expertise to implement robust security mechanisms.

Maintain metadata. While some data sets are considered "self-explanatory," metadata and documentation are critical. For example, highway crashes may appear to be a straightforward data set. On closer examination, you may find that data from one jurisdiction is gathered using different definitions for serious injuries than another. Data may be collected using a mixture of electronic and manual processes with different quality assurance processes applied. Newer data sets may be provisional and subject to further updates. Metadata and documentation become even more important when data is used in calculations to support TPM. Two individuals can use the same raw data and measure definition, but execute calculations differently depending on the context and interpret the results completely differently.

Metadata should be maintained at both the data set and data element level. Data set metadata covers information such as source, spatial and temporal scope, quality, and access classification. Data element metadata covers meaning, origins, usage, value domain, and format. Standards for data set level metadata can be found in International Organization for Standardization (ISO) 19115 and the Office of Management and Budget's (OMB) Project Open Data (POD) Schema. Standards for data element level metadata can be found in ISO/IEC 11179.

Proper metadata and documentation that is frequently updated and audited can ensure that confusion and interpretation variations are minimized. Metadata and documentation must be properly versioned so that data processing spanning different versions of metadata can be interpreted and processed properly.

For more information...

1. ISO 19115—Geographic Information-Metadata https://www.iso.org/standard/53798.html
2. ISO/IEC 11179—Information Technology-Metadata Registries http://metadata-standards.org/11179/
3. NCHRP Project 17-75, Leveraging Big Data to Improve Traffic Incident Management (TRB, 2019—publication forthcoming) http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4051
4. Integrating Emerging Data Sources into Operational Practice (FHWA, 2017) https://rosap.ntl.bts.gov/view/dot/34175
5. NCHRP Project 08-36 Task 130, Inventory and Assessment of Methods for Making Collected Transportation Data Anonymous (TRB, 2016) http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP08-36(130)_FR.pdf

Step 3.2

Load & Integrate Data

Establish repeatable data loading processes. Ad hoc data loading conducted in a rushed manner is a recipe for disaster. Repeatable processes need to be set up and, ideally, automated to load and transform raw data into a form suitable for use. When a problem occurs with a data load, procedures should be in place to roll back and then repeat the process once the issue is identified. Sometimes, a series of loads are needed to refresh data in various repositories. For example, new bridge inspection data may be loaded into a staging database for review and quality assurance. The data may then be transferred to the bridge management system database for analysis and to the agency's road inventory system. These data flows should be thoroughly tested, automated, and well-documented. Accurate and detailed documentation is essential, especially when data loads occur infrequently and there are multiple systems and staff from different business units involved.

Store both raw and processed data. Storing transformed performance data in addition to raw data can facilitate analysis and reporting.

Make use of data integration tools. There is a wide array of commercial and open source tools available supporting data integration processes. Some tools are geared to building extract-transform-load processes for data warehouse environments; others are geared to big data sets. Several excellent tools focus on integrating geospatial data. Use of these tools requires expertise and involves a learning curve, but it can save a great deal of time for data loading and integration tasks while also reducing the risk that errors are introduced through highly manual processes.

Case G

At MARC, use of automated processes and commercial data integration tools for maintaining key data sets has greatly simplified the process of querying, which means MARC is able to dedicate more time to analyzing data, not just collecting it.

Case K

Virginia DOT integrated pavement condition data with data on planned paving projects to produce performance monitoring reports that tracked anticipated versus actual changes in condition and the likelihood of achieving performance targets. The data integration effort relied on standardization of several data elements across two databases.

For more information...

1. NCHRP Synthesis 523: Integration of Road Safety Data from State and Local Sources (TRB, 2018) http://www.trb.org/Main/Blurbs/177990.aspx
2. NCHRP 08-36, Task 131: Transportation Data Integration to Develop Planning Performance Measures (TRB, 2017) http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP08-36(131)_FR.pdf
3. Informational Guide for State, Tribal and Local Safety Data Integration (FHWA, 2016) https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa16118.pdf
4. Data Integration Primer (FHWA, 2010) https://www.fhwa.dot.gov/asset/dataintegration/if10019/if10019.pdf

Step 3.3

Assess & Improve Data Quality

Data quality assessment. Poor quality data may have significant impacts on calculated performance metrics and therefore impact TPM decisions. Step 2.2 discussed the importance of planning for data quality as part of data acquisition and outlined the contents of a data quality management plan. However, there may be already-existing data sets needed for TPM that are of unknown quality. A data quality assessment can be conducted to determine suitability of a data set for use in TPM. Quality assessment can consider multiple characteristics, including completeness, currency, accuracy, and consistency. Data accessibility and interoperability are also sometimes considered. Assessing data quality involves establishing data quality metrics and measurement methods. For example, a metric for crash data completeness might be the percentage of data records that are missing a location code. This could be measured through a simple data query. Accuracy is typically assessed through a combination of independent verification for a sample of the records and application of validation checks to make sure measured values are within expected ranges.

"Data that is loved tends to survive."

Kurt Bollacker

Quality management. Quality management is a continuous process that starts prior to data acquisition and continues through the entire data life cycle. It should include analysis and flagging of data records that fail specific quality policies and thresholds. For example, pavement roughness measurements less than 30 inches/mile or travel speeds over 150 mph might be flagged as suspect.

Case D

I-95 Corridor Coalition data use agreements contain explicit data quality specifications that ensure 3rd-party-provided data meets required quality standards to support TPM.

It is important to find the right balance when planning for data quality improvement. All too often, agencies spend large amounts of resources attempting to clean, scrub, and validate data—only to find that there continue to be data issues regardless of how much time and energy is spent in cleaning. Perfection becomes the enemy of good, and agencies end up never fully using the data to inform decisions. Worse, the department (or person) responsible for the data hides it or prevents others from using it due to potential issues, fear, liability, etc. As soon as data (in any form) become available, it can and should be analyzed for data quality and consistency. The act of analyzing data, even when it has not been cleaned or validated, is important for guiding and informing potential users, applications, and data investment decisions.

Annotating (not discarding) suspect records. When suspect data records are encountered, a methodical process should be followed to flag these records and address the gaps in a carefully planned manner. Bad data records should not be summarily deleted because this could cause downstream analysis problems. Discarding bad data could negatively impact calculations if data gaps are not properly addressed. One way to address suspect or missing records is to fill gaps with historical data or otherwise imputed data. When this approach is used, these imputed records must be flagged to ensure that TPM decisions account for this simulated/modeled input.

In some situations, marking and tracking bad data can provide important information that can be used to improve future data quality. For example, a crash data manager might observe a pattern of inaccurate or incomplete crash records from one particular source. As another example, a traffic sensor may exhibit a specific pattern where it reports erroneous data (or does not report data) every day during the 8am hour. If bad data is discarded or not otherwise tracked, this particular failure (and its cause) may never become evident and subsequently continue to impact metrics based on that data.

For more information...

1. Development of a Computational Framework for Big Data-Driven Prediction of Long-Term Bridge Performance and Traffic Flow (Midwest Transportation Center, 2018) https://rosap.ntl.bts.gov/view/dot/36042
2. Crash Data Improvement Program Guide (NHTSA, 2017) https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812419
3. National Performance Management Research Data Set (NPMRDS)—Speed Validation for Traffic Performance Measures (Oklahoma Department of Transportation, 2017) http://www.okladot.state.ok.us/Research/FinalRep_2300_FHWA-OK-17-02.pdf

Capabilities Checklist: Store & Manage Data

Basic

- ☐ Data needed for TPM is stored in databases that are managed and regularly backed-up to provide protection from unauthorized access and corruption.
- ☐ Back-ups are tested on a regular, established cycle (e.g., monthly).
- ☐ Quality control procedures are in place to flag records that do not meet established validation criteria.
- ☐ Data dictionary information (metadata) is maintained and stored in a standardized fashion.
- ☐ Annual data snapshots are created for coordinated reporting across data programs.

Advancing

- ☐ Hardware and software requirements for data storage, updating, integration, and access are understood.
- ☐ Central data repositories have been established to integrate data from multiple sources and provide source data for reporting and analysis.
- ☐ Cloud and hosted storage options are considered for larger and more complex data sets.
- ☐ Data retention policies and archiving protocols have been updated to reflect lower storage costs and analysis of TPM business data needs.
- ☐ A range of data storage options are available to support databases with high transaction volumes and memory-intensive calculations as well as archived data retained for future use.
- ☐ Standards have been adopted to enable combining data from different sources.
- ☐ Data from multiple sources are fused to assemble a more complete and accurate data set than would be possible from any single source.
- ☐ Where appropriate, edge computing techniques are used—involving data processing at the source (e.g., at the site of the field sensor) rather than within a centralized repository.

Do's and Don'ts

Do:

- ✔ Consider cloud storage to reduce or minimize the agency's IT footprint and make it easier to scale storage up or down based on need.
- ✔ Explore hosted solutions from partners—universities, consultants, and sister agencies—to provide cost-effective approaches to managing large and complex data sources.
- ✔ Explore how fusing of disparate data sources can add value to your analysis and capabilities.
- ✔ Build or hire expertise in statistical analysis and computer programming to effectively analyze and transform data into TPM-related information.
- ✔ Adjust your agency's data retention policies and storage architectures so that potentially useful data isn't destroyed permanently.
- ✔ Establish repeatable, automated, and documented data-loading processes.
- ✔ Take advantage of commercial data integration tools.

Don't:

- ✘ Delete older data. The minute you get rid of it, you'll find you need it again.
- ✘ Delete erroneous data records. Flag them instead.
- ✘ Aggregate data sets to the lowest common denominator to save on storage space.
- ✘ Let the allure of "big data" technologies prevent you from continuing to invest in proven solutions.
- ✘ Rely on ad hoc approaches to loading and integrating data.

Step 4
Share Data

This step includes sharing transportation performance data across business units within an agency, across agencies, or with the general public. This includes but is not limited to transmitting data and reports to meet reporting obligations.

"There's a digital revolution taking place both in and out of government in favor of open-sourced data, innovation, and collaboration."

Kathleen Sebelius, Former United States Secretary of Health and Human Services

Agencies benefit from sharing data through improved coordination across jurisdictions, enhanced understanding of joint priorities, and leveraging of investments.

Note: this step focuses on the mechanics of data sharing and reporting, including tool selection. See Step 5 for a discussion of data analysis and Step 6 for a discussion of communicating data.

Step 4.1

Establish Reporting & Presentation Infrastructure

Select and deploy analysis and reporting tools. Data analysis and reporting tools that are available to agency staff are a critical element in making effective use of data. These can include tools that fuse "siloed" data from disparate sources, tools that fill in gaps (missing data), and those that identify or screen data 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. Tools include commercial business intelligence packages that support both traditional reporting as well as dashboards: GIS tools, statistical analysis packages, and specialized tools geared to particular types of performance data—for example, asset management systems and analytics platforms for congestion performance reporting.

While it is unlikely that a single reporting and analysis tool can meet all of the agency's needs, it is important to keep in mind that every new tool requires support to bring on new releases, train users, and troubleshoot issues. It is best to follow a disciplined and coordinated process of defining needs and requirements and considering whether existing tools are sufficient prior to bringing on a new tool.

Make build versus buy decisions. Developing the appropriate analytics software and databases that make the data easier to analyze and accessible to end users can be a significant hurdle for agencies. For an agency to build successful tools independently, they will typically need to draw upon the expertise of software engineers, system architects, user interface and user experience design specialists, developers, and project managers. The tools will need to be maintained over time; therefore, ample documentation and knowledgeable staff are needed that can be called upon over the course of many years to keep the tools up to date. Building complex tools with extremely small teams can be risky and costly to an agency.

Because of the high barrier to entry and continuing maintenance costs of developing custom tools, many agencies are now choosing to either purchase off-the-shelf tools or to leverage tools that other

Case E

Maryland State Highway Administration (SHA) uses Regional Integrated Transportation Information System (RITIS) visual analytics to combine disparate data sets and derive valuable information as part of after action reviews for operational improvements.

agencies/universities have already paid to develop. This effectively creates a pooled-fund approach to software development and maintenance. This approach is becoming easier for those agencies who are unaccustomed to purchasing services and for those who have historically not adopted tools and products that were not developed in-house or even within their respective states.

Whether an agency decides to build their own tools, hire consultants to build custom tools, or leverage existing tools, the following items should be considered.

In-House Development

- **Allocate ample time** to working on requirements for usability, functionality, and recruiting multiple user groups to get an understanding of expected usage.
- **Find an experienced partner.** Attempt to procure the services of a consultant who has performed similar work for other agencies. Analysis tools 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. Ensure the chosen provider has documented procedures that can be shared with software engineers and data analysts. Some providers have multistate/agency steering committees that collectively drive the features of the archive products to ensure they are constantly meeting user needs.

Purchasing Tools

- More and more states and MPOs are starting to purchase probe-based speed data; however, not as many agencies are investing in tools to analyze the data that enables better decisions. Probe data vendors, for example, have analytic tools that are sold at prices that are less expensive than the effort needed to reproduce those tools inside of the agency. These tools dramatically improve the

productivity of analysts by making the data much easier to analyze. In addition, these tools provide capabilities to agencies that might have previously taken months of effort to produce.
- Most vendors can provide return on investment (ROI) examples—including case studies from previous applications—showing how much money other organizations have been able to save by investing in 3rd-party data analytics tools or 3rd-party data. The I-95 Corridor Coalition has produced these types of ROI documents for its member agencies showing the benefits of some of their probe and incident data analytics products. More information can be obtained at www.i95coalition.org, or by reaching out to any 3rd-party data provider.

Purchasing Services

- For agencies that are not comfortable using analytic tools and are not interested in doing in-house data analysis, hiring outside consultants or university support may prove to be a viable 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 to bigger analysis tasks—adding layers of complexity and building on 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 mandating that consultants use specific tools, technologies, or techniques to deliver a solution. New technologies, methodologies, and tools are developed quickly and often. 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.

For more information...

1. NCHRP Project 03-128, Business Intelligence Techniques for Transportation Agency Decision Making (TRB, report forthcoming) http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4352
2. Development of a Travel-Time Reliability Measurement System (Minnesota Department of Transportation, 2018) http://www.dot.state.mn.us/research/reports/2018/201828.pdf
3. Implementation of Probe Data Performance Measures (Pennsylvania Department of Transportation, 2017) https://rosap.ntl.bts.gov/view/dot/32283

Step 4.2

Establish Data Standards & Formats

Take advantage of data standards. There are a number of data standards that can be adopted for agency data sets and/or used when sharing transportation system performance data between agencies (see Table 2). Some data standards cover data dictionary information (data elements and their definitions); others are more comprehensive and specify data formats, message structures, and technical mechanisms and protocols for sharing.

"The wonderful thing about standards is that there are so many of them to choose from."

Grace Murray Hopper

Data standards can make sharing processes easier. However, standardization should not be a prerequisite for sharing. It is more beneficial to share non-standard data than to not share anything. While standards are absolutely necessary in some instances (such as for vehicle-to-vehicle safety communications), the use of standards can break down in practice. Standards may become cumbersome because they try to address every possible data element and use case, or alternatively, standards are extended with custom fields and effectively lose the benefit of being a standard.

Agencies may be asked to comply with a standard imposed by an external entity as a condition for data sharing. This can have unintended consequences if that standard requires data to be "dumbed down" to the lowest common denominator to satisfy the needs of the external entity. In order to comply with this standard and remain on budget, agencies may permanently modify their data to match that standard and therefore lose significant value of that data for future use.

This issue of standards becomes even more challenging when dealing with big data. Unstructured data and crowdsourced data are rarely standardized or clean, but still may have substantial value to an agency to support TPM. The key to successful data sharing is to adhere to standards when possible, but not at the cost of losing insight or capability from non-standard data sets.

Table 2. Example data standards related to TPM.

Data Type	Applicable Standards
Safety	American National Standards Institute (ANSI) Standard D16 Manual on Classification of Motor Vehicle Traffic Crashes Model Minimum Uniform Crash Criteria (MMUCC) (https://crashstats.nhtsa.dot.gov/Api/Public/Publication/812433) Model Inventory of Road Elements (MIRE) (https://safety.fhwa.dot.gov/rsdp/downloads/fhwasa17048.pdf)
Pavement Condition	HPMS Field Manual—defines pavement data elements AASHTO Standard R43-13, Standard Specification for Transportation Materials and Methods of Sampling and Testing, Standard Practice for Quantifying Roughness of Pavement
Bridge Condition	FHWA National Bridge Inspection Standards
System Performance	IEEE 1512-2006 Standard for Common Incident Management Message Sets for use by Emergency Management Centers ITE TMDD 3.3 ITE TMDD Traffic Management Data Dictionary (TMDD) Standard for Center to Center Communications ASTM E2665-08 Standard Specifications for Archiving ITS-Generated Traffic Monitoring Data
Other	Open Geospatial Consortium—variety of standards for geospatial data All-Roads Network of Linearly Referenced Data (ARNOLD) manual—provides guidance and best practices for building linear referencing systems (LRS) covering all public roads National Information Exchange Model (NIEM)—provides a common vocabulary that enables efficient information exchange across diverse public and private organizations.

Select file formats. Certain file formats have advantage over others when it comes to sharing data between agencies. For example, exchanging PDF files containing detour plans may make sense on an individual case basis, but it significantly reduces the ability to automatically process information and incorporate it in TPM processes. Ideally, data should be formatted in a machine-readable format that provides the most flexibility for integration in TPM tools.

Common data file formats found in open data platforms include JSON, XML, CSV, and KML.

For more information...

1. Open Geospatial Consortium http://www.opengeospatial.org/standards
2. Project Open Data https://project-open-data.cio.gov/
3. General Transit Feed Specification http://gtfs.org/
4. National Information Exchange Model https://www.niem.gov/
5. National Bridge Inventory Resources https://www.fhwa.dot.gov/bridge/nbi.cfm
6. USDOT JPO ITS Standards Program https://www.standards.its.dot.gov/
7. Manual on Classification of Motor Vehicle Traffic Crashes Eighth Edition—ANSI D16 (Association of Transportation Safety Information Professionals, 2017) http://www.atsip.org/ANSI_Ver_2017_D16.pdf
8. HPMS Field Manual (FHWA, 2016) https://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/
9. Traffic Monitoring Guide (FHWA, 2016) https://www.fhwa.dot.gov/policyinformation/tmguide/tmg_fhwa_pl_17_003.pdf
10. All Roads Network of Linear Referenced Data (ARNOLD) Reference Manual (FHWA, 2014) https://www.fhwa.dot.gov/policyinformation/hpms/documents/arnold_reference_manual_2014.pdf
11. AASHTO Standard R43-13, Standard Specification for Transportation Materials and Methods of Sampling and Testing, Standard Practice for Quantifying Roughness of Pavement, 34th/2014 Edition (AASHTO, 2014)

Step 4.3

Publish Data

Designate authoritative data sources. Authoritative data sources for performance measure calculation should have been established as part of Step 1.3—Identify Analysis and Reporting Requirements. In preparation for publication, it is also important to designate authoritative sources for the computed performance measures and for any contextual data to be provided in the reports. Only designated authoritative sources should be used for reporting. Following this guideline will ensure that information released to the public is consistent and quality-checked.

Determine what data to share. The growing "open data" movement is creating the need for agencies to decide what data to proactively make available to the public, what data to provide on request, and what data to keep restricted. Several states have developed policy guidance on data classification. For example, the District of Columbia defines five levels:

- Level 0—Open (the default classification)
- Level 1—Public, Not Proactively Released (e.g., due to potential litigation risk or administrative burden)
- Level 2—For District Government Use (exempt from the Freedom of Information Act but not confidential and of value within the agency)
- Level 3—Confidential (sensitive or restricted from disclosure)
- Level 4—Restricted Confidential (unauthorized disclosure can result in major damage or injury)

DC has adopted the philosophy that data should be open by default and restricted only if there is a reason to do so.

Select data sharing methods. Sharing methods can vary from very basic file transmission, such as FTP, to more complex asynchronous, persistent transmission methods such as subscriptions, web services, and others. Open data sharing platforms such as data.gov have been established at the federal level and by many state agencies. While simple methods may be quick and inexpensive to implement, they can, in some situations, diminish the value of shared data. For example, files posted to an FTP site once a day introduce unnecessary latency and reduce certain TPM capabilities.

To support system performance management, agencies should strive to share data in near real time and at the highest possible resolution in order to provide the most flexibility and usefulness. For example, the Maryland State Highway Administration (SHA) allows external entities to securely subscribe to their real-time operations data sharing interface, which pushes incident information out to subscribers as it is entered by operators into their Advanced Traffic Management System (ATMS) platform. This approach allows partners to integrate this data into their system and become aware of significant incidents as soon as they occur (as opposed to minutes or hours later). This is important because it enables better real-time tracking of incident clearance times, responder activities, and other measures that are often requested by senior managers.

Share data within the agency. Departments within agencies often invest in data collection and data services to satisfy specific needs. For example, operations groups may procure and install sensors to support real-time operations. Planning groups may install different devices with a slightly different configuration to support planning and modeling needs. However, agencies frequently fail to evaluate existing investments within the agency. For example, there could be significant overall cost savings if agency departments evaluated existing data sets within the agency and adjusted the existing configurations or agreements rather than going through a completely separate procurement process. This is particularly true for larger agencies

Share data with other agencies. Sharing data with other agencies provides significant benefits to all parties, as well as the traveling public. Access to other agencies' data allows a more holistic approach to TPM, as well as better coordination in efforts to improve performance. Some of the challenges to sharing data with other agencies include data sharing methods, formats, and agreements.

It is also important for agencies to develop methods for the integration of external data. Separating relevant data from noise is an important exercise that can have a significant impact on TPM output. For example, an agency integrating incident data from a neighboring agency may want to only focus on external incidents close to the border, major regional incidents, or external incidents that may have an impact on the agency's area of responsibility. This means that for TPM purposes, the agency must develop a policy for identifying incidents of significance to the agency and its system while avoiding the trap of throwing away too much data that may be of use in the future.

Case C

The Florida DOT created an open data portal for sharing data both internally and with the public. Several FDOT business units had already begun to post important data sets online at various web sites. The portal provided a central location for data sharing, making it easier for people to locate available data.

Case F

The Metropolitan Area Transportation Operations Coordination (MATOC) Program enhanced real-time data sharing among agencies. This allowed agencies to become aware of incidents more quickly, to respond more quickly, to clear the incident more quickly, to alert travelers more quickly, and to develop standard operating procedures (SOPs) that account for impacts of regional and cross-jurisdictional events.

Sharing data with the public. Public agencies have a responsibility to provide best possible service to their customers: the general traveling public. One component of this responsibility is sharing of agency's performance with the public. While some data elements may be sensitive, most transportation data can be shared to better inform the public regarding system performance. In addition to open data, agencies can provide easily digestible and interactive reports regarding system performance.

One challenge with open data is that it is exposed to a general public that has varying levels of understanding of raw data. This can lead to distorted interpretation of data. While this is a real challenge, it should not be a barrier to sharing data with the public. Agencies can provide metadata, documentation, and sample applications to help users better understand raw data and its potential uses.

Application programming interfaces (APIs) allow users to develop their own data ingestion and processing applications and add value to existing data sets. In order for the agency to effectively distribute information, it must be able to share data via APIs for integration with other applications and systems. For example, the City of Chicago publishes APIs for historical congestion estimates, average daily traffic counts, and other TPM-related data sets. Similarly, the New York State DOT publishes APIs for historical traffic and transit events across the state and New York City.

Consider data sharing agreements. When agencies share data with each other, there is frequently a need for some type of an agreement or memorandum of understanding. Interagency agreements—especially those between states—have sometimes been difficult to negotiate because of governing law language and other restrictive terms and conditions. Because of this, many agencies are now opting to make their data "open" to the public—eliminating the need for data sharing agreements. Other agencies opt for informal, hand-shake agreements. Informal agreements can work well, though agencies with significant staff turnover will want to document any agreements in place.

Investigate public–private data sharing arrangements. Over the last decade, the private sector has emerged as an important partner when it comes to transportation data to support TPM. Private-sector data providers have been able to leverage technology and innovation in ways that public agencies are often unable to do. The concept of sharing data with the private sector has become both more important and more prevalent in recent years. Not only are agencies benefiting from obtaining new data sets from the private sector, but they are also benefiting from the private-sector value-add to the existing agency data sets.

Agencies must be careful about negotiating data sharing contracts with private-sector entities. In particular, agencies should pay particular attention to data use restrictions and seek maximum flexibility in use of data. This includes the ability to share data with universities and partner agencies and the ability to generate and share reports and summaries with the general public. Agencies, in turn, should treat the private sector as equal partners who can assist in disseminating information to the public and providing valuable insight in customers' behavior and travel patterns.

Provide tools for easy data access. Data has little value if it is not easily accessible. With continued improvement in bandwidth capabilities, web-based tools and data portals are becoming the norm. These tools allow users to log in and access data from anywhere with an internet connection. In addition to web-based access, the user interface and efficiency of the applications are critical. Poor user interfaces can make it difficult to understand what data and capabilities are available. Similarly, executing a query on a data set and waiting several hours or even days to receive an answer is unacceptable. Users must be able to quickly define a question and receive a response to make data and information useful. This means that agencies need to go beyond establishing databases or big data platforms and ensure that appropriate tools exist to access, visualize, and manipulate data for TPM. In many cases, more than one type of tool will be required to meet the needs and skill sets of different types of users. For example, some agencies make available one reporting package for technical staff and "power users" and a second for more casual users.

Case D

The I-95 Corridor Coalition collaboratively developed a public–private partnership between member agencies and 3rd-party data providers to take advantage of the latest private sector data offerings. They created a liberal and flexible model data use agreement that has become the "gold standard" for agencies and consortiums across the country for over a decade.

For more information...

1. Data Presentation on Transportation Agency Websites: Trends and Best Practices (Caltrans, 2017) http://transweb.sjsu.edu/sites/default/files/1501-data-presentation-on-transportation-agency-websites-trends-and-best-practices.pdf
2. Uses of Geospatial Applications for Transportation Performance Management (FHWA, 2016) https://www.gis.fhwa.dot.gov/documents/Uses_of_Geospatial_Applications_for_Transportation_Performance_Management_Case_Studies.pdf
3. State of the Practice on Data Access, Sharing, and Integration (FHWA, 2016) https://rosap.ntl.bts.gov/view/dot/35860
4. NCHRP Synthesis 460: Sharing Operations Data Among Agencies (Transportation Research Board of the National Academies, 2014) http://www.trb.org/Publications/Blurbs/170868.aspx
5. Geospatial Tools for Data Sharing: Case Studies of Select Transportation Agencies (FHWA, 2014) https://rosap.ntl.bts.gov/view/dot/12147

Capabilities Checklist: Share Data

Basic

- ☐ Employees are aware of key performance data sources within the agency.
- ☐ There are clear agency policies in place that data should be shared unless the need to protect it is demonstrated.
- ☐ There are protocols defined for how to share data to meet different needs that consider use of state and federal open data portals and hosted or cloud solutions.
- ☐ Open data portals are used to share data.
- ☐ Data explanations are provided in "plain English" to help users understand meaning, sources, and limitations.

Advancing

- ☐ Data governance and stewardship structures have been established to facilitate communication about data sharing and identify opportunities for synergies across business units for collaborating or combining data sources.
- ☐ Data sharing agreements are used (internal to an agency and between an agency and its partners) that specify what data will be shared, when and how it will be shared, and establish a clear understanding of data limitations and expectations for use.
- ☐ Data are shared in formats that are designed to meet the needs of different users, which may include standard reports, data feeds, and dashboards.
- ☐ Data with sensitive elements are sanitized for public distribution.
- ☐ Data contracts and sharing agreements are reviewed to ensure that agency flexibility is retained.

Do's and Don'ts

Do:

- ✔ Strive to open your data up to partner agencies and the public.
- ✔ Make sure that your data sets are ready to be shared by putting in place some standard criteria (no sensitive information, passed basic quality review, from an authoritative source, etc.)
- ✔ Treat other agencies as partners with whom you want to share your data so that you can improve systems safety and reliability.
- ✔ Put your data into standard formats when it is simple and improves upon your capabilities.

Don't:

- ✕ Assume a one-size-fits-all data feed will work for both the public and your agency partners.
- ✕ Sign data sharing agreements with restrictive "governing law" language.
- ✕ Let a lack of standardization become an excuse for not sharing data.

Insight

Step 5

Analyze & Use Data

The **Analyze & Use** step begins once data are converted into information. Data consist of values and figures that on their own have limited value. Information is a result of data that have been processed, organized, and interpreted to provide insight. Analyzing and using data for TPM involves consumption of data by analysts, planners, managers, engineers, and operations personnel to inform decision making or direct real-time system management. In the Store & Manage step, reporting and analysis tools are selected and configured. In the Analyze & Use step, these tools are used to support decision making. Moving from installation of a tool to productive use of the tool requires, at a minimum, the following:

- Identification of the intended users and uses for the tool;
- Designation of one or more individuals to develop specialized expertise with the tool (or engagement of a consultant to play this role);
- Training and support for additional users of the tool; and
- Iterative application and adjustment to tool parameters and configuration.

"Distinguishing the signal from the noise requires both scientific knowledge and self-knowledge; the serenity to accept the things we cannot predict, the courage to predict the things we can, and the wisdom to know the difference."

Nate Silver

Step 5.1

Analyze Trends

Assemble data. Assemble historical performance data for as many years as possible. If there have been changes to measurement methods, document when these occurred, but do not discard the older data. Even if there are discontinuities in the trend line, each section of the line can still be instructive for understanding how performance changed within each applicable time period.

Review and analyze the data. Plot the data to visualize variations over time. If appropriate, apply smoothing techniques such as moving averages to reduce noise. Use statistical techniques to distinguish the underlying trends in the data from seasonal variations and one-off variations due to events or other exogenous factors. Involve someone with expertise in statistics to be sure that the methods being applied are valid for the data being analyzed.

Case E

Maryland SHA used an incident timeline tool and graphics showing queue buildups and delay costs to help convince the responder community to change its policies for blocking lanes.

For more information...

1. "Time-Series Review of Highway Performance Monitoring System Data," Transportation Research Record (TRB, 2018) https://journals.sagepub.com/doi/10.1177/0361198118767415
2. Applying Safety Data and Analysis to Performance-Based Transportation Planning (FHWA, 2015) https://safety.fhwa.dot.gov/tsp/fhwasa15089/data_anl.pdf
3. Silver, N. The Signal and the Noise: Why So Many Predictions Fail—But Some Don't (Penguin Books, 2012)
4. Trends in Non-Fatal Traffic Injuries 1996-2005 (NHTSA, 2008) https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/810944

Step 5.2

Identify Patterns & Causes

Visualize data. Many people think of visualization as an end product—something produced after an analysis is complete to help communicate the results (or a story) to the public. While this is often the case, visualization can also be leveraged during the analysis life cycle as a way to better understand what is in your data, to identify outliers, and even to point out flaws that exist in your data. Interactive visual analytics can lead to insights earlier in the TPM process, sometimes more so than at the end.

Interpret data. Involve a group of experienced analysts in interpreting the observed trends. Look for correlations between performance trends and factors such as changes in revenues or budget allocations, fuel prices, economic conditions, or legislation/regulation. Use statistical packages and available analytical tools to analyze correlations. Develop insights that can be communicated to stakeholders (see Step 6—Present & Communicate Data.)

Case A

Arizona found in its Long-Range Transportation Plan process that for some performance areas, good outcome-oriented performance curves can be established. Where this was not possible, however, ADOT relied on simple curves reflecting the percent of identified needs met at a given allocation level. The lesson was to "not let the perfect become the enemy of the good."

For more information...

1. Approaches to Presenting External Factors with Operations Performance Measures (FHWA, 2018) https://ops.fhwa.dot.gov/publications/fhwahop18002/fhwahop18002.pdf
2. Historical Performance Evaluation of Iowa Pavement Treatments Using Data Analytics (Midwest Transportation Center, 2017) https://intrans.iastate.edu/app/uploads/2018/03/iowa_pvmt_tx_historical_perfomance_eval_using_data_analytics_w_cvr.pdf
3. An Analysis of the Significant Decline in Motor Vehicle Traffic Fatalities in 2008 (NHTSA, 2010) https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811346

Step 5.3

Predict Future Performance

Create predictive models. Develop realistic assumptions about future work based on available revenues, program budgets, and improvement programs. Use available analytical tools to predict future performance based on funding levels and/or specification of planned improvements. These include pavement and bridge management systems, safety analysis tools, travel demand models, and other specialized simulation tools. In addition, predictive analytics tools are available that make use of a variety of statistical techniques, including machine learning to predict future performance based on available data.

Applying analytical tools typically involves initial calibration—adjusting model parameters so that predictions are in line with observed conditions. This is followed by an iterative process of testing different assumptions and reviewing results for reasonableness. Every model has limitations; it is the role of an analyst to understand and explain these limitations.

Predictive models generally require specialized expertise to set up and use. Significant modeling tasks can be outsourced if this expertise is not available in-house. However, staff with analytical skills, patience, and interest in modeling can be trained to take on ownership and apply these tools as well as oversee work of contractors.

There are variations in available predictive tools for different performance areas, and it can take time to develop robust modeling capabilities. Agencies can start with basic approaches to prediction that rely on expert judgment and rules of thumb. As long as methods are clearly documented and caveats are stated, these approaches can provide value.

Case B

Caltrans developed a unified approach to presenting predictions of asset performance and need that combined results from mature pavement and bridge management system runs with analytical methods based on available data and expert judgement for other assets. The approach involved combining data from multiple disparate data sources that were at different levels of completeness and based on different analysis methodologies at varying levels of sophistication.

For more information...

1. FHWA Pavement Management Quarterly Webinar—Pavement Performance Modelling (2018) https://connectdot.connectsolutions.com/p9octngm9vv9/
2. Bhavsar, P., Safro, I, Bouaynaya, N., Polikar, R., & Dera, D. "Machine Learning in Transportation Data Analytics," Data Analytics for Intelligent Transportation Systems, pp 283-307 (Elsevier Inc., 2017)
3. Predicting Travel Time Reliability Using Mobile Phone GPS Data (Microsoft Research, 2016) https://www.microsoft.com/en-us/research/wp-content/uploads/2016/11/GPS_flow_sensing_travel_times.pdf
4. Paz, A., Veeramisti, N., & de La Fuente-Mella, H. Forecasting Performance Measures for Traffic Safety Using Deterministic and Stochastic Models (IEEE, 2015) https://www.researchgate.net/publication/308864167_Forecasting_performance_measures_for_traffic_safety_using_deterministic_and_stochastic_models

Step 5.4

Establish/Update Targets

Integrate results of trend analysis and predictive analysis. Establish a baseline value based on the trend line. Integrate the results of trend analysis and performance predictions to set targets for future performance.

Document the analysis. Documentation should include data sources, the steps taken to prepare and combine them, any key assumptions or parameters used (e.g., inflation rates), and observations about data anomalies or correlations. Good documentation will enable the analysis process to be repeated in the future by other staff members and will serve as a valuable resource if questions come up about the results.

For more information...

1. Safety Target Setting Factsheet (Arkansas Department of Transportation, 2018) http://www.tpm-portal.com/wp-content/uploads/2016/02/ARDOT-Target-Setting-Safety.pdf
2. Safety Performance Management Targets for 2018 (California Department of Transportation/Office of Traffic Safety, 2018) http://www.dot.ca.gov/fed-liaison/docs/Safety-Performance-Management-Targets-for-2018.pdf
3. Performance Management Target Setting Webinar—Pavement and Bridge (PM2) (California Department of Transportation, 2018) http://www.dot.ca.gov/assetmgmt/documents/Webinar_Slides.pdf
4. NCHRP Report 706: Uses of Risk Management and Data Management to Support Target-Setting for Performance-Based Resource Allocation by Transportation Agencies (Transportation Research Board of the National Academies, 2011) http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_706.pdf
5. NCHRP Report 666: Target-Setting Methods and Data Management to Support Performance-Based Resource Allocation by Transportation Agencies: Volume I: Research Report: Volume II; Guide for Target-Setting and Data Management (Transportation Research Board of the National Academies, 2010) http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_666.pdf

Case A

Arizona DOT's Long-Range Plan process developed performance curves for different investment categories, including preservation, modernization, and capacity. These curves were used to analyze the impacts of a change in investment on different performance outcomes.

Capabilities Checklist: Analyze & Use Data

Basic

- ☐ Analysts are aware of and taking advantage of existing commercial off-the-shelf, open source, and publicly available tools for analysis, visualization, forecasting, and scenario analysis.
- ☐ Analysts are trained in use of data analysis and visualization tools.
- ☐ Private-sector or university contractors are used to provide data analysis services as alternatives to setting up analysis capabilities in-house.
- ☐ Data are available that are sufficiently accurate to meet analysis requirements.
- ☐ Visualization and analysis tools are used to explore and discover data anomalies and limitations.
- ☐ Data preparation and analysis tasks are well defined and planned to ensure sufficient calendar time and staff resources.
- ☐ Analysts are able to identify trends and causal factors.
- ☐ Data element meanings, data transformations, and analysis assumptions are documented.

Advancing

- ☐ Predictive models for key transportation performance measures are validated based on multiple cycles of application.
- ☐ Targets are established based on predictive analysis relating revenues and programmed work to performance results.
- ☐ Data mining is conducted to support "back-casting"—which involves starting with a future vision and analyzing current and historical data to estimate changes required to move from the current situation to the future vision.
- ☐ Cooperative arrangements across agencies have been established to transform data into information (e.g., the state DOT performs analysis of travel-time reliability, computes measures for each facility, and provides the data for use by MPOs and local agencies).
- ☐ Predictive analytics and machine learning techniques are applied for predicting asset failure probabilities and other performance measures.

Do's and Don'ts

Do:

- ✔ Begin to analyze and visualize available data to help uncover potential quality issues.
- ✔ Document your analysis process so that someone else could trace your steps in the future.
- ✔ Explore 3rd-party tools to get your agency up and running quickly with data analytics capabilities.
- ✔ Invest in developing in-house analysis capabilities or, if that isn't feasible, engage consultants that specialize in data analysis.
- ✔ Treat your consultants as trusted team members.

Don't:

- ✕ Wait until all of your data is "cleaned" or perfect before beginning an analysis.
- ✕ Build an analysis tool in-house unless you are sure about what you are getting into and you are confident that you will have staff to support the tool in the future.
- ✕ Be so detailed and stringent in your analysis requirements that costs escalate out of control.
- ✕ Just go with your current on-call consultant or low-bid contractor (if they aren't data analysis experts).

Step 6

Present & Communicate Data

The **Present & Communicate** step involves developing effective ways of communicating the message and story behind the data.

"The greatest value of a picture is when it forces us to notice what we never expected to see."

John Tukey

The process of communicating performance results is likely to lead to questions about the data and analysis. Data analysts should anticipate that there will be iteration between the communication and analysis steps. The need for data improvement or augmentation may also be identified as new questions arise. Over time, these improvements will strengthen the agency's ability to make effective use of data to improve performance.

Step 6.1

Develop & Communicate Performance Stories

Tell the story. Once data are successfully translated into information, it is important to provide context and the "so what" surrounding that information. One of the most effective ways to accomplish this is through a storytelling approach. Information consumers (the audience) must buy into the story for the information to be effective. The focus of the story must not be on data and information, but on the message that the information is supporting. For example, when the New Jersey Department of Transportation and the Delaware Valley Regional Planning Commission were trying to convey the importance of specific roadway projects to senior managers and public officials, they were challenged to communicate about complex performance measures related to reliability, safety, congestion trends, economic impacts, and more.

After many unsuccessful attempts at producing thorough reports for decision makers, they tried an information visualization approach. They developed "elevator pitch" brochures that conveyed, primarily through graphics, the performance measures related to individual projects. The visualizations contained within the brochures could be easily interpreted by both engineers and the public. Accompanying narratives were short, and the brevity of the brochures meant that more people ultimately read and understood the message.

When making a new investment in sensor infrastructure to support TPM, the message should not be that 200 more sensors will provide more data about congestion. Instead, the message should be that the new sensors will allow operators to more quickly identify traffic slowdowns, which will enable signal timing changes to ensure that inbound commuters make it to work on time. The audience must be able to relate to the outcomes and understand how they are being affected by changes in performance of the system.

Case H

New Jersey DOT developed project assessment summary pamphlets that tell a compelling story about how investment in a project benefitted the general public.

In 2012, an article in *Governing* magazine reported, "The [Gray] Notebook tells Washington citizens pretty much whatever they might want to know about how their transportation system is working.…The first Gray Notebook—as it came to be called because of the color of its cover—was published in 2001, and legislators loved it. Two years later, those legislators approved a 5-cent increase in the gas tax to fund new transportation projects."

Communicate with visualizations. Information visualization is critical in supporting successful data presentation and communication; visualization is used to help people understand and interpret the data. Human perceptual skills are remarkable—we can identify trends, clusters, gaps, and outliers extremely quickly. The presentation of data and information in a visual way should be easily processed to identify the primary message. Some of the visualizations frequently used in TPM due to their simplicity and effectiveness include heat maps, timelines, choropleths, arc diagrams, etc., but there are hundreds of visualization strategies, methods, and devices that can be leveraged, depending on the analysis or communication need.

But just as visualizations can be powerful in communicating a message, they can also often be misleading. It is critically important to be consistent in presenting information to avoid introducing a bias to the user. There are many known data visualization approaches that can skew the message—truncating Y-axes, omitting data, correlating causation, and many others. In order to present information in the most objective way, it is important to provide the methodology, definitions, metadata, and any other contextual information.

For more information...

1. Communicating Performance (AASHTO) http://communicatingperformance.com/
2. The Colorado Transportation Story https://www.codot.gov/programs/colorado-transportation-matters/statewide-transportation-plans/statewide-transportation-plans
3. Visualization and Communication in Pavement Performance (Midwest Transportation Center, 2018) https://rosap.ntl.bts.gov/view/dot/36445
4. NCHRP Web-Only Document 226: Data Visualization Methods for Transportation Agencies (TRB, 2016) https://www.nap.edu/catalog/24755/data-visualization-methods-for-transportation-agencies
5. "Truth, Transparency and Transportation," Governing Magazine, September 13, 2012 http://www.governing.com/blogs/bfc/col-washington-state-transportation-gray-notebook-transparency.html
6. Tufte, E. The Visual Display of Quantitative Information (Graphics Press, 2001)

Case A

Arizona's 2040 Long-Range Transportation Plan broke new ground in its methods for aligning Arizona stakeholders' priorities with competing 25-year transportation needs. ADOT innovated by using a multi-objective decision analysis software platform with intuitive visual elements like slider bars, dashboards, and data visualizations that helped stakeholders see the consequences of alternate funding choices on performance outcomes.

Capabilities Checklist: Present & Communicate Data

Basic

- ☐ Managers and analysts meet to review and interpret performance results.
- ☐ Story lines for performance results are developed, reviewed, and communicated.
- ☐ Training is offered to internal staff to build skills in data presentation and communication.
- ☐ Staff have capabilities to present data in a variety of formats tailored to the needs of different audiences, including heat maps, thematic maps, timelines, and other infographics.
- ☐ A combination of narrative and graphical presentation is used to communicate performance information.

Advancing

- ☐ Feedback from data consumers is sought and used to improve communication of information to different target audiences.
- ☐ Individuals with expertise in data visualization and communication are available to support development of performance data products.
- ☐ Social media is used to communicate key results or draw people to more detailed communication products.
- ☐ Specialized visualization and analysis environments have been developed—e.g., virtual reality simulators.

Do's and Don'ts

Do:

- ✔ Leverage visualization tools for your data analysis *and* for communicating TPM to the public and decision makers.
- ✔ Employ "best practices" in visualization that aim to communicate with users, not deceive them.
- ✔ Leverage 3rd-party visualization tools and/or professionals to support your analysis.
- ✔ Use visualization to support your narrative.

Don't:

- ✕ Wait until the end of your project to begin to interpret the results.
- ✕ Use all the bells and whistles in a chart or visualization tool; clean and simple graphics tell compelling stories.
- ✕ Try to be too complicated with your visualizations; you're trying to tell a story, not confuse people.
- ✕ Expect visualization alone to tell your story; some supplemental explanatory text will be required.

Cases

Case A

Arizona DOT Long-Range Plan Investment Trade-offs

Like many states, Arizona has a gap between the transportation needs of its growing population and funding available to pay for those needs. Consequently, Arizona DOT (ADOT) and its stakeholders must make difficult trade-offs across transportation investment priorities. This case shows how ADOT utilized data to make these trade-offs as part of its What Moves You Arizona 2040 Long-Range Statewide Transportation Plan update (LRSTP or 2040 Plan). The 2040 Plan process used decision science techniques in combination with stakeholder engagement to quantify the likely performance outcomes of several investment scenarios designed by ADOT staff and stakeholders. This process guided the selection of a recommended investment choice (RIC) for the plan, which has provided the starting point to invest approximately a billion dollars a year in alignment with consensus transportation priorities of Arizona's citizens and businesses.

Overview

Arizona's LRSTP four-step planning process relied equally on data, public engagement, and use of multi-objective decision analysis (MODA) software. First, it mined rich engineering data about capital and operating needs and revenues. Second, it used these data within a public engagement process to identify stakeholder goals and priorities. Third, it used the MODA software to enable stakeholders to explore performance projections of Arizona's future transportation safety, congestion, and infrastructure condition under various alternate transportation futures propelled by divergent investment strategies. Finally, it used stakeholder input to inform a recommended investment strategy.

Multi-Objective Decision Analysis (MODA)

Multi-objective decision analysis (MODA) is a tool for resolving resource allocation problems where choices involve trade-offs among competing objectives that feature sacrifice of one objective for the sake of another. MODA uses data about stakeholders' preferences and decision outcomes to guide selection of optimum choices. Initially, the relative importance ascribed by stakeholders to different choices is scored using weighting techniques. Subsequently, outcomes of different choices are evaluated in terms of their relative alignment with stakeholders' priorities to arrive at an optimum solution.

Foundation: Specify & Define Data

Estimates of capital and operating needs by transportation investment category. Arizona Department of Transportation oversees a statewide system of major highways and supports transit, rail, aviation, and non-motorized transportation facilities around the state. The 2040 Plan used established data sources and modeling tools such as FHWA's Highway Performance Monitoring System (HPMS), National Bridge Investment Analysis System (NBIAS), and Highway Economic Requirements System—State Version (HERS-ST) to document $89.5 billion in baseline 25-year needs for all of the state's major transportation investment categories:

- ***Preservation investment needs*** to maintain pavement and bridges in good repair;
- ***Modernization investment needs*** for upgrades like safety improvements and intelligent transportation systems;
- ***Expansion investment needs*** for added lanes, new roadway alignments, or interchanges;
- ***Operations and maintenance investment needs*** for routine work, like patching potholes, fixing guardrails, mowing, and snow removal; and
- ***Non-highway investment needs*** for transit, rail, non-motorized, and aviation modes.

These estimates were derived from segment-by-segment data and analysis of all engineering work needed to achieve and maintain an acceptable level of performance throughout the state for each major investment category. In combination with funding information, they provided a baseline indicator of the potential investment trade-offs Arizona faces: highway revenue forecasts suggested that $23 billion would be available over the 2040 Plan's 25-year time horizon compared to $53 billion in highway needs.

Statewide transportation system goals and performance outcome measures. Building on information developed about the state's transportation needs, ADOT—with extensive input from its stakeholders—set broad goals for the state's transportation system in 2040 and measurable yardsticks for gauging progress toward them:

- ***Improve mobility, reliability, and accessibility.*** Implement critical/cost-effective investments to improve access to multimodal transportation and optimize mobility and reliability for passengers and freight. (Measures: congestion, speed, travel delay)
- ***Preserve and maintain the system.*** Maintain, preserve, and extend the service life of existing and future state transportation system infrastructure. (Measures: pavement and bridge deficiencies; maintenance spending)
- ***Enhance safety.*** Continue to improve and advocate for transportation system safety for all modes. (Measures: fatalities and serious injuries)

Stakeholders' relative priorities for these three transportation system-related goals helped inform the build-out of possible scenarios for investment across major investment categories like preservation, modernization, and expansion. The measures associated with each goal provided the basis for comparison of each scenario's system performance outcomes and its consequences for spending trade-offs.

Insight: Analyze & Use, Present & Communicate Data

Stakeholders' priorities and scenarios for Arizona's future transportation performance. The 2040 Plan broke new ground in its methods for aligning Arizona stakeholders' priorities with competing 25-year transportation needs. ADOT innovated by using a MODA software platform with intuitive interface elements like slider-bars, dashboards, and data visualizations that helped stakeholders see the consequences of alternate funding choices on performance outcomes. Importantly, the software does not pick the "best" investment strategy; rather it allows users to compare strategies and choose which one they like best based on data about performance outcomes. Both stakeholders and staff at ADOT were enthusiastic about the software's ability to show performance trade-offs in real time as users moved funding from one investment area to another.

The MODA software was customized with Arizona-specific "performance curve" algorithms for each investment need category, including preservation, modernization, and capacity.[1] The curves show how a rise or fall in spending changes performance outcomes. ADOT invested considerable effort in deriving each curve by processing raw transportation data that depicted performance outcomes at different spending levels to establish generalized rules. In the preservation category, for example, pavement and bridge curve algorithms were built from outputs of infrastructure management system data. In the capacity category, the performance curve algorithm was derived from projected changes in delay and reliability as sets of proposed projects are built.

Through a series of workshops and webinars, ADOT used the decision-support software to develop several distinct "alternative investment choice" scenarios that showed how different spending strategies might affect the state's transportation system performance:

- ***"Current Plan" Investment Scenario.*** ADOT projected future performance based on continuation of current capital spending without any changes.
- ***"Agency Plan" Investment Scenario.*** ADOT used an interactive process to discuss ADOT staff and outside stakeholders' preferences for investment (gathered via a webinar and an in-person workshop). Through this process, a consensus based on projections of future performance relative to a baseline was developed about allocation of resources.
- ***Public Investment Scenario.*** ADOT used MetroQuest, a web-based public engagement software platform, to gather general public input on highway investment priorities. Nearly 6,000 individuals provided their opinions about transportation priorities and potential trade-offs.

Arizona's recommended investment choice (RIC). At the culmination of ADOT's 2040 Plan process, the three scenarios developed during the planning process informed the 2040 Plan's RIC (as shown in

---

[1] Because annual operations spending levels are determined independently by the Arizona legislature and ADOT does not have the ability to allocate these funds to highway capital spending, Operation and Maintenance needs were excluded from the scenarios.

Figure 4) by providing a sense of different stakeholders' priorities for use of ADOT resources and the projected performance outcomes associated with them. The combination of data, tools for projecting future needs, public engagement, and MODA in the scenario development process helped build support for tough trade-off decisions about limiting expansion spending to avoid a future decline in system preservation performance.

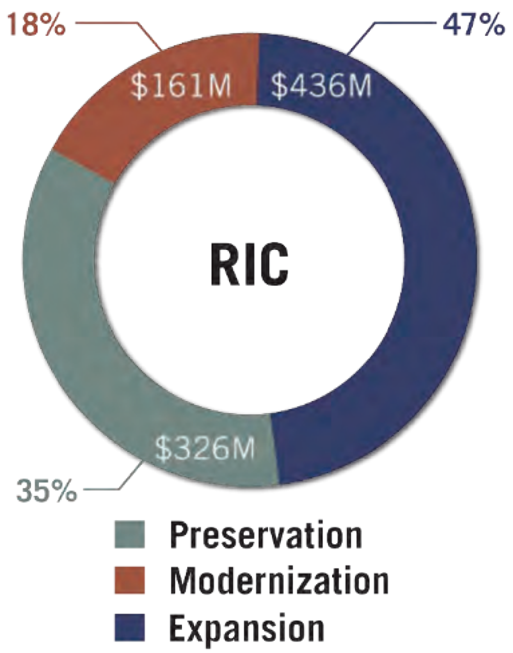


Figure 4. Annual funding levels—Arizona Long-Range Plan recommended investment choice (RIC).

Development of RIC using scenarios and MODA tools engaged stakeholders in Arizona in new ways and educated them about the importance of thinking about priority weighting, measures of performance, the necessity of trade-offs, and target setting. Based on the successful application of these techniques for the 2040 Plan, a similar MODA-based process will likely be used for translating What Moves You Arizona 2040 into ADOT's programming process.

Success Factors

Using a MODA-based scenario approach for long-range planning helped ADOT achieve a more informed recommendation for how to allocate scarce transportation funding in the future. Success factors included:

- **Taking advantage of available data sets and tools.** HPMS, NBIAS, HERS-ST, and asset management systems were used to create projections of future performance under varying investment levels.
- **Incremental approach.** ADOT recognized that advancement in analysis capabilities would be an incremental process. The agency was able to establish outcome-oriented performance curves for certain investment areas (e.g., expansion, preservation). For investment areas where ADOT could not develop outcome-oriented curves, it based the curves on the percentage of identified needs met at a given allocation level (e.g., safety, technology). While the methodologies for the curves differed, the resulting analysis was still valuable in informing investment decisions.
- **Interactive analysis.** ADOT provided interactive tools that enabled stakeholders to explore implications of different investment strategies. Stakeholders and staff alike benefitted from the ability to see information about performance trade-offs in real time as they collaborated to explore shifts in funding from one investment area to another.
- **Opening up the black box.** The interactive process educated stakeholders about the impacts of various model parameters and assumptions; for example, in the ADOT MODA tool, performance curves, performance thresholds, and criteria weights were key drivers of the results.
- **A data-informed decision-making philosophy.** ADOT recognized and acknowledged that decisions are informed—not made—by the data and analysis results. Communicating this philosophy alleviated stakeholder concerns that the analysis did not consider other non-quantifiable factors.

Challenges & Lessons

Communicating with stakeholders. Key challenges and lessons learned were related to communicating technical information to stakeholders in a clear and succinct manner and making sure they understood definitions, implications of assumptions, and limitations of the analysis:

- During the 2040 Plan development, some stakeholders were confused about the difference between "preservation" and routine "operations and maintenance." This may have led to preservation being underweighted.
- Many participants struggled with the pairwise comparison, which asks a respondent to identify the relative priority between two different

investment options and served as the basis for criteria weighting. This was either because they were not comfortable with some of the comparisons (e.g., how can you compare preservation and safety), or because they simply did not understand its purpose.
- The MODA approach did not enable users to consider synergies in spending, such as the benefits to safety or mobility that might come from increased preservation spending. Integrating consideration of investment synergies across performance areas is an area identified for future improvement.

For more information...

- Arizona Long Range Statewide Transportation Plan https://www.azdot.gov/planning/transporation-programs/state-long-range-transportation-plan
- Arizona DOT Point of Contact: Statewide Planning Manager

Case B

Caltrans State Highway System Management Plan

Caltrans developed an integrated State Highway System Management Plan (SHSMP) that summarizes conditions of existing transportation assets, communicates funding needs, and details projected funding levels. The plan addresses an extensive array of asset types, two different major investment programs, and forty different subprograms. This case illustrates how data across different sources at different levels of completeness can be integrated and put on a comparable footing for analysis and presentation.

Overview

At Caltrans, capital projects to preserve, rehabilitate, or replace existing transportation assets are included in the State Highway Operations and Protection Program (SHOPP). Caltrans manages a separate Highway Maintenance (HM) program for smaller maintenance projects. Together the annual budgets for these programs are projected to total over $4 billion per year over the next 10 years.

The California Streets and Highways Code requires Caltrans to prepare periodic updates to its SHOPP and HM programs. Historically these updates were made separately, in some cases drawing upon different data sources. In 2017, Caltrans developed a new, integrated SHSMP that incorporates state requirements for preparation of both a ten-year plan for the SHOPP and five-year HM plan. The SHSMP (shown in Figure 5) includes a Needs Assessment and Investment Plan to help guide the management of the state highway system and related infrastructure. The plan covers thirty-four different SHOPP subprograms and six maintenance subprograms. These subprograms address physical assets including but not limited to

- pavement
- bridges
- drainage systems
- lighting
- signage
- guardrail
- transportation management systems
- water/wastewater treatment
- rest areas
- facilities

Foundation: Specify & Define, Obtain Data

Data for the SHSMP were obtained from several different sources. For major assets (pavement and bridge), Caltrans has well-defined management systems and analytical processes. For selected assets (e.g., traffic management systems), Caltrans has established asset inventory databases but no formal analytical processes. For other assets (e.g., drainage systems), Caltrans has partial inventory information and relies on a sample of the inventory together with estimates of the extent of missing data to determine the overall size of the inventory.

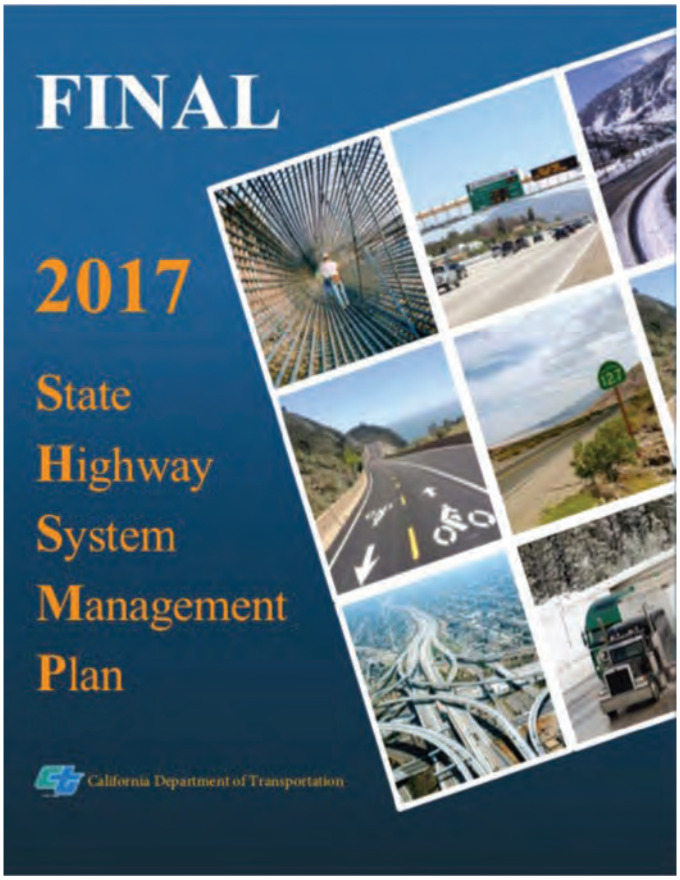
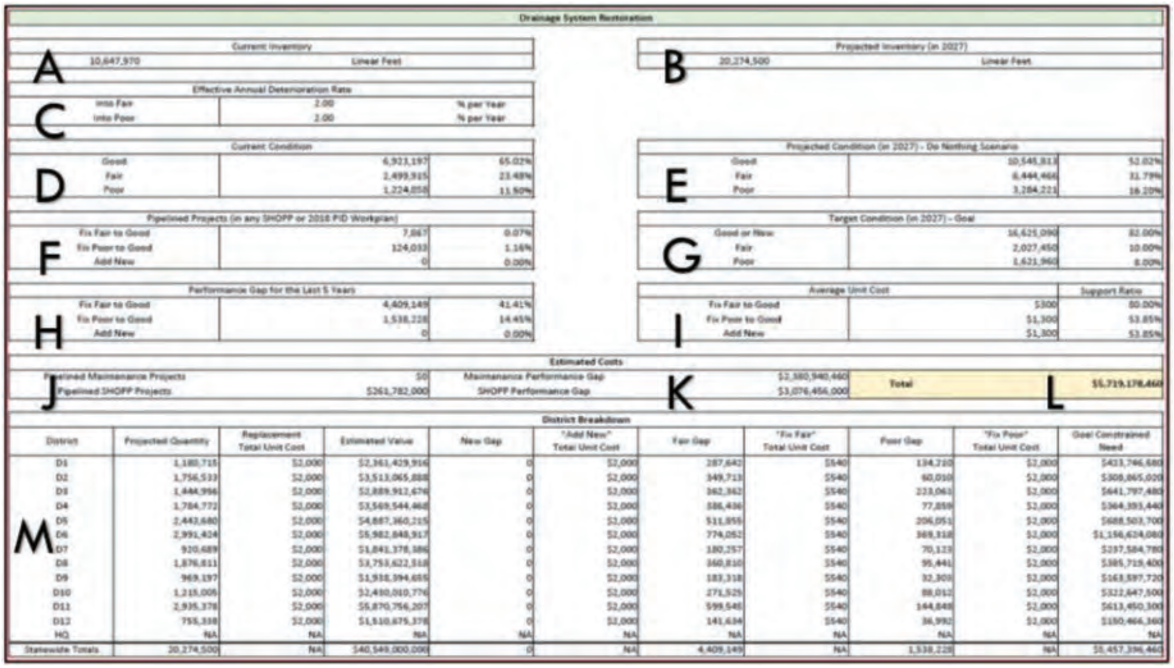


Figure 5. Caltrans SHSMP. (Courtesy of California Department of Transportation)

Reporting: Store Data

To facilitate development of the 2017 SHSMP, Caltrans requested data from asset data owners in whatever format was available and then developed a standardized means for representing data on the asset inventory, its condition, and investment needs. This standard representation was implemented in a spreadsheet, with separate sheets for each investment sub-program. Figure 6, reproduced from the SHSMP, illustrates the format used for summarizing and reporting the data. The summary view includes 13 distinct sections, including the asset inventory, projected inventory, predicted investment level, and total investment need.



Drainage System Restoration

Current Inventory	
10,647,970	Linear Feet

Projected Inventory (in 2027)	
20,274,500	Linear Feet

Effective Annual Deterioration Rate		
Into Fair	2.00	% per Year
Into Poor	2.00	% per Year

Current Condition		
Good	6,923,197	65.02%
Fair	2,499,915	23.48%
Poor	1,224,858	11.50%

Projected Condition (in 2027) - Do Nothing Scenario		
Good	10,545,813	52.02%
Fair	6,444,466	31.79%
Poor	3,284,221	16.20%

Pipelined Projects (in any SHOPP or 2018 PID Workplan)		
Fix Fair to Good	7,867	0.07%
Fix Poor to Good	124,033	1.16%
Add New	0	0.00%

Target Condition (in 2027) - Goal		
Good or New	16,625,090	82.00%
Fair	2,027,450	10.00%
Poor	1,621,960	8.00%

Performance Gap for the Last 5 Years		
Fix Fair to Good	4,409,149	41.41%
Fix Poor to Good	1,538,228	14.45%
Add New	0	0.00%

Average Unit Cost		Support Ratio
Fix Fair to Good	$300	80.00%
Fix Poor to Good	$1,300	53.85%
Add New	$1,300	53.85%

Estimated Costs					
Pipelined Maintenance Projects	$0	Maintenance Performance Gap	$2,380,940,460	Total	$5,719,178,460
Pipelined SHOPP Projects	$261,782,000	SHOPP Performance Gap	$3,076,456,000		

District Breakdown

District	Projected Quantity	Replacement Total Unit Cost	Estimated Value	New Gap	"Add New" Total Unit Cost	Fair Gap	"Fix Fair" Total Unit Cost	Poor Gap	"Fix Poor" Total Unit Cost	Goal Constrained Need
D1	1,180,715	$2,000	$2,361,429,916	0	$2,000	287,642	$540	134,210	$2,000	$423,746,680
D2	1,756,533	$2,000	$3,513,065,888	0	$2,000	349,713	$540	60,010	$2,000	$308,865,020
D3	1,444,956	$2,000	$2,889,912,676	0	$2,000	362,362	$540	223,063	$2,000	$641,797,480
D4	1,784,772	$2,000	$3,569,544,468	0	$2,000	386,436	$540	77,859	$2,000	$364,393,440
D5	2,443,680	$2,000	$4,887,360,215	0	$2,000	511,855	$540	206,051	$2,000	$688,503,700
D6	2,991,424	$2,000	$5,982,848,917	0	$2,000	774,052	$540	369,318	$2,000	$1,156,624,080
D7	920,689	$2,000	$1,841,378,386	0	$2,000	180,257	$540	70,123	$2,000	$237,584,780
D8	1,876,811	$2,000	$3,753,622,518	0	$2,000	360,810	$540	95,441	$2,000	$385,719,400
D9	969,197	$2,000	$1,938,394,655	0	$2,000	183,318	$540	32,303	$2,000	$163,597,720
D10	1,215,005	$2,000	$2,430,010,776	0	$2,000	271,525	$540	88,012	$2,000	$322,647,500
D11	2,935,378	$2,000	$5,870,756,207	0	$2,000	599,546	$540	144,848	$2,000	$613,450,300
D12	755,338	$2,000	$1,510,675,378	0	$2,000	141,634	$540	36,992	$2,000	$150,466,360
HQ	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Statewide Totals	20,274,500	NA	$40,549,000,000	0	NA	4,409,149	NA	1,538,228	NA	$5,457,396,460

A. Current inventory of physical assets
B. Projected future inventory of physical assets
C. Annual deterioration rates used to calculate projected condition
D. Current condition of physical assets
E. Projected future condition of physical assets (do-nothing scenario)
F. Asset quantities from work in the pipeline
G. Performance targets (not constrained)
H. Performance gaps
I. Average unit costs for repair and associated support ratio
J. Dollar value of unfunded future commitments
K. Dollar value necessary to close the performance gap
L. Total need to achieve the performance target
M. District level breakdown of inventory, gaps and needs

Figure 6. A SHSMP summary sheet with sections labeled. (Courtesy of California Department of Transportation)

Insight: Analyze & Use, Present & Communicate Data

A critical step in developing the integrated SHSMP was establishing a common approach to characterizing asset conditions and investment needs across investment areas and asset types. This was needed both to help communicate the plan and to standardize the analysis approach to allow for meaningful comparisons across the different areas. A complicating factor was that in some cases, such as for pavement, Caltrans had already implemented a sophisticated approach to predicting conditions and establishing needs, while in other cases such an approach had not yet been implemented.

Caltrans' approach to this step was to define criteria for good, fair, and poor condition for each asset/investment type and then characterize existing conditions in these terms. In some cases, such as where an asset's condition is assessed strictly in terms of age (e.g., for traffic management systems), the approach was further simplified to include good and poor conditions only.

Caltrans staff then established, for each asset/investment category, the likelihood of deterioration from good condition to fair condition (or poor, if fair is not defined) and from fair condition to poor condition (if fair condition is defined). For cases where a formalized management system had been established, these values were calculated based on model runs from the existing system. Alternatively, staff established the values based on supplemental analysis and/or expert judgment. Figure 7 shows an example of this calculation for bridges. Here 75 percent of the existing inventory is in good condition, and 0.45 percent is expected to drop to fair condition annually. An additional 22 percent is in fair condition, and 0.75 percent of these assets are expected to drop to poor condition annually.

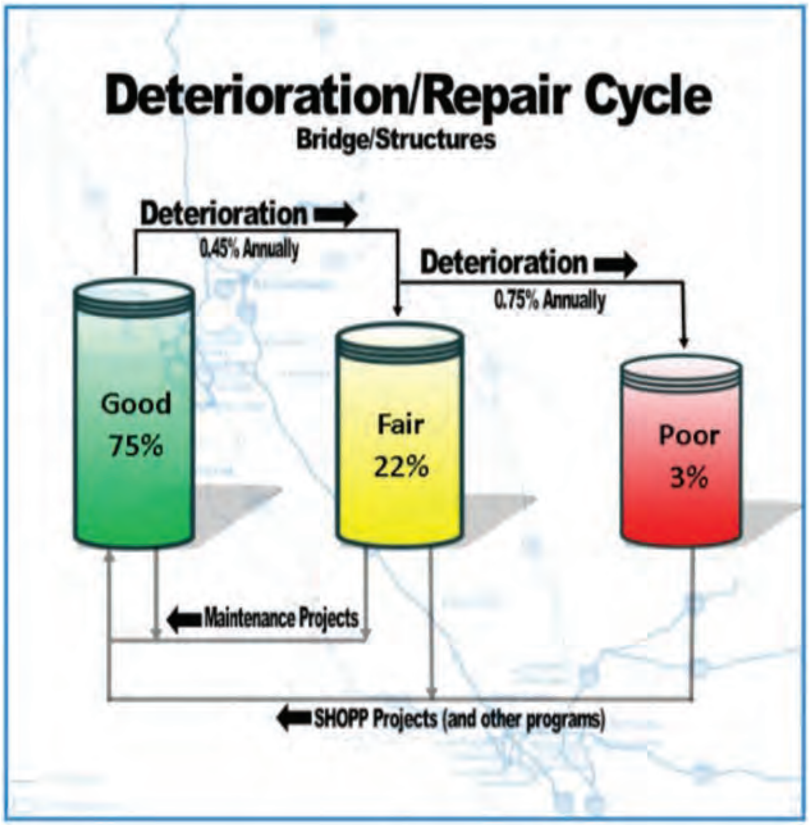


Figure 7. Example representation of good/fair/poor condition and predicted deterioration. (Courtesy of California Department of Transportation)

With this representation, future conditions at the end of a 10-year period can be approximated based on initial conditions, expected budget, and the average treatment cost in fair and poor condition. Caltrans performed this calculation for each asset/investment area to predict future conditions and needs. The end result was the development of an integrated plan using a common, easily communicated representation to establish and illustrate expected investment levels over a wide range of assets and investment areas in two different funding programs.

Success Factors

- **Alignment between analysis and available data.** The investment categories in Caltrans' SHOPP and maintenance programs are well defined, and in many cases they are aligned with specific asset classes. This simplified the process of determining what data were needed to develop the SHSMP and help address the organization of data in the plan.
- **Common performance measures across assets.** Development of a common approach for analyzing and summarizing asset/investment data was a critical step in preparing an integrated plan. One key insight was that using a simplified approach to presenting data (e.g., good/fair/poor asset conditions) allowed for an effective means to summarize data in cases where a more complicated approach was used for analysis while also providing a basic analytical approach in cases where only summary data were available on the asset inventory and its condition.
- **Streamlined presentation.** The resulting SHSMP uses standardized graphics for communicating conditions and deterioration rates. Details on each asset/investment area are included in an appendix to the document. Use of a standard approach helped simplify the presentation of the materials and streamlined document preparation.

Challenges & Lessons

Lack of integrated data system. Caltrans did not, as of the development of the 2017 SHSMP, have an integrated system for collecting and managing the data used to prepare the document. Caltrans is exploring the feasibility of implementing an integrated asset management system to help support development of the SHSMP and other related documents and plans in the future.

For more information...

- Caltrans 2017 State Highway System Management Plan http://www.dot.ca.gov/assetmgmt/documents/SHSMP.pdf
- Caltrans Point of Contact: State Transportation Asset Engineer

Case C

Florida DOT Transportation Data Portal

Open data portals have been developed at the federal, state, and local levels. The U.S. Federal Open Data Portal, deployed in 2009, provides access to more than 300,000 data sets on data.gov. U.S. DOT has created their own open data portals, including its.dot.gov/data/. States are also developing their own open data portals.

This case describes one state's approach to creating a data portal. Florida DOT's Transportation Data Portal was developed to help FDOT, the public, contractors, and other 3rd-parties locate and utilize data for internal and external projects. FDOT's portal is "a platform for locating data related to the core mission of the Florida Department of Transportation."

Overview

FDOT's goal was to provide a resource to allow people "to explore and download open geospatial data; analyze and combine open data sets using maps; develop new web/mobile applications, and more." FDOT's available transportation data includes everything from GIS shapefiles describing transportation facilities, aerial photography, documents, manuals, real-time and historical traffic counts, summary statistics, interactive web applications, assets, software, and much more. They created a data portal that meets the needs of multiple stakeholders, including internal employees, those doing business with FDOT, and the public.

Reporting: Share Data

FDOT took a multi-faceted approach to reporting and sharing data with their stakeholders. FDOT is a relatively decentralized agency with seven large districts and many different business units. Each business unit had already been making select data sets available on disparate websites. The Central Office eventually stepped in and decided to consolidate their transportation data, reports, and other resources together in a single site. Some business units provided their data directly to this new site while others simply had their existing websites linked from this central data portal. While this makes it slightly more difficult for users to locate everything they might need, the DOT is clearly making great strides toward open government and open data.

Insight: Present & Communicate Data

FDOT's data portal (shown in Figure 8) is more than just a listing of live and archived data sets. They also have applications, reports, and other materials that provide insights, present the data in a digestible format, and otherwise allow for interactive exploration.

Figure 8. FDOT's transportation data portal provides links to many FDOT data resources across different business units.

Figure 9 shows a user exploring aerial photography for a part of the state. Users can search for photos by year, specific date, location, format, etc. Results for small queries are shown immediately, while larger queries may require additional time or retrieval options that go beyond online access.

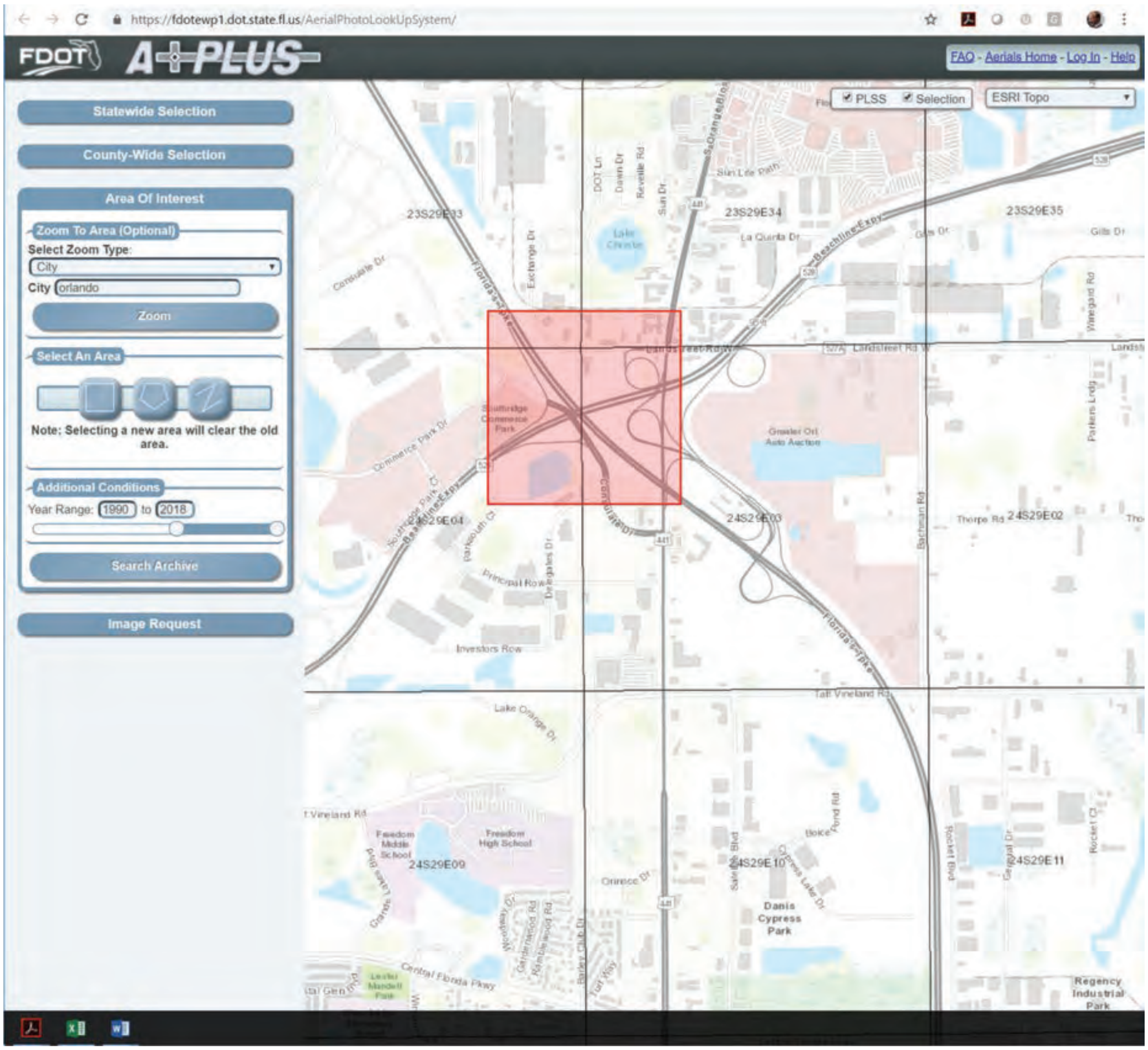


Figure 9. A screenshot of Florida's Aerial Photo Lookup System.

Figure 10 shows another interactive application on Florida's transportation data portal that enables a user to explore traffic monitoring sites, counts, districts, and summary statistics. There are many additional summary reports and analyses that can be explored with other applications—even apps dedicated to exploring data about the location of wildflowers, meadows, and other beautification projects on or adjacent to Florida roadways.

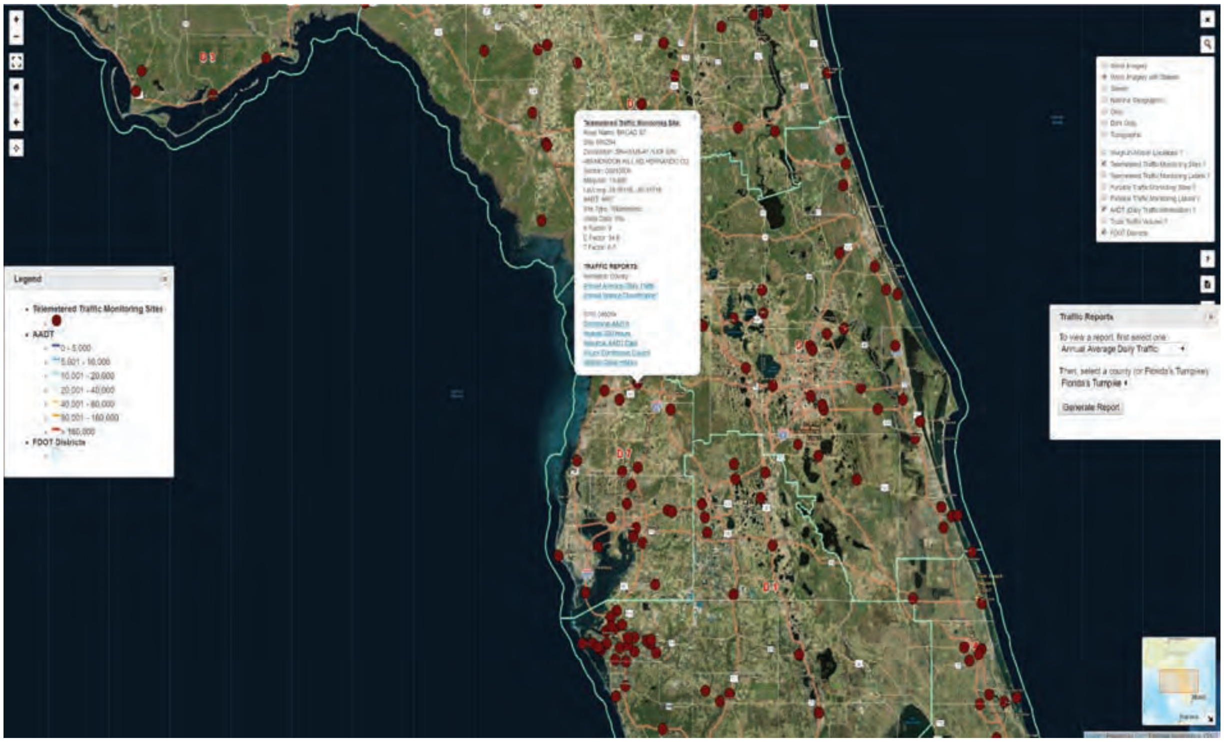


Figure 10. In addition to raw data sets, FDOT makes many data exploration websites available to the public. The image above is of FDOT's continuous count stations.

Success Factors

- **Legislative action:** In 2011, Florida's governor issued Executive Order 11-03 establishing the Office of Open Government. This order required the state to establish and maintain a website providing ready access to accountability information and required each Florida agency to establish an Open Government contact.
- **Leading by example:** Several business units within the DOT had already started to post important data sets online. This was done for several reasons, including trying to proactively keep consultants and the public from flooding FDOT phone lines and inboxes with data requests, thus freeing up employees to conduct other business. Certain FDOT business units also believed that providing data to the public could potentially spur innovative solutions to FDOT's growing transportation problems. The business units that were already successfully sharing their data with the public could tout their success and show the positive ROI to other business units that had not yet begun to share data.

Challenges & Lessons

Agency buy-in. Not every business unit initially thought that providing all of their data online was a good idea. Providing data to the public can open up the business unit (and the managers of the data) to increased scrutiny. Some data managers felt they had little to gain and much to lose. Others feared that making their data more easily accessible would lead to a devaluation of their own job. They were no longer the gate-keepers of the data, and some thought that was a job security concern. The lesson learned is to anticipate these concerns and address them head-on.

Organization and presentation. FDOT has tried very hard to make the bulk of their data available online. This includes a wide variety of data sets, including maps, asset inventory, and traffic counts. Because of the diversity in data offerings, organizing the data in a way that makes it easily discoverable has been a challenge. Also, some business units were already making their data available through disparate web pages. Joining all of these web pages together into a truly centralized one-stop shop is an ongoing challenge.

For more information...

- Florida Office of Open Government https://www.flgov.com/open_government/
- FDOT Transportation Data Portal http://www.fdot.gov/agencyresources/mapsanddata.shtm
- FDOT Open Data Hub https://gis-fdot.opendata.arcgis.com/
- U.S. DOT's ITS JPO Data Portal https://www.its.dot.gov/data/

Case D

I-95 Corridor Coalition Probe Vehicle Data Procurement

In 2007, the I-95 Corridor Coalition issued an innovative, multistate request for proposals (RFP) for purchase of private-sector, probe-based travel time. The individuals behind this procurement had several goals in mind:

1. Demonstrate the value of 3rd-party probe-based speed data to states,
2. Make it easier for states to procure data uniformly, and
3. Remedy private-sector acceptable use issues that had plagued agencies in prior contracts.

As a result of this RFP and follow-up efforts, the I-95 Corridor Coalition established one of the most liberal and flexible data use agreements that has become the "gold standard" for agencies and consortiums across the country for over a decade. This case illustrates how, while procurement mechanisms and data products have changed over the last decade, the underlying foundation of data ownership, acceptable use, quality expectations, etc., has not.

Overview

DOTs have been procuring traffic data and information services from the private sector for many years. Data use agreements—documents that state what can and cannot be done with private-sector data—are a standard component of these public–private data procurements. Unfortunately, many agencies end up with data use agreements that heavily favor the private sector and severely limit the agency's ability to utilize the data in a way that benefits everyone.

Most agencies write RFPs for data from the private sector in a vacuum. They may forget to talk to other stakeholders in their own agency—procuring the data for a single use only. They may not think strategically about future applications of the data. They may not seek out "lessons learned" from other DOTs who have procured similar data. And worst of all, they may neglect to specify acceptable use terms at all—leaving it completely up to the data provider.

For example, many agencies deployed and owned Closed Caption Television (CCTV) infrastructure, but contracted with 3rd parties to stream those videos internally and with their customers (television stations, agency traveler information sites, etc.). In those contracts, agencies ended up having to pay to view their own video or share it with others. The 3rd party monetized an asset that wasnottheirs by taking advantage of agencies' inexperience in negotiation of acceptable use agreements.

Similarly, some 3rd-party speed sensor data providers negotiated the installation of private-sector sensors on the public right-of-way in exchange for allowing the agency to view the data coming from those sensors. While at a glance that seems to be a reasonable partnership, that data exchange came with many strings attached, effectively preventing agencies from doing useful things with the data (like posting travel times on variable message signs or on the web) unless the agency paid significant additional fees. In effect, agencies traded valuable right-of-way for a data set of very limited value due to acceptable use agreements.

Foundation: Specify & Define, Obtain Data

To develop an agency-friendly and private-sector beneficial agreement, the I-95 Corridor Coalition's I-95 Vehicle Probe Project (Figure 11) took a bold approach of defining acceptable use terms that incorporated multistate sharing, perpetual use, data quality standards, unrestricted non-commercial use, and other innovative clauses.

Figure 11. I-95 Corridor Coalition's Vehicle Probe Project.

Agencies exchanged best and worst practices from their previous procurements and contracts and generated a list of necessary components for this agreement. Academics and consultants provided insights on the future of planning and operations to cover a wide range of situations. No stakeholder was left out of the discussion, and the Coalition did not shy away from demanding certain terms and conditions.

The RFP also encouraged data collection and delivery approaches that did not rely on installation of equipment on public rights-of-way as a means to spur innovation and protect agencies from trading their most valuable asset. Data quality specifications (including latency, accuracy, and availability) ensured that providers were held accountable for the quality of data they provided, regardless of technology used. Financial penalties ensured that providers maintained the quality of their data and processes throughout the length of contract. The University of Maryland was tasked to be an impartial and objective 3rd-party validator of data quality to ensure fair and equitable evaluation. Perhaps most importantly, the I-95 Corridor Coalition (and all state participants who agreed to the terms of the data use agreement) retains the rights to use data in any way it sees fit in perpetuity, with some limitations in sharing with other private entities for commercial purposes. This allowed 3rd party data providers to retain their right to resell data in commercial markets and maintain their competitive edge.

This flexibility allowed agencies to use data for a variety of applications ranging from real-time operations and traveler information to planning and project prioritization. The general public also benefitted from an empowered DOT that could suddenly react more quickly, provide better traveler information, and make data-informed decisions with respect to spending and construction.

Because the public sector did a good job setting clear expectations from the beginning, 3rd party data providers saw a marked improvement in their relationship with the public sector. Public-sector participants did not feel restricted or taken advantage of, and more time could be spent on innovating products and building relationships. Agencies and private-sector data providers truly worked together in a mutually beneficial partnership.

Success Factors

- **Collaboration.** Because agencies approached this procurement as true collaborators, they were able to leverage their collective knowledge and experience in contracting and procurement to create a strong, public, agency-friendly data use agreement.
- **Willingness to share.** The I-95 Corridor Coalition emphasized the need to share data across jurisdictional borders to support TPM efforts across entire regions. The importance placed on cross-jurisdictional sharing has led to innovation.
- **Strong champion.** The executive leadership of the Coalition was a strong proponent of this project. This leadership helped to push for the collaboration mentioned above and was extremely forceful in demanding some of the more innovative terms and conditions that had heretofore not been requested of data vendors.
- **Governance.** The Coalition established a steering committee made up of leadership from each state. This ensured that no single interest could dominate and kept all of the states actively involved.
- **Focusing on the end result.** This highly successful data use agreement was possible because agencies focused on the end result and allowed the private sector to innovate and meet the needs of agencies in a mutually beneficial manner. Agencies worked together to exchange knowledge and experience and create a common vision for better service to the general public.

Challenges & Lessons

Disparate needs across agencies. Each agency may have different priorities or focus areas depending on their geography—urban versus rural nature—or modes and types of traffic traversing their networks. In the past, this led agencies to think that collaborative solutions would not meet their unique needs. However, the I-95 Corridor Coalition data use agreement showed that despite differences, agencies have many common goals, and with carefully drafted language that is not overly prescriptive, agencies could rely on innovative private-sector solutions to satisfy many disparate missions.

Change in established vision of public–private relationships. Through this process, both agencies and private-sector data providers had to adjust to a new kind of relationship. Agencies had to let go of their previous tendencies to view vendors with suspicion, while private-sector providers had to make some concessions in short-term profits in order to establish a positive relationship that would result in overall long-term benefit. Both had to act in true partnership.

For more information...

- I-95 Corridor Coalition Probe Data RFP http://i95coalition.org/wp-content/uploads/2015/02/RFP_82085N_Final_Final.doc?x70560
- I-95 Corridor Coalition Vehicle Probe Project Documents http://i95coalition.org/projects/vehicle-probe-project/
- I-95 Corridor Coalition Data Use Agreement http://www.i95coalition.org/wp-content/uploads/2015/02/VPPII_DUAv9_signed.pdf
- Data Use and Application Guide http://i95coalition.org/wp-content/uploads/2015/03/008-7G_VPP_DATA_USE_Report_Final_April_2011.pdf?x70560

Case E

Maryland State Highway Administration's Incident After Action Reviews

Many agencies conduct after action reviews (AARs) for major incidents. The purpose of the AAR is to bring together the key operations, personnel, and responders involved in the incident in order to reflect on the successes and failures of their response with the goal of improving future performance. This case highlights the power of data to build a common understanding of what happened and a consensus on needed improvements.

Overview

The most mature transportation operations agencies conduct weekly AARs on all types and categories of events in order to build teams, enhance communication, and continually improve. Some agencies use a manual tracking process—operators and responders fill out paper forms to document AARs. These forms may be completed days or even weeks following an incident and thus rely on foggy, imprecise memories. Then, information from the forms is compiled and used to write reports or discuss the incident in small groups. This manual process can be tedious, leading to less complete data capture and a less-than-enthusiastic group of AAR participants.

The Maryland State Highway Administration (MD SHA) has taken a decidedly different approach to conducting AARs that involve automated, electronic data capture and reporting. The result is a more effective AAR that is engaging, easy to conduct, and informative. This approach encourages more frequent AARs and allows the operational response of the agency to be quantified and tracked over time.

Foundation: Specify & Define Data

MD SHA identified the need for incident data in their business and strategic plans (see "For more information…" at the end of this chapter). They collect data about where incidents are occurring, how many vehicles are involved, which responders (both agencies and vehicle types) have been notified about the incident, when they arrived on the scene, and when they departed. They capture the name of responders, certain tasks that they performed on scene, the road surface conditions, and lane closings/openings over the course of the incident, as well as other operator notes related to information flows, radio communications, and more. These data were deemed necessary not only for real time operations management, but also for MD SHA's annual performance evaluation and benefit analysis, which helps to justify their operations program and annual budget.

Reporting: Store & Manage Data

Maryland's operations platform has been customized and refined since the mid-1980s to collect highly detailed data. The data are collected and stored at MD SHA. Data are archived, time-stamped, and attributed to individual incidents.

Incident data are then combined with data from ITS devices [Dynamic Message Signs (DMS), Closed Caption Television (CCTV) images, volume and speed detectors, signals] and probe-based speed data used to derive queue buildups and congestion levels.

Maryland SHA's data is transmitted in real time to the Regional Integrated Transportation Information System (RITIS) platform that supports reporting and analysis for AARs and performance evaluation.

Insight: Analyze & Use, Present & Communicate Data

Maryland has developed a series of reporting and visualization tools that are used to help spur discussion and facilitate a dialogue among the responders and operations personnel.

They begin by producing a timeline graphic of the response to the incident (Figure 12).

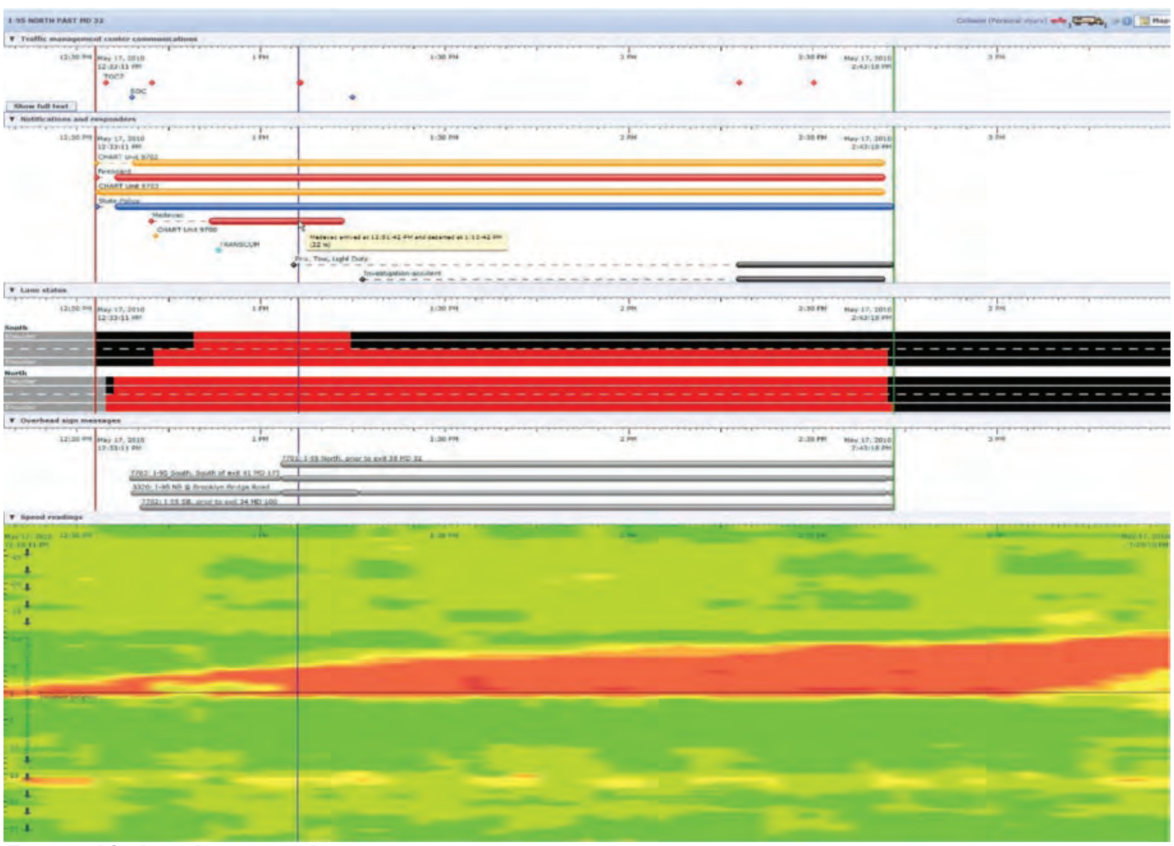


Figure 12. Incident timeline.

This timeline includes every event recorded during the incident, including when responders were notified about the incident, when they arrived, and when they departed. It includes communication logs, DMS activations, queue buildups, photos and videos of the event, and indications of which lanes were blocked over time.

Within the timeline, they can expand the list of operator notes and communication logs to see the flow of information between the responders. Figure 13 shows the communication log.

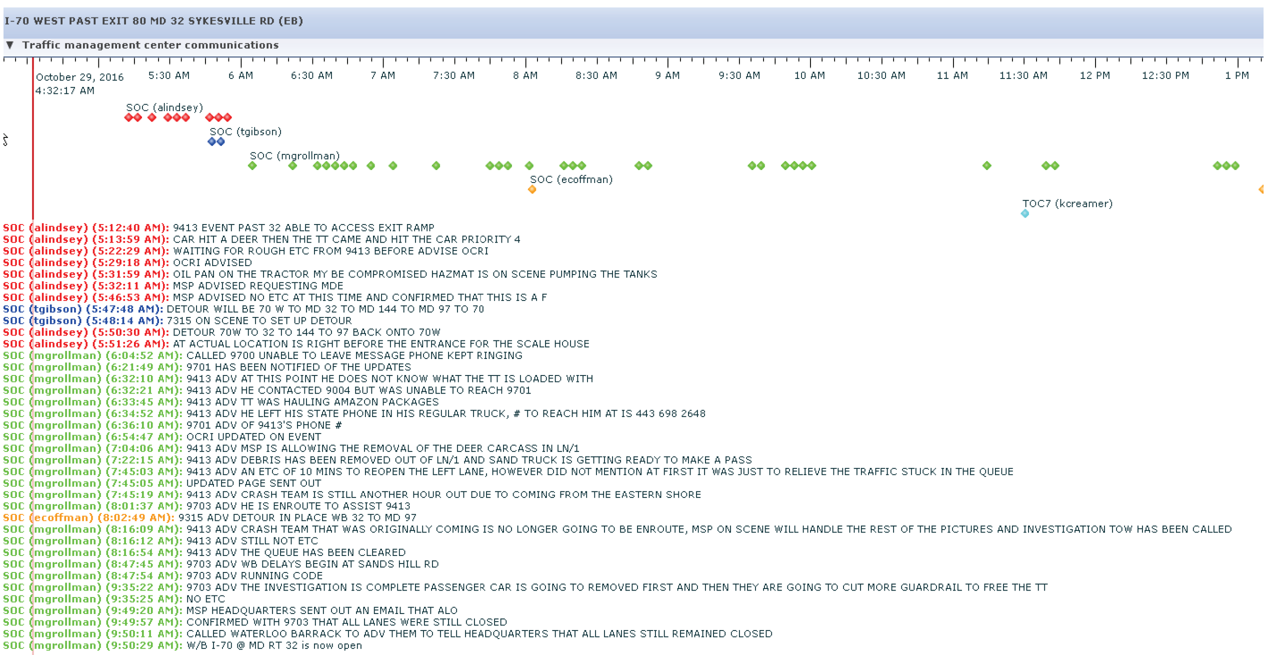


Figure 13. Communication log.

They also generate animated maps that show queues building (and receding) over time (Figure 14). These maps typically show adjacent roads and arterials so that the agency can better understand how their actions affect others. Animated maps can also be placed side by side to showcase traffic during a particular incident compared to normal traffic conditions.

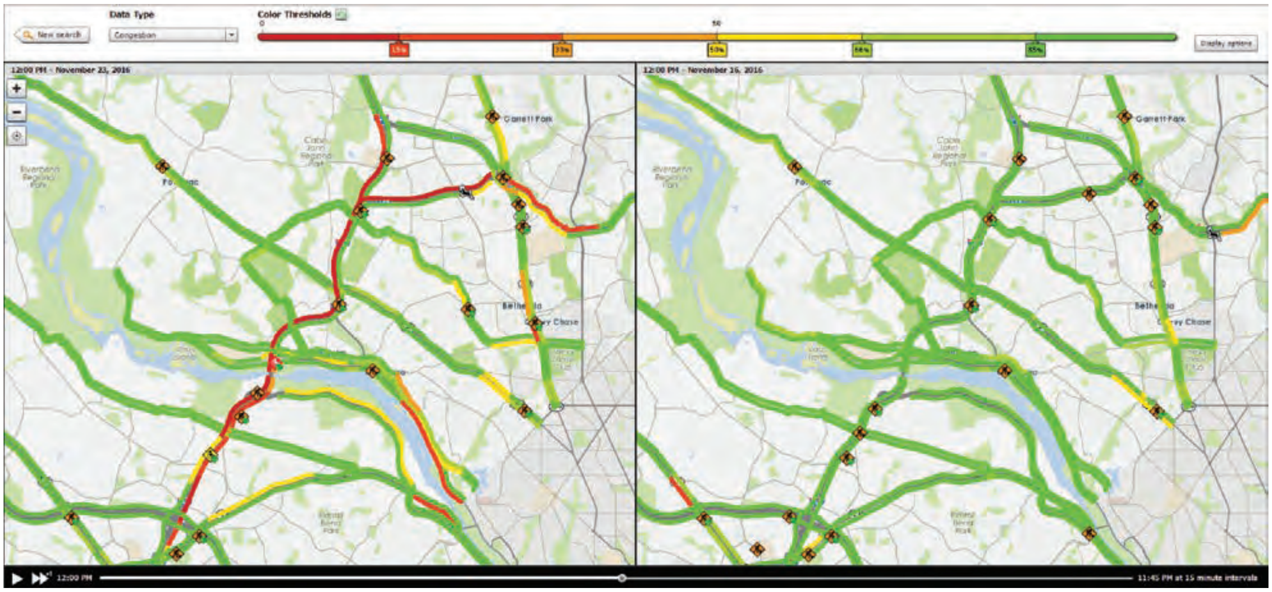


Figure 14. Animated maps.

Side-by-side congestion scan graphics (Figure 15) also show how queues built up and subsided during the day of the event compared to similar days of the week when no incidents occurred.

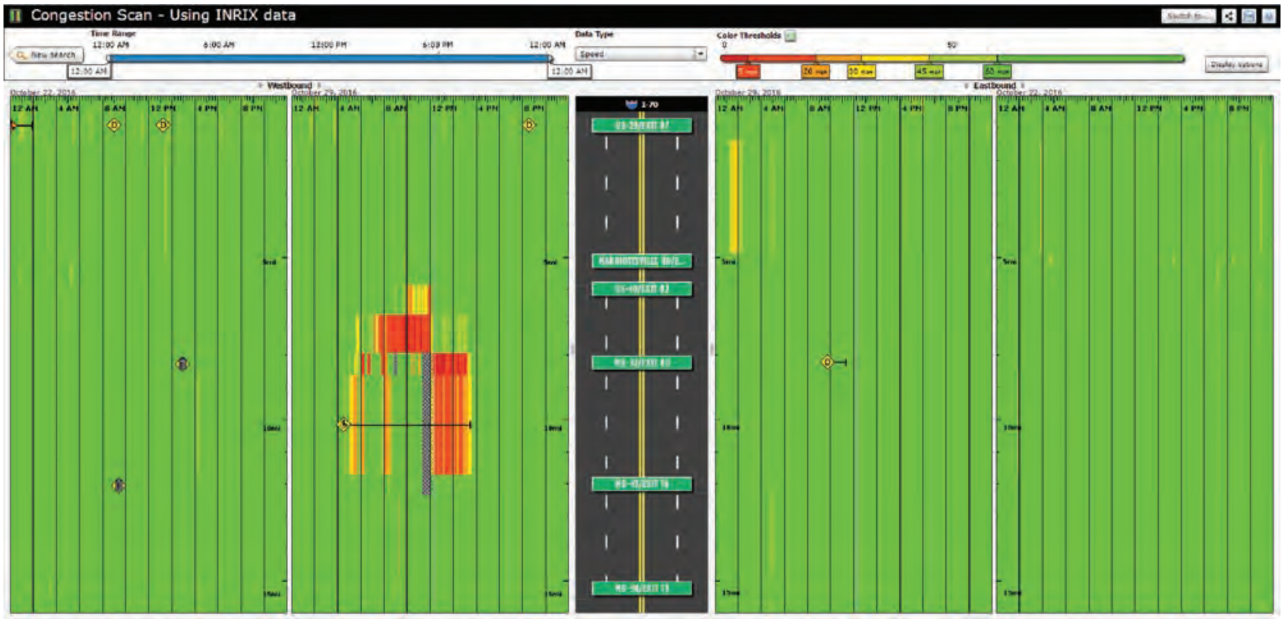


Figure 15. Congestion scan.

Heat maps then help the agency understand if this location is a high-crash location (Figure 16).

Figure 16. Heat maps.

Video images at different locations and time points (Figure 17) are available to provide additional documentation of the incident.

Figure 17. Example incident documentation supporting AARs.

Finally, user delay cost graphics (Figure 18) help the agency visualize the social financial cost of the delays. For example, the cost of typical user delay on I-495 in Maryland (including the connecting arterials) would be about $150k/weekday. However, during one particularly bad incident (shown above from images captured by MD SHA in RITIS), the cost of user delay skyrocketed to over $1.2M. This conservative estimate did not account for delays in the opposite direction of travel, excess fuel consumption, emissions, or the cost of equipment damage.

Figure 18. User delay cost graphics.

In less than an hour, MD SHA is able to create powerful slide decks from the above data reports. These slide decks are then sent out to the responders who participate in AAR meetings in-person or remotely.

During AAR meetings, the slide decks documenting the incident are used to facilitate a discussion. Each agency comments on what the data is telling them, including how they responded, how they communicated, etc. This process helps the responders determine what they did well and what they could have done better. The RITIS reports validate the discussion and lend credibility to the responders' accounts of what happened. These reports also help the responders to understand the societal impacts of their actions (or inactions), which can help drive policy. Two examples illustrate how insights from the reports led to changes in operations policies:

- In one AAR meeting, MD SHA was able to make the case that the state police needed to adjust how they manage their tow list to ensure that towing companies have the necessary equipment for heavy duty operations. This recommendation was based on clear data from the incident timeline that showed multiple tow dispatches and long arrival times.
- Firefighters tended to close many lanes (or all lanes) of the roadway during their response, resulting in high delay costs. However, information captured in the timeline tool, queue graphics, user delay cost graphics, and additional reports on secondary incidents have

helped convince the responder community to change its policies for blocking lanes and move more toward a stronger quick-clearance mentality.

Success Factors

- **Specifying and obtaining the right data.** The performance reports described above were only possible because MD SHA spent considerable time over the last couple of decades defining data needs and dedicating funding and operator training to ensure that all necessary data is collected.
- **Analysis tools tailored to decision maker needs.** Analysis tools provide quick access to data and show the benefits of quick-clearance practices and the value of transportation systems management and operations programs.
- **Effective visualizations.** The reports and visualization provide the agency ammunition for requests for funding, positions, and equipment. MD SHA's early investments in data and analytics are paying off.
- **Commitment to data-driven decision making.** In the past, AARs were more about "war stories" than data analysis. As a result, they reinforced or justified existing behaviors rather than provide an opportunity for new insights. Now, however, data, tools, and processes are in place to conduct regular AARs, and those tools provide data-backed conclusions. The agency can be more confident in its decision making, and the tools assist MD SHA in making the case to external (and internal) partners about improving current practices. Over time, the agency will be able to analyze trends along individual corridors and quantify the effects of actions taken based on the AARs.

Challenges & Lessons

Making the case for investing in data. Operators already face demanding jobs, and asking them to collect more data was an uphill battle. Early education and advocacy were needed at all levels. Senior management had to be convinced that the extra workload would be worth the effort.

Funding and implementation. Even after the agency made the decision to collect more data, it took a great deal of time to raise funds and enhance systems to add new data fields, train staff, and see a return on the investment. It is important to keep in mind that implementation takes time and to set appropriate expectations.

For more information...

- The Maryland DOT CHART Strategic Planning website can be found here: https://chart.maryland.gov//readingroom/RR_StrategicPlanning.asp
- A video of the Statewide Operations Center (SOC) Operations Manager for MDOT discussing their AAR reporting procedures using a fatal incident example can be found here: https://vimeo.com/207690734#t=567s
- Maryland DOT Point of Contact: State Operations Center Manager
- RITIS Point of Contact: University of Maryland's CATT Lab Director

Case F

MATOC Regional Operations Evaluation

Following experiences from the 9/11 attacks and other major incidents, transportation officials from Maryland, Virginia, the District of Columbia, and the Washington Metropolitan Area Transit Authority (WMATA) committed to share and coordinate their transportation systems' conditions and information management during regional incidents. They formed the Metropolitan Area Transportation Operations Coordination (MATOC) Program. MATOC is staffed like a small traffic operations center. The staff integrate system technologies, improve procedures and planning, and provide accurate and timely transportation information to the public. This enables participating agencies to work together to make travel smoother and safer. This case describes how the MATOC program has been able to use its shared pool of operational data to support a quantitative analysis of benefits and costs in support of a regional coordination program.

Reporting: Store & Manage, Share Data

Through MATOC, transportation agencies from Maryland, Virginia, and the District of Columbia realized that they could benefit from coordinating not only during significant events, but also to support their day-to-day operational needs. MATOC integrates disparate agency systems and policies to support real-time coordination and information dissemination. MATOC uses the RITIS to provide a common platform for data sharing and analysis. As the program grew, the number of jurisdictions contributing data sets grew as well.

Insight: Analyze & Use Data

Metropolitan Washington Council of Governments (MWCOG) engaged an independent contractor to perform a benefit-cost analysis of the program. As illustrated in Figure 19, the analysis evaluated the response to three representative incidents in the region where MATOC was involved. These incidents were compared to calibrated model incidents where MATOC actions were not present—resulting in estimated costs that were annualized based on historical data.

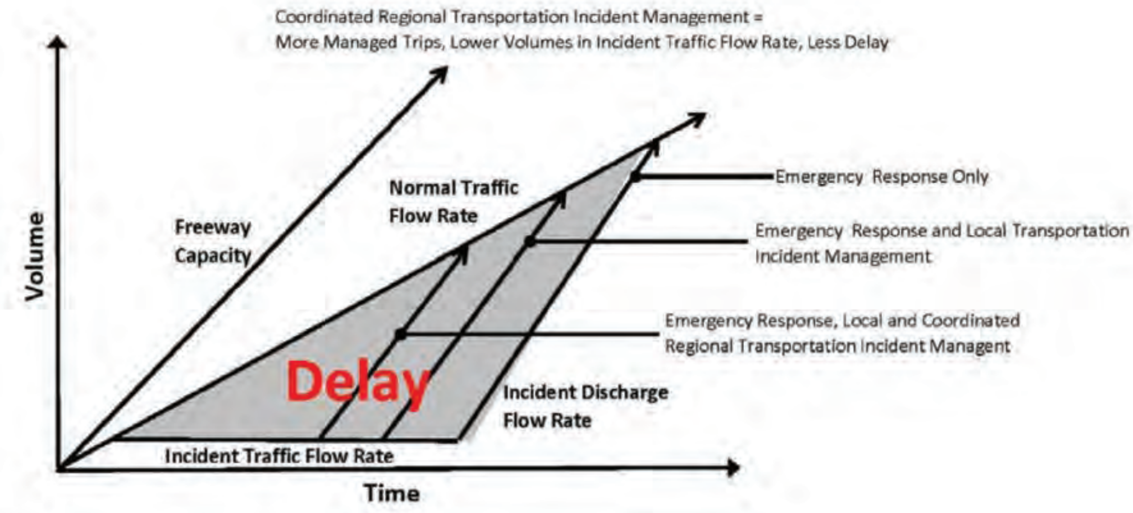


Figure 19. Link-based speed visualization.

The analysis found that an average of 224 police-reported vehicle-related accidents occur per day across the region. MATOC helps to manage approximately ninety regionally significant incidents per month. Annual benefits of direct MATOC action were estimated at $12.9 million in mobility savings, which includes a greenhouse gas savings of more than $500,000.

This is a conservative estimate and does not include the costs of secondary incident reduction.

The positive ROI stemmed primarily from enhanced real-time data sharing among agencies. This allowed agencies to more quickly become aware of incidents, respond, clear the incident, and alert travelers, and to develop standard operating procedures that account for impacts of regional and cross-jurisdictional events. Using the RITIS as a data sharing, warehousing, visualization, and dissemination platform, agencies had easy access to regional performance measures data that included detailed incident and incident response information, as well as flow information from traditional sensors and probe vehicles. A sample RITIS incident information display is shown in Figure 20. The byproducts of this data sharing and collaboration were improvement of each agency's data (since others were relying on it) and the ability to provide better and more relevant traveler information.

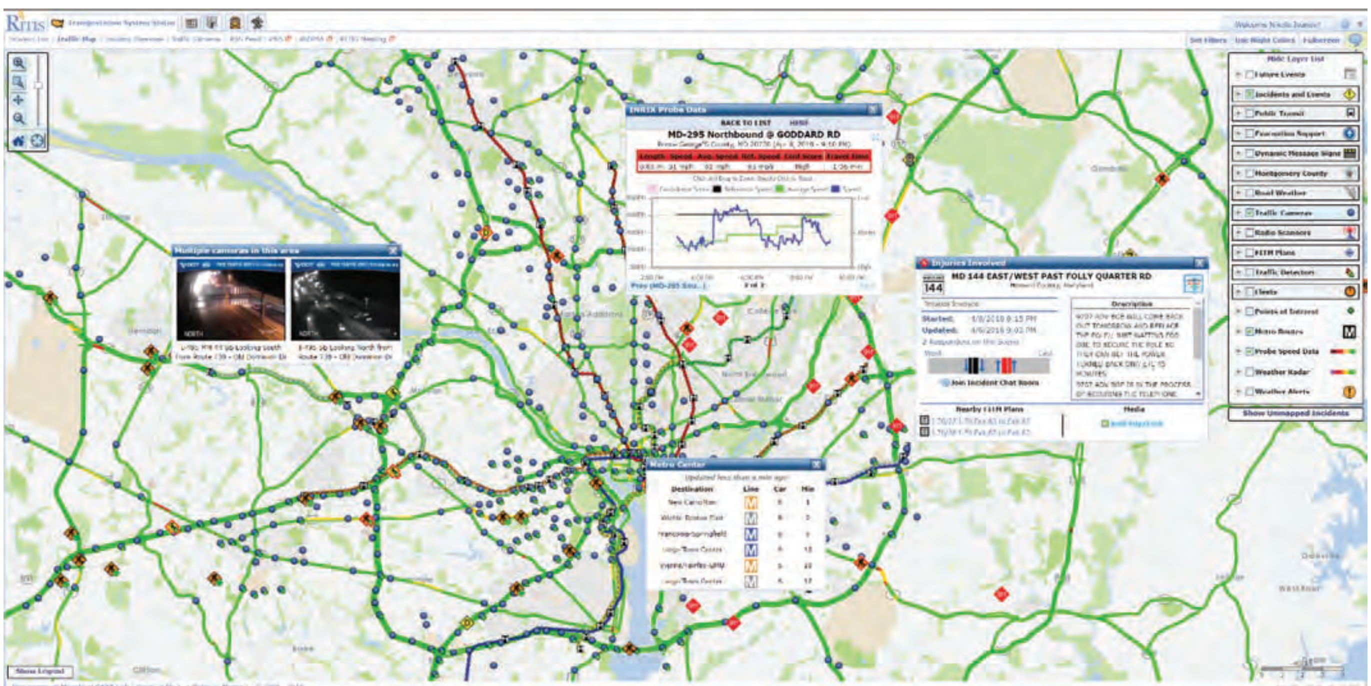


Figure 20. RITIS incident information.

Agencies can now analyze their performance across jurisdictions using a system that removes institutional barriers and data challenges.

The ability of the region to make a case for a program solely focused on data sharing and collaboration removed any doubt about the value of sharing in improving safety and mobility in the region. Presenting a benefit-cost ratio of 10:1 for participating agencies convinced the decision makers to continue to invest in the program and tailor it to the needs of the region as it continues to grow and change.

Success Factors

- **Interagency collaboration.** The pooled operational data from multiple agencies enabled the benefit-cost evaluators to look at benefits and costs as they pertain to the entire region, not just a single agency or jurisdiction.
- **Exposing data as a quality improvement strategy.** As data were exposed to a larger audience, a virtuous cycle of quality improvement and data utilization occurred.
- **Benefit-cost analysis to sustain support.** While 9/11 and other major incidents provided the initial impetus for MATOC, conducting a benefit-cost analysis was instrumental to sustaining support for the program. Historical data enabled before-and-after comparisons of incident response that provided the basis for the analysis.

Challenges & Lessons

Providing a baseline. One of the largest challenges when it comes to evaluating benefits of a program such as MATOC is establishing a baseline performance level. Unlike a capital improvement project that provides a capacity increase, the value of quicker communication or cross-jurisdictional coordination is more difficult to establish. However, the availability of supporting data prior to establishment of MATOC and after the program, and the ability to model incidents, allowed the independent evaluators to calculate tangible benefits of the program.

For more information...

- MATOC Benefit Cost Analysis White Paper http://www1.mwcog.org/uploads/committee-documents/Yl5ZVlZc20100607114406.pdf
- MATOC Website https://matoc.org/
- RITIS https://ritis.org/
- MATOC Point of Contact: MATOC Facilitator

Case G

Creating a Team of Data Experts to Support TPM at the Mid-America Regional Council

Mid-America Regional Council (MARC) is the eight-county, bi-state Kansas City Region's Metropolitan Planning Organization (MPO); it covers four counties in Missouri and four counties in Kansas. Like many MPOs and other transportation agencies, MARC has seen a growing need to gather and analyze data in order to monitor and interpret trends, shape and track progress toward performance goals, and ensure compliance with federal TPM requirements.

In response, MARC has developed new staff capabilities and data management practices that give the agency a better ability to handle data analytics and performance management tasks. These new capabilities come in the form of two new "data developers" who dedicate roughly a quarter of their time to TPM compliance. Prior to their hiring, much of MARC's data-related effort was concentrated on manually obtaining, organizing, and cleaning data.

Foundation: Specify & Define Data

Prioritizing data sets for automation. With the ever-growing importance of data in mind, MARC created a data coordination committee charged with improving agency-wide governance and administration for data management. The committee included liaisons from each department, as shown below in Figure 21.

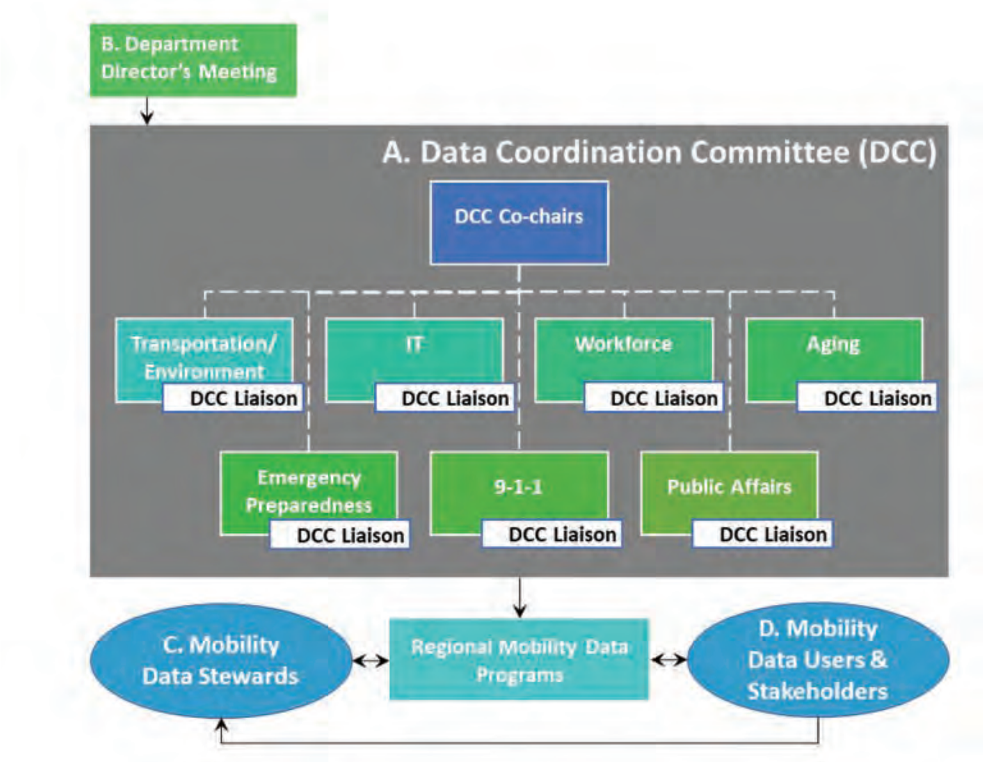


Figure 21. MARC data coordination committee.

After inventorying all data housed within MARC, the committee created a "top ten" list of priority data sets for automation. The list was split evenly between transportation- and census-related items. The transportation-related items covered pavement and bridge conditions, safety/crash statistics, National Performance Management Research Data Set (NPMRDS) information, transit route information, and what MARC calls "network attributes" (e.g., functional classification, National Highway System designation).

Repurposing staff positions to adapt to changing needs. In order to move forward with its plan to automate data processes, MARC repurposed two open positions, including a GIS specialist and a demographer, into data developer positions capable of creating and managing systematic workflows for data gathering and organization . While GIS and demographics support needs still exist at MARC, advances in mapping software and web data tools have reduced the amount of time required for these tasks, and MARC concluded it no longer needed full-time employees for those specific functions.

Data specialists with public-sector experience. In its search for data developer hires, MARC listed programming and application development skills, along with data analytics and visualization, as base competencies. While not listed as requirements, candidates with planning and public-sector experience were given a strong preference. The data developers MARC hired both had public-sector experience, in addition to the required programming skills; one had worked in the GIS division of a public entity; the other had worked for a transportation engineering firm serving public-sector clients. Both had experience in dealing with various stakeholders and in handling data sets related to public policy issues.

Ensuring systems meet the needs of the users. Once hired, the data developers were immediately inserted into the data coordination committee to coordinate their roles with the committee's goals of enhancing MARC's data management processes. To ensure the data developers understand how the data they manage is ultimately used, they regularly participate in meetings with transportation planning staff to discuss high-level data needs. More focused, detail-oriented breakout sessions are conducted as well. These meetings serve as a feedback mechanism to compare the data needs of the MPO with the systems created by the developers to address those needs.

Reporting: Store & Manage Data

Automating data delivery has greatly improved efficiency. Prior to MARC's hiring of the data developers, accessing data to respond to requests or to meet reporting requirements was time-consuming. Data sets were typically downloaded into Excel files that users then had to manually edit in order to extract the relevant information. Analysts were not necessarily aware of what information was available within the agency and would spend considerable time searching through files and folders in MARC's servers to see which data sets had already been saved.

The developers have since created automated processes to obtain data sets and import them into SQL databases. The databases have front-end interfaces that greatly simplify the process of querying them to extract the information MARC needs. This has resulted in MARC being able to dedicate more time toward reviewing and analyzing the data. As information needed to meet federal requirements was given priority, automated processes have been implemented and SQL databases developed for data concerning pavement and bridge conditions, safety measures, and system performance.

Insight: Analyze & Use Data

Beyond standard reporting: diving deep into the data. While automating data processes was the initial impetus behind MARC's hiring of the data developers, a valuable byproduct of the efficiency gained in automation was the ability for MARC to review data in ways that it previously had not realized were possible. Of particular note is how MARC now uses the crash data. As a bi-state MPO, MARC receives crash data from two different state DOTs in two different formats. Prior to automation, MARC only took the high-level data points such as fatalities and injuries from the files as further processing would be too time-consuming. With the use of a commercial GIS data integration platform, the data developers were able to combine the raw crash data files from the separate DOTs into one comprehensive dataset that allows MARC to analyze data points it previously was not even aware were included in the files. This includes location data that could now be presented spatially, adding much depth to their analyses. Also, because the process is automated, MARC is now able to refresh this data quarterly, which leads to more timely analyses.

The time saved in managing data, combined with new tools to better visualize the data, have contributed to MARC's ability to conduct special analyses for planning studies and special projects and, in general, to think of new and more inventive ways to analyze the data they have. Whereas before MARC primarily used the data simply to answer the questions required of it, it can now use the data to come up with its own questions and form hypotheses.

Success Factors

- **Focus.** Developing a top ten list of data elements/sets for automation helped MARC determine the right skill sets to look for, and also helped the developers focus on high priority projects immediately.
- **Communication.** Having the data developers participate in the performance management team meetings ensures that they have a

good understanding of how the data is to be used so that they can develop the proper systems to acquire and organize the data.

- **Automation of routine processes.** The data developers were able to create automated processes for data collection and subsequently developed databases with user-friendly querying capabilities from which to access the data.
- **Deployment of data integration and analysis tools.** The data developers brought new skill sets and introduced innovative tools that vastly improved MARC's ability to access and analyze data. While storage and management greatly improved efficiency, it was the introduction of the data integration platform that truly allowed MARC to see the possibilities for more robust data analyses.

Challenges & Lessons

From "We have two dedicated data developers!" to "We *only* have two dedicated data developers?" MARC has greatly benefitted from the addition of the data developers and the successful systems they have been able to implement. However, the demand for their services is beginning to outpace their capacity. TPM will always take priority given the federal requirements, but MARC now has to focus on balancing the data developers' workloads. That entails better defining their roles and responsibilities.

Data analysis only gets you so far. MARC cited "institutional inertia" as a challenge, specifically the difficulty in convincing stakeholders to appropriately consider data analyses when making critical decisions. Much of the data MARC is now able to process was not available ten years ago. That lack of availability led people to make decisions that were slightly more political in nature. Unfortunately, that practice carries on through today even though the data is now readily accessible and can more easily be analyzed.

A useful committee is often composed of people too busy to sit on it. MARC staff understand the importance of the data coordination committee and its meetings in ensuring the developers have proper guidance, but finding staff bandwidth to keep the meetings going has been a challenge. To address this, MARC is planning on restructuring certain staff roles so that participating in the committee meetings becomes an explicit responsibility.

For more information...

- MARC data webpage http://www.marc.org/Data-Economy
- MARC Point of Contact: MARC Principal Transportation Planner

Case H

New Jersey DOT Project Assessment Reporting

Motivated by a desire to be more open in communicating the reason for selecting projects and describing the impacts of projects, the New Jersey Department of Transportation (NJDOT) began to develop the capabilities to analyze individual projects, quantify impacts, and communicate those impacts in a meaningful, reproducible, and easily understandable way. Prior to 2015, this was extremely difficult due to three challenges:

1. The high cost of conducting an analysis,
2. A general lack of before-and-after data, and
3. The difficulty in distilling complex project and mobility data into a digestible format for the public and decision makers.

NJDOT adopted a Tactical-Level Asset Management Plan (T-LAMP) for better project and program development and to demonstrate to senior leadership, legislators, and the public effective, transparent funding expenditures with positive performance results.

Foundation: Specify & Define Data

The data for this initiative came from INRIX (for speeds and travel times) and NJDOT (for incident, event, and construction location data). All of these data are then provided to the RITIS Platform, where they are fused and integrated. The RITIS Platform is then used for an analysis and for generating the graphics that are inserted into various reporting templates.

The performance reports are largely automated within the RITIS Suite. However, the tools still require basic training and domain expertise. The agency was successful at implementing before-and-after studies because they had a couple of dedicated employees (one transportation engineer and one planner) that took the time to attend training sessions, participated in user groups, and collaborated with other agencies who were also using the tools. They also had the support from senior leadership to utilize the new system.

Reporting: Store & Manage Data

The RITIS Probe Data Analytics Suite includes a series of tools that crunch most of the numbers on behalf of the user. The user, however, must know which analysis needs to be performed and how to select proper parameters in the application and is responsible for interpreting the results and scrutinizing them with additional investigation as needed. The bulk of the work is consolidating all of the results from the analytics and turning them into a story that conveys meaning to the audience.

The use of the analytics tools did not require any advanced degrees, and the amount of training needed was minimal; however, the transportation domain expertise of the planner and transportation engineer made interpretation of the results much easier to digest and insert into a narrative that would be more readily digestible by the public. The agency focused primarily on projects that were ranked as the most impactful. The Garden State Parkway is listed as the second worst congested location, as shown in Figure 22.

Rank	Location	Direction	Average Duration	Average Max Length (miles)	Occurrences	Impact Factor
1	Atlantic City Expy W @ NJ-54/Exit 28	WESTBOUND	1h 30m	5.78	49	25,471
2	Garden State Pkwy N @ Washington Ave/Exit 37	NORTHBOUND	1h 21m	8.54	36	24,892
3	NJ-347 S @ NJ-47/Delsea Dr.	SOUTHBOUND	1h 7m	4.06	79	21,483
4	Garden State Pkwy S @ NJ-52/Exit 30	SOUTHBOUND	1h 49m	5.57	27	16,383
5	NJ-47 N @ US-9	NORTHBOUND	32m	1.34	362	15,552
6	NJ-47 N @ CR-670/E Creek Mill Rd	NORTHBOUND	1h 19m	5.03	39	15,493
7	Garden State Pkwy S @ Shell Bay Ave	SOUTHBOUND	31m	0.79	587	14,317
8	Atlantic City Expy E @ Garden State Pkwy/Exit 38A	EASTBOUND	2h 8m	4.39	23	12,930
9	NJ-47 S @ CR-550 Spur/Paper Mill Rd	SOUTHBOUND	29m	1.59	252	11,588
10	Garden State Pkwy N @ Burlington Toll Barrier	NORTHBOUND	2h 33m	12.36	5	9,457

Figure 22. Top 10 bottleneck locations—state and authority roadways.

Figures 23–25 illustrate some of the other system outputs:

- Figure 23: Congestion Scan—showing impacts (before/after) of a construction project.
- Figure 24: Performance Summary Report—showing impacts (before/after) of a signal removal project.
- Figure 25: System Trend Map—showing impacts (before/after) of a construction project designed to reduce congestion and travel times to the beach.

Figure 23. Congestion scans depict congestion before and after the completion of a project on the Garden State Parkway.

Before (2012) — Performance Summaries for GARDEN STATE PKWY using INRIX data, Southbound January 2012 through May 2012

	Planning time (minutes) 7:00 AM - 10:00 PM	Speed (mph) 7:00 AM - 10:00 PM	Travel time (minutes) 7:00 AM - 10:00 PM
Monday	8.48	59.21	7.79
Tuesday	8.57	58.55	7.88
Wednesday	8.69	58.77	7.85
Thursday	8.63	58.43	7.89
Friday	8.74	58.53	7.88
Saturday	8.76	59.16	7.80
Sunday	8.09	59.96	7.69
Weekends	8.66	59.56	7.74
Weekdays	8.71	58.70	7.86
All Days	8.70	58.94	7.83

After (2016) — Performance Summaries for GARDEN STATE PKWY using INRIX data, Southbound January 2016 through May 2016

	Planning time (minutes) 7:00 AM - 10:00 PM	Speed (mph) 7:00 AM - 10:00 PM	Travel time (minutes) 7:00 AM - 10:00 PM
Monday	8.07	63.56	7.26
Tuesday	7.99	64.64	7.14
Wednesday	8.00	63.39	7.28
Thursday	7.82	64.93	7.10
Friday	8.08	64.61	7.14
Saturday	7.96	64.42	7.16
Sunday	7.77	64.67	7.13
Weekends	7.88	64.55	7.15
Weekdays	8.02	64.21	7.18
All Days	7.95	64.32	7.17

Before (2012)	8.70	58.94	7.83
After (2016)	7.95	64.32	7.17
Change	(0.75)	5.38	(0.66)
Change %	-8.62%	9.13%	-8.43%
	Favorable	Favorable	Favorable

Condition/ Change	Planning Time	Speed	Travel Time
Before	8.70	58.94	7.83
After	7.95	64.32	7.17
Change	-0.75	5.38	-0.66
Change %	-8.62%	9.13%	-8.43%
Result	Favorable	Favorable	Favorable

Figure 24. Before-and-after performance from the Garden State Parkway signal removal project.

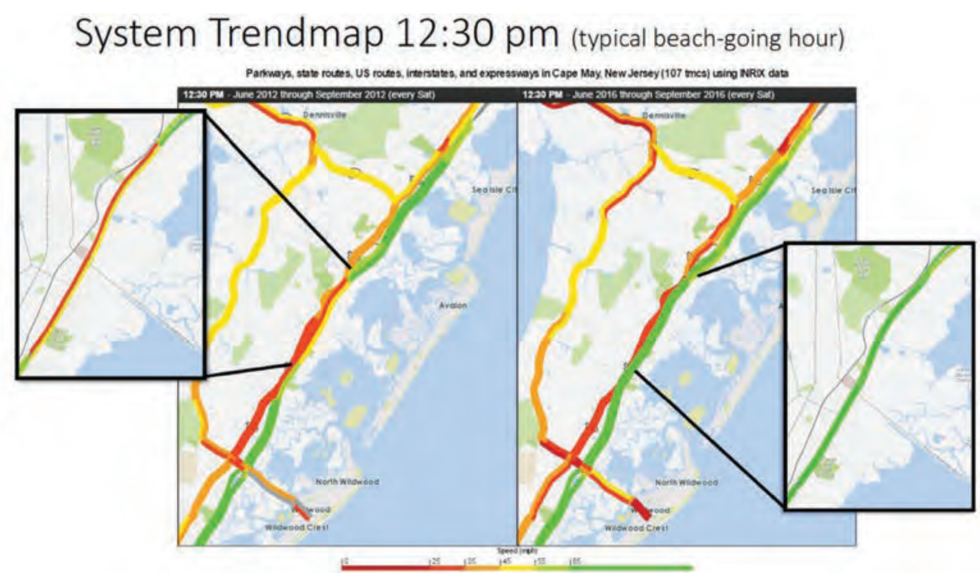
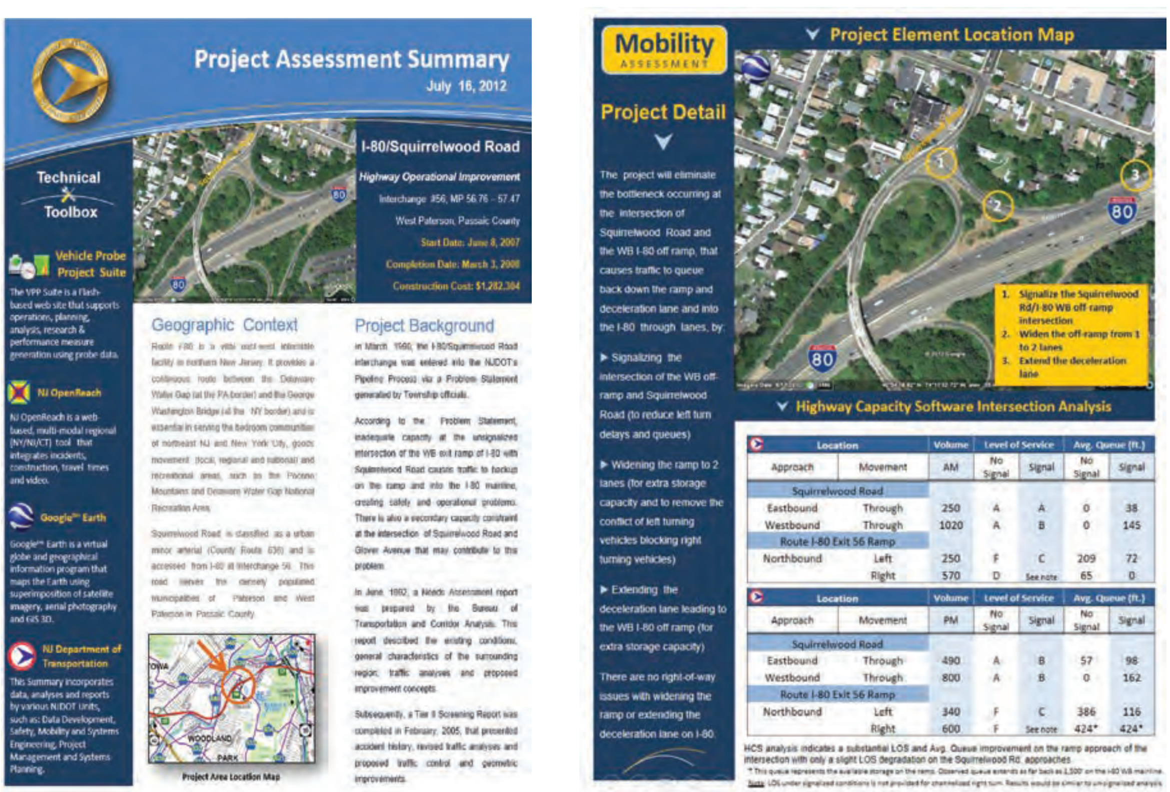


Figure 25. NJ Garden State Parkway—before-and-after conditions of a major construction project designed to reduce congestion and travel times to the beach.

Insight: Analyze & Use, Present & Communicate Data

Simply providing a set of graphs, maps, and charts from the analytics tools is not sufficient. The agency must write a narrative around the results that interprets the results, points out meaningful observations, and tells the "why" within the performance report. This can be a challenge, and the agency was able to leverage in-house journalists, marketing experts, and other professionals who are trained at communicating performance and otherwise complex transportation analysis for a public audience.

For example, both NJDOT and MPOs have developed 11x17 pamphlets that are meant to be handed out to the public and other officials that quickly tell the story of the project (Figures 26 and 27).



Front cover

Page 1, inside

Figure 26. NJDOT Project Assessment Summary Pamphlet—Part 1.

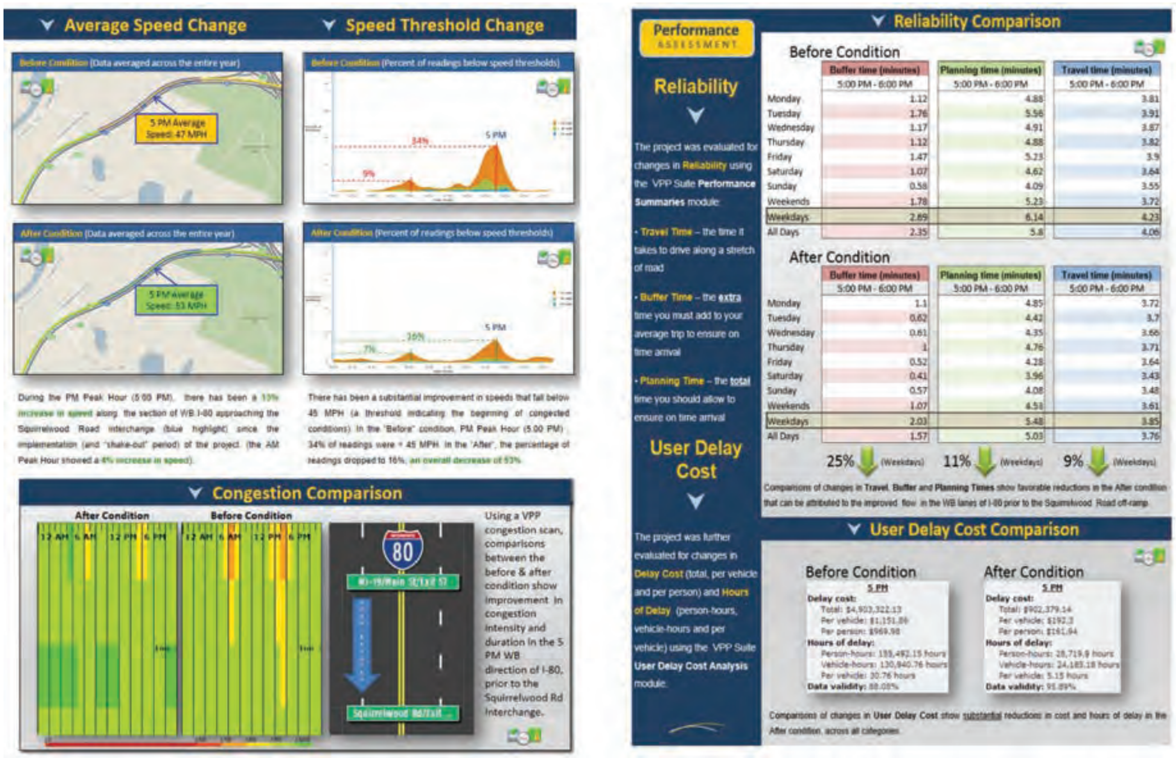


Figure 27. NJDOT Project Assessment Summary Pamphlet—Part 2.

Most of these documents are posted online, and some are also distributed in print form. The new reports and online publications have been well received; the graphics are more engaging than prior reports, and they tell a story that is relatable and understandable to a customer. The agency is now able to update the documents more quickly and have been able to conduct more before-and-after studies in less time than before. This leads to the agency being perceived as more responsive to the public and more capable.

The agency is also realizing significant cost savings in using this approach and the data analytics platform. Before having their data fused and available in analytics, they spent upwards of $20k for a before-and-after study with a small consultant team. Now they can conduct the analysis in-house in just a few hours. NJDOT estimates they are saving $475k/year and 4,475 person-hours on conducting these studies annually.

Success Factors

- Ease of access to the data through graphical user interfaces

- Powerful data visualizations that make data interpretation easy
- A champion within the agency that takes the time to learn to use the data/tools
- Trust in the data and tools (which comes from having a relatively open system that fully documents how data is interpreted and how calculations are made)

Challenges & Lessons

- Making the adjustments needed to transition from prior methodologies that relied heavily on modeling and simulation to a new methodology that is based on actual data
- Building confidence in the new data and methodologies on the part of both technical staff and agency leadership
- Concern that new transparency could potentially show that some projects may not have had the desired impact

NJDOT was able to overcome these challenges because of the open lines of communication from those conducting the work up to senior leadership. It was fairly easy to sell management on switching from a modeling and simulation-based methodology (one that relied heavily on assumptions and piecemeal data collection) over to a methodology that leveraged continuous field measurements (probe data) that had been vetted to be extremely accurate. The I-95 Corridor Coalition had an existing program in place to validate data from the probe data providers, which gave NJDOT leadership more confidence in the data. Lastly, agency team members had been participating in a Probe Data Analytics User Group (also hosted by the I-95 Corridor Coalition) that exposed staff to other agencies conducting similar types of analysis using the Probe Data Analytics Suite. Seeing their peers work on similar endeavors gave the agency more confidence in the direction that they were heading.

For more information...

- Presentation on the NJ DOT Complete Team https://tinyurl.com/completeteam
- Use Case Description https://www.ritis.org/usecases/assessment
- RITIS Platform www.ritis.org
- Project Prioritization Demo https://vimeo.com/179829037

Case I

Ohio DOT Winter Performance Management

Ohio has spent significant time developing and operationalizing their Transportation System Management and Operations (TSMO) Dashboard, which includes 20 performance measures. Because Ohio experiences a fair share of winter weather, some of their performance measures are dedicated specifically to snow and ice recovery. It can be difficult or impossible to keep roadways clear of snow and ice during storms. However, travelers must be able to continue to utilize the roadways even after major weather events. In light of this challenge, Ohio DOT (ODOT) has developed a specific activity and outcome-based performance measure called "recovery time," which allows the agency to evaluate their operational strategies to most effectively manage roadways during the winter season, incentivize recovery efforts, and communicate with both management and travelers.

Foundation: Specify & Define, Obtain Data

In order to effectively manage operations and winter weather response, ODOT collects speed data along impacted corridors prior to, during, and after snowstorms. In the past, ODOT used traditional static sensors to collect speed information, but with emergence of 3rd-party-provided speed data, ODOT has been able to collect speed data statewide at a much more granular level.

The agency also monitors atmospheric and roadway conditions using Roadway Weather Information Stations (RWIS) to identify when the storms begin and end.

A snow or ice event begins when 40% of a county's RWIS stations detect either snow or freezing rain, paired with the following criteria:

- The air temperature or pavement temperature is below 34°, AND
- The speed drops more than 10 mph below its expected value on at least 25% of designated routes within the county (minimum two).

A snow and ice event is considered complete when the following are true:

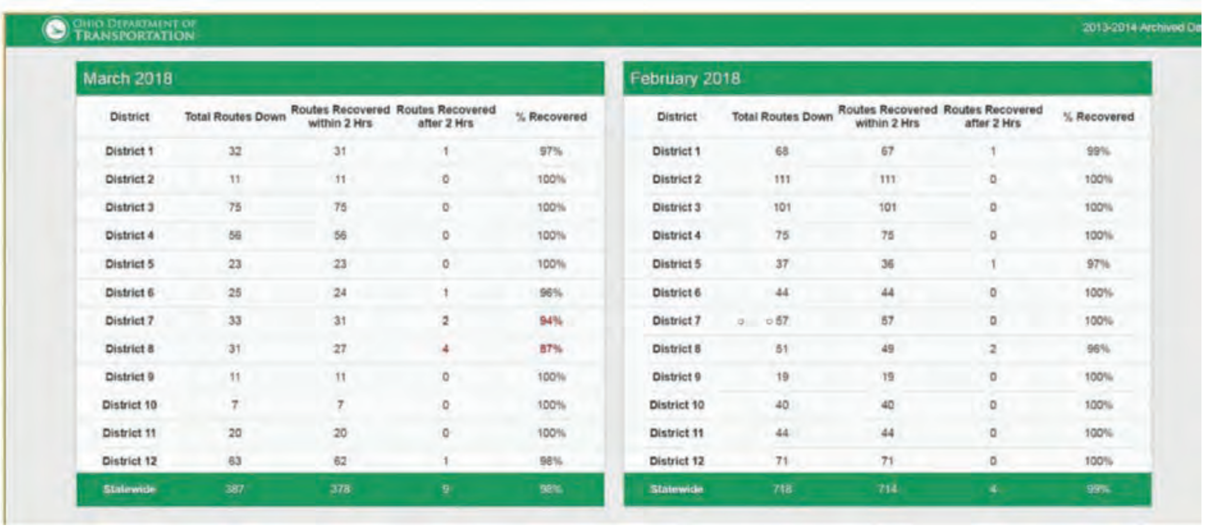
- At least 60% of the county's RWIS stations are reporting "None" or "Rain" as the precipitation type, AND
- The wind speed detected from RWIS stations drops below 15 mph (to account for drifting snow).
- ALSO: A new snow and ice event does not begin within two hours.

When the above is true, the performance clock starts for each route. The time from the end of the event until speeds recover is called the **recovery period**.

Each county in the state has the goal of two hours for the recovery period on each designated route. The recovery period officially ends once speeds recover to within 10 mph of their expected values for at least an hour. Any route that has not recovered within the two-hour goal will be reported.

Insight: Analyze & Use Data

To view speed information and calculate the recovery period, ODOT has several tools at its disposal. First, they have ability to analyze data using traditional tools and techniques, such as Microsoft Excel and database queries. Second, ODOT also has access to a hosted data visualization platform, Probe Data Analytics, which allows the agency to quickly evaluate the conditions prior to, during, and after the storm. The user can select corridors or regions of interest and animate link-based speed information across multiple time periods to determine when the system is operating at pre-storm speeds. Lastly, the agency has developed its own internal performance measures dashboard that leverages data from the above sources to produce a drill-down-capable summary with supporting graphics. Figure 28 shows the platform's snow and ice event dashboard.



March 2018

District	Total Routes Down	Routes Recovered within 2 Hrs	Routes Recovered after 2 Hrs	% Recovered
District 1	32	31	1	97%
District 2	11	11	0	100%
District 3	75	75	0	100%
District 4	56	56	0	100%
District 5	23	23	0	100%
District 6	25	24	1	96%
District 7	33	31	2	94%
District 8	31	27	4	87%
District 9	11	11	0	100%
District 10	7	7	0	100%
District 11	20	20	0	100%
District 12	63	62	1	98%
Statewide	387	378	9	98%

February 2018

District	Total Routes Down	Routes Recovered within 2 Hrs	Routes Recovered after 2 Hrs	% Recovered
District 1	68	67	1	99%
District 2	111	111	0	100%
District 3	101	101	0	100%
District 4	75	75	0	100%
District 5	37	36	1	97%
District 6	44	44	0	100%
District 7	57	57	0	100%
District 8	51	49	2	96%
District 9	19	19	0	100%
District 10	40	40	0	100%
District 11	44	44	0	100%
District 12	71	71	0	100%
Statewide	718	714	4	99%

Figure 28. Partial screenshot of ODOT's Snow and Ice Event Dashboard.

It can be seen in this figure that in April 2018, District 4 met their performance objective of recovering routes within two hours on only 11 of the 13 routes in their district—thus receiving a recovered score of 85%.

As the recovery period became an established measure of performance, the agency was able to implement operational strategies that improved the overall customer satisfaction level. For example, ODOT was able to focus their attention on storm-impacted areas with the slowest recovery period and/or districts that struggled to recover all of their routes to deploy additional roadway treatments for future storms. ODOT also leveraged existing RWIS data to evaluate roadway conditions (like salinity content, surface temperature, etc.) and a combination of CCTV and speed information (from probe data providers and sensors) to further identify potential sources of slow recovery time and dispatch service patrols to assist stranded motorists or direct travelers to alternate routes through traveler information resources and media campaigns.

Their interactive dashboard also allows for drill-down capability—enabling users to look at specific district and specific event performance, as illustrated in Figure 29. This type of drill-down capability made the measures much more insightful as ODOT began to answer the question of why certain districts and routes were performing worse than others.

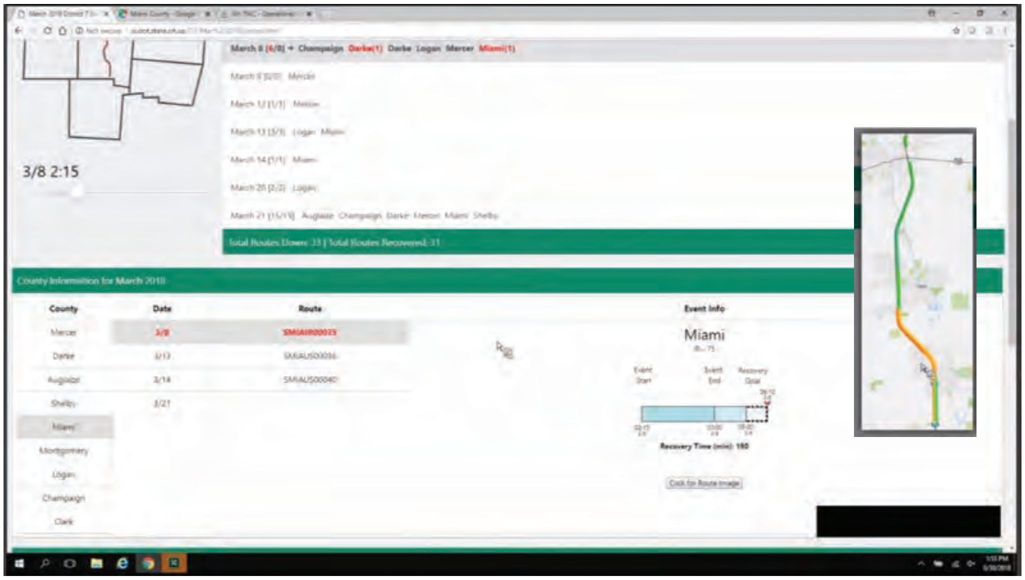


Figure 29. Dashboard drill-down view.

Drilling down in the dashboard lets the user see how each event was managed, whether it hit the two-hour recovery period, missed, etc. Clicking on a specific location draws a diagram of the road depicting which segments did not recover soon enough.

The dashboard resource view (shown in Figure 30) displays which resources were used: overtime hours, brine, equipment, etc. The district can even include written feedback explaining why it believes it was not able to meet the specified target.

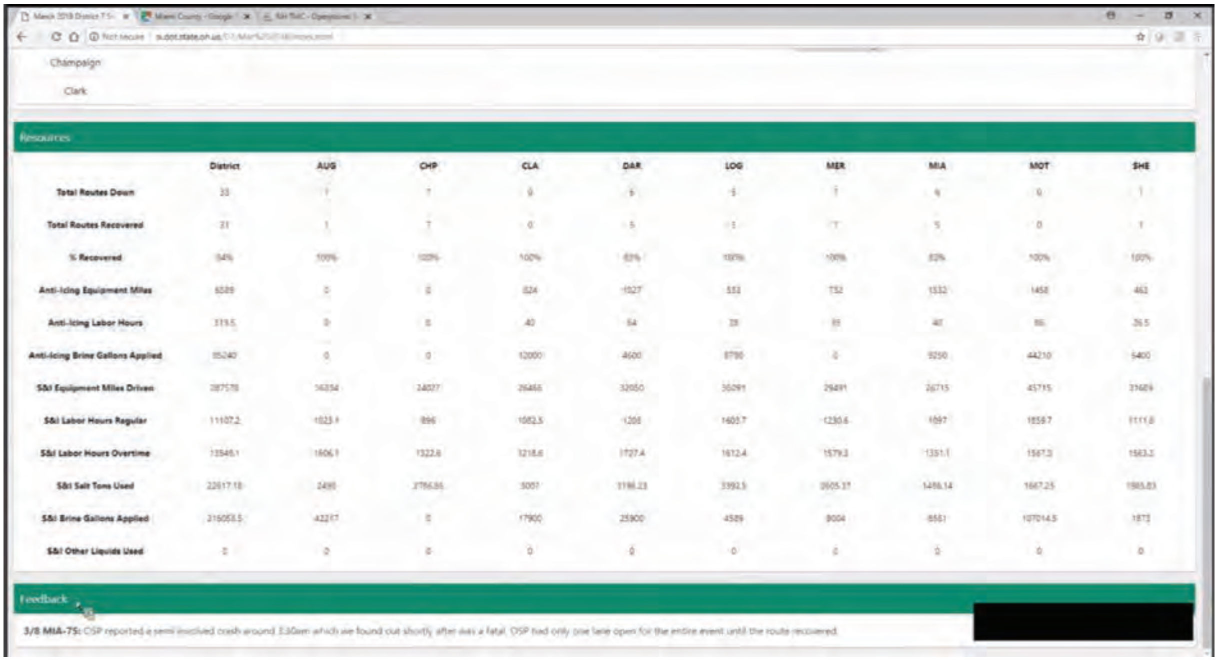


Figure 30. Dashboard resource view.

Success Factors

- **Customer-focused performance measure with connection to operational actions.** ODOT spent a good deal of time developing their TSMO plan and identifying specific measures that would best reflect customer expectations with respect to system performance after a winter storm. Recovery time is the type of measure that is easily communicated to decision makers and general public, yet it also ties well with operational actions that directly influence it.
- **Use of available data and tools.** ODOT was able to utilize existing data sources and tools to calculate this new measure that provided better insight into winter performance management. They also leveraged internal staff to develop their own dashboard and develop computational methods.
- **Workforce capabilities.** ODOT believes it has a large number of younger, computer-savvy engineers who have thoroughly embraced these measures. These staff resources have enabled ODOT to use the measures to make real decisions within the agency.

Challenges & Lessons

Defining scope of measure's effectiveness. As ODOT makes progress in improving recovery periods, they have internally discussed implementation of the same measure *during* the storm. In effect, the goal could be to bring speeds up to pre-storm values while the storm is ongoing. However, many believe this goal may be in direct conflict with safety-related goals that may require that travelers slow down during poor weather conditions. In this context, recovery period may not be the appropriate measure to track *during* a storm.

Automation. The DOT faced a number of technical challenges associated with the scale of this effort. They had to leverage data and technologies that already existed and find a way to blend these data into a system and processes that could compute measures in real time as a storm moves through an area, provide actionable insights, and allow for operational changes to be made in real time. The level of effort involved in designing and implementing such a system was significant. There has also been a significant amount of manual effort to populate the system and keep it functioning—especially in the early years of development.

For more information...

- Webinar Recording on Building TSMO Performance Measures https://youtu.be/SCDnwBDhN5k?t=2345
- FHWA Best Practices for Road Weather Management case study https://ops.fhwa.dot.gov/publications/fhwahop12046/rwm16_michigan1.htm (this case study is of Michigan DOT, which uses a similar measure as ODOT)
- Ohio DOT Point of Contact: Administrator, Office of Traffic Management

Case J

Pennsylvania DOT's Statewide Transportation Operations Data Warehousing Business Plan

The Pennsylvania Department of Transportation (PennDOT) has taken a holistic approach to data planning—involving all possible stakeholders—to try to improve the agency's data capabilities while reducing costs. This foundational work is key to their long-term success.

Overview

Every agency generates transportation-related data and must store those data in a way that enables easy access and management. All too often, agencies collect data in silos. One department generates centerline mapping files in a standalone GIS environment; a second department collects speed and volume data for planning and federal reporting using in-pavement sensors; a third department collects speed data from a mix of probes and above-ground sensors; a fourth department collects and manages toll collection data; and so on.

This approach is often organic and usually happens because agencies are large and complex. However, in other agencies, an ad hoc approach can be intentional. Business units fight for resources, become territorial over their own data, or can be uncomfortable with others becoming aware of their efforts. Ad hoc data management (or lack of management) approaches increase agency costs, limit capabilities, and can lead to a toxic culture.

The more mature agencies take a holistic view of data collection and management—pooling resources to understand data needs, data assets, data gaps, management, and accessibility.

Foundation: Specify & Define Data

In 2015, PennDOT undertook a significant effort to lower the cost of doing business through the consolidation of the state's transportation data assets. Leadership recognized that the transportation community generates a considerable amount of transportation operations data—such as traffic volume data, incident data, asset information, and speed data—without an effective means to share the data between one another.

The concept of a Pennsylvania Statewide Transportation Operations Data Warehouse (PA STODW) was born. The idea behind the STODW was to increase planning, operations, and research capabilities within the state, significantly reducing the cost of doing business, dramatically improving agency capabilities, and improving coordination with partners in transit, tolling, freight, local governments, and emergency management agencies. Implementation of a PA STODW would provide access to the best-available transportation operations data in support of the safe, secure, and efficient movement of people and goods on Pennsylvania roads.

Figure 31 illustrates an early vision for the data warehouse.

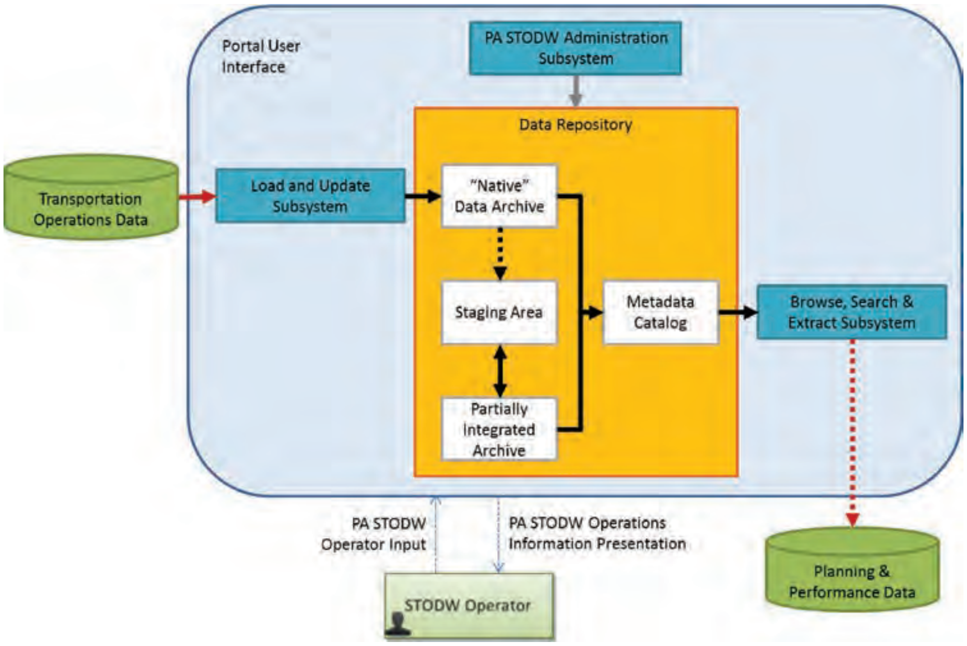


Figure 31. Early PA STODW system concept.

To test the viability of this vision, PennDOT assembled a team to assist in the development of a Concept of Operations (ConOps) and then to develop a business plan with a technical analysis of alternatives and cost analysis. This team, which consisted of leadership within PennDOT's Bureau of Maintenance and Operations and several dozen stakeholders, developed a set of overarching goals for the PA STODW:

1. Support objectives-driven, performance-based transportation operations planning.
2. Support a robust transportation operations performance measures program.
3. Provide integrated and interconnected Enterprise IT and ITS technology and data.
4. Support development of data-driven strategies and actions in the planning process that lead to an integrated, efficient transportation system.
5. Improve cooperation and collaboration within PennDOT and between PennDOT and other public agencies for the sharing of transportation operations data.

Five specific objectives were established in support of these goals, including:

- Establish a PA STODW architecture that is open, receptive and adaptable; is consistent with developing national standards; provides opportunities for private/public partnerships; and encourages and supports interagency cooperation.
- Develop and integrate traffic monitoring, traffic surveillance and incident reporting, roadway and roadside equipment, and environmental information throughout Pennsylvania, as appropriate.
- Define how operations information is collected, processed, archived, shared, and distributed.
- Define the interfaces and information flows among/between subsystems, stakeholder organizations, and PA STODW users.
- Assist in developing, prioritizing, and addressing proposed transportation operations technology and data-related investments.

In the concept development phase, PennDOT assembled a large group of stakeholders that included nearly every office and division within the agency. PennDOT also involved stakeholders from 18 external agencies, including MPOs, municipalities, transit, commercial freight operators/organizations, emergency management agencies, etc. During multiple concept development and outreach meetings, these agencies were asked to describe their agency's data wants and needs and to provide a list of any existing data assets that they might be willing to share. Figure 32 illustrates the deployment concepts that were explored.

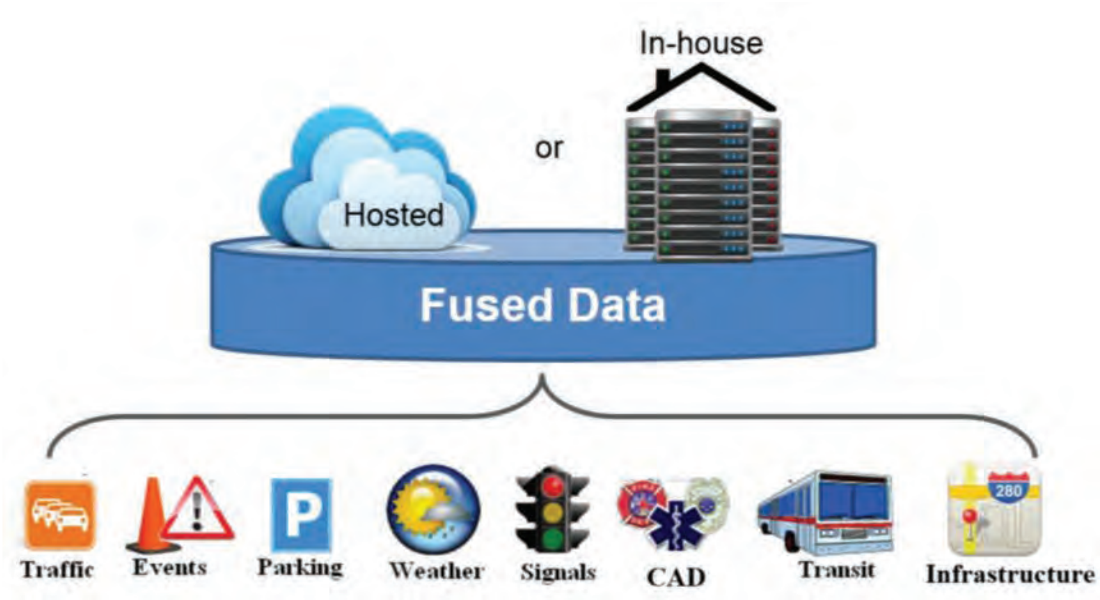


Figure 32. The STODW deployment alternatives concepts.

Success Factors

- **Stakeholder involvement.** Through frequent stakeholder meetings and considerable outreach, this project was much more successful than it otherwise would have been. The stakeholder engagement identified data needs and existing capabilities (and inabilities) with respect to data. These stakeholder meetings served to galvanize the state's data owners and data users and produced a much more effective end result.
- **Leadership buy-in.** The Chief of Traffic Operations in PennDOT had a strong desire to see this project succeed. He was an effective communicator who could easily convey the intent and justification for the project in a way that secured buy-in from others within the agency. Through his leadership, he was able to successfully build a team that shared his vision.
- **Alignment with *existing* agency goals.** PennDOT's State Transportation Advisory Committee's 2015 Transportation Performance report states that "PennDOT is committed to accountability for results and transparency of operations" and that "PennDOT must continue to provide leadership and collaboration to its partners in continuing to modernize transportation products and services." This foundational data project directly aligned with this commitment and therefore was easier to justify to agency funders.

Challenges & Lessons

Conflicting priorities. While nearly everyone in the agency (including many external stakeholders) all viewed this as an important project, it was also seen as a bit of a distraction from other priorities within the department.

Lack of immediate results. The result of this first phase was still very foundational. The deliverable was a plan and a list of recommendations. The implemented system that would truly affect people wouldn't be realized until 1 or 2 years later, and therefore it was a bit difficult to get people to remain excited and committed to the project over a longer time frame after the initial stakeholder engagement activities.

Turf wars. When it came to evaluate alternative implementation strategies, certain internal departments would attempt to convince the ConOps developers to sway the recommendations in their favor—giving them ultimate control over the final system. These internal turf wars had the potential to steer the agency toward a potential solution that might not have been optimal.

For more information...

- Pennsylvania State Transportation Advisory Committee reports & studies web page http://www.talkpatransportation.com/advisory-committees/tac-reports-studies
- PA STODW Concept of Operations Document and Business Plan and Cost Analysis Report No. FHWA-PA-2016-006-130108
- PennDOT Point of Contact: Chief of Transportation Operations

Case K

Virginia DOT’s Pavement Monitoring Program

The Virginia Department of Transportation’s (VDOT’s) pavement management systems and tools were developed to provide network-level needs assessment and maintenance prioritization support, as well as project-level insights to guide detailed project development and delivery. In 2012, VDOT recognized a need to strengthen the links between the decision-support tools and analysis, and field project selection. The objective was to better ensure the alignment of investment decisions with the established statewide performance targets.

This case describes VDOT’s implementation of a pavement performance target monitoring program, which formalized the methodology for establishing district-specific pavement performance targets and the processes for monitoring paving project development and execution toward these district-specific targets. Through routine monitoring of planned paving outcomes in comparison to the established measures, VDOT is able to provide early warning of any anticipated deviations, allowing district pavement managers time to adjust project plans as needed.

Overview

VDOT has a well-established pavement management methodology that includes annual pavement condition collection and needs assessment, establishment of statewide pavement condition targets, and a performance-based budgeting process. VDOT's Central Office Maintenance Division has responsibility for data collection and analysis; districts have primary responsibility for pavement maintenance, rehabilitation, and reconstruction project selection and development.

Foundation: Specify & Define, Obtain Data

In order to implement the performance target monitoring program, VDOT integrated data between its pavement management system (PMS) and a separate application used at the district level to develop paving contracts—the pavement maintenance scheduling system (PMSS). Planned project information from PMSS is periodically transferred into the PMS, where it can be used to evaluate projected outcomes of planned paving against the established performance targets.

Reporting: Store & Manage Data

District paving status reports are compiled by Central Office Pavement Management based on PMS analysis of current pavement condition information, district-specific performance and paving targets, as well as the most up-to-date planned project information available from PMSS.

The status reports provide a comparison of baseline performance and paving targets with planned and/or actual work accomplishments from the two systems. Reporting is completed at key milestones in the development, execution, and delivery of a paving project:

- Routinely during project development,
- Immediately following project advertisement and award, and
- Routinely during the construction season.

The status reports provide aggregated statistics, as discussed below.

Baseline Performance Targets

These condition targets are established according to district and highway system and are summarized from financially constrained PMS optimization analysis results. These targets are established for the percentage of pavement forecasted to be in "Fair" or better condition (termed the "percent sufficient") based on the state pavement condition measure.

Baseline Paving Targets

Paving targets are based on the same analysis results used to establish the baseline *performance* targets, however they identify the total lane miles to be paved within each of VDOT's four treatment categories: Preventative Maintenance (PM), Corrective Maintenance (CM), Restorative Maintenance (RM), and Reconstruction/Major Rehabilitation (RC). It is important to note that each of these treatment categories represents a range of specific pavement maintenance actions, which are selected on a project-specific basis, at any eligible locations within the network.

Planned Paving

This is the number of lane miles scheduled in each of the four VDOT treatment categories, calculated based on detailed planned project information imported from the PMSS. These are presented for comparison to the baseline paving targets, as aggregated from district planned pavement maintenance projects.

Projected Performance of Planned Paving

Based on the district's planned paving locations and treatment selections (as imported from PMSS), these targets are summarized from PMS condition forecasting based on current conditions, network-level deterioration modeling, and the modeled benefits of planned paving. The results are presented to provide the predicted percentage sufficient for comparison to the baseline performance targets.

Flagged Treatment Locations

District pavement treatment selections are flagged if they are substantially different from the unconstrained need identified by the PMS based on detailed pavement condition, pavement structure performance, and traffic information available for the location.

Specifically, a location is flagged if the district's planned treatment differs more than a single treatment category from the unconstrained need (e.g., a PM treatment is planned where unconstrained needs analysis suggests RM is needed). This ensures that the district is making reasonable project-level maintenance decisions at the locations selected for investment, while providing the flexibility to allow project-level adjustment where necessary.

Actual Work Accomplishments

As planned paving is advertised, awarded, and delivered, it is necessary to follow up on actual work accomplishments to ensure these plans are ultimately constructed. Currently, this process requires pen-and-paper data collection in the field, followed by manual update of the PMS. However, this will become a more viable part of the performance monitoring process as tools to support GIS-based field validation of planned paving and automated data transfer to the PMS are implemented.

Insight: Analyze & Use Data

These reports are intended to drive project-level decision making in each of the districts toward an optimal "mix of fixes" based on the district's current pavement condition and pavement maintenance allocations in a way that balances statewide, network-level strategy and decision-support analysis with project-level engineering and decision making.

Metrics surrounding planned projects vs. targets are the emphasis in quarterly reporting meetings, as this is where VDOT pavement management staff have the most control over possible outcomes.

By providing routine comparison of expected outcomes of planned work to district-specific paving and performance targets, VDOT pavement managers are provided the information necessary to ensure the district's project-level decision making is in alignment with network-level pavement management strategy and goals. This is accomplished in a manner that is objective and transparent and allows for iterative improvement of both district decision-making processes and Central Office network-level decision-support tools.

Figure 33 shows a sample district status report. This report highlights

- Current, targeted, and predicted paving performance;
- Predicted performance and planned paving as compared to baseline targets; and
- Expected outcomes of the district's current plans.

This information is all included as part of the standardized reporting products, which are discussed within regular paving status meetings.

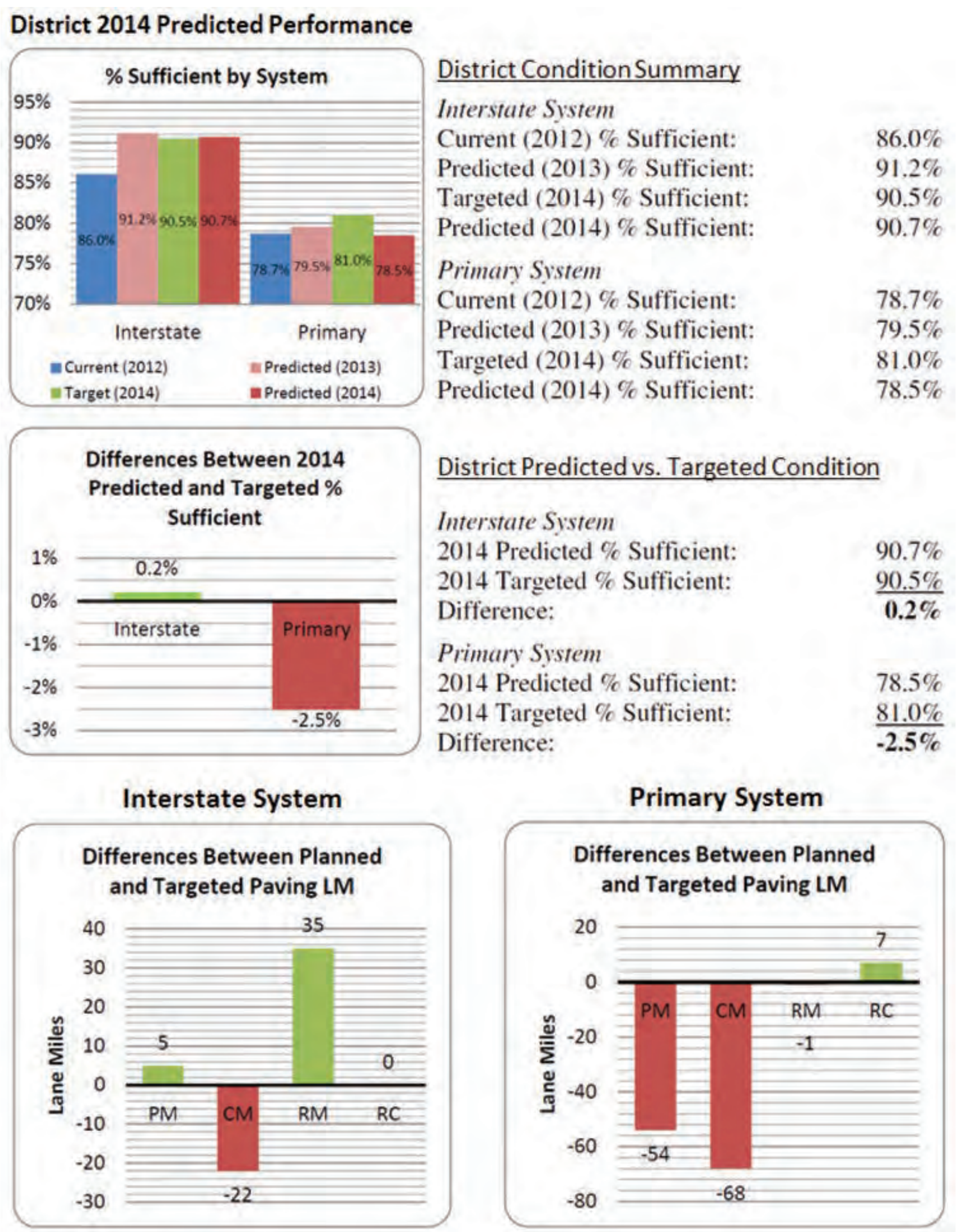


Figure 33. Sample pavement monitoring report.

A district pavement manager is expected to use the reported information to address discrepancies between district-specific project plans and network-level goals for their program. In this example, the district manager would quickly identify that current plans achieve the performance goals for the Interstate system, however, the primary system will underperform if current plans are not improved.

By reviewing the differences between planned and targeted paving by treatment category, the pavement manager should see that a reasonable explanation for the underperformance of the primary system is that the current maintenance program is overemphasizing RC at the expense of low cost, effective PM and CM treatments. By acting to reduce planned RC to allow funding of more cost-effective maintenance strategies, the district should be able to bring the predicted performance of the primary system closer to the desired performance target.

Additionally, a tabular report is provided that flags any locations where the district-selected maintenance treatment does not reasonably align with the PMS unconstrained treatment recommendations. In this case, the district pavement manager could review this report to identify locations where a costlier treatment has been planned over the PMS-recommended maintenance treatment. This may allow for efficiencies to be identified, generating savings that could be used for additional maintenance on the primary system.

Success Factors

- **Planning and data modeling to facilitate system integration.** Through iterative improvement of both PMS and PMSS data models over the course of several years, automated processes to transfer information between these two systems have been developed.
- **Tapping into available data sources.** Optimal selection, planning, and execution of maintenance projects are the most direct ways for a DOT to improve asset conditions on their network. VDOT identified the information available during project development and execution that could be used to predict the influence of planned paving on the network-level performance and made this information available to PMS analytical tools. This allowed the department to integrate network analysis with project decision making without additional data collection or reporting burden to district staff.
- **Linking paving schedules to performance targets.** Arming district pavement managers with the information needed to understand network-level implications of project-level investment decisions reinforces good pavement management practice. Field input generated through elevated attention to the PMS analysis also exposed previously unrecognized opportunities to improve the decision-support tools.

- **Building report review into business processes.** Pavement target monitoring analysis is integrated into the project development process through an easily consumed reporting format and in a way that is respectful of field decision makers' local experience and expertise. The information in these reports is the focus of quarterly performance meetings, as well as routine status meetings with Central Office Pavement Management program leaders, which ensures district attention and action.

Challenges & Lessons

VDOT faced challenges bridging network-level analysis to project-level decisions in a manner that would support routine execution as part of the pavement target monitoring process.

Bridging PMS analysis to project decision making. VDOT worked to balance the use of the network-level analysis to ensure that decision quality would not be reduced through overly prescriptive use of network-level analysis and that distrust of modeling output would not develop where network analysis did not align with the project-level observations.

Automation of data exchange. An efficient process to translate planned paving information from PMSS to the PMS was required. Developing a solution required significant effort on the part of IT and business staff to review and update existing data models, as well as address location referencing issues between the two systems. Proactive engagement of IT and/or Enterprise Architecture staff to help identify formal, sustainable solutions to the exchange of information between established IT systems is recommended.

For more information...

- VDOT Point of Contact: Maintenance Division Assistant Division Administrator, State Infrastructure

Appendix A: Capabilities Checklists

Step 1: Specify & Define Data

Determine what types of data are needed, how data will be used within TPM business processes, and, based on this, specify attributes, scope, level of spatial granularity, and frequency of updates.

Basic

- ☐ The business need for data has been identified and documentation of this need is available for future reference.
- ☐ An inventory of existing agency data sources has been compiled.
- ☐ Managers of the units responsible for data collection can describe the primary users and uses of that data.
- ☐ Data requirements to meet internal and external performance reporting requirements are defined and documented—including attributes, scope, and granularity.
- ☐ Location referencing methods for performance data are established to enable linkages with other agency data sets.
- ☐ Updated frequencies for new data are defined and documented.
- ☐ Authoritative data sources have been designated for performance measure calculations.

Advancing

- ☐ Discussions about data requirements are not constrained by the status quo—they reflect what is important to know about transportation performance in order to improve.
- ☐ Data needs are identified to support the entire TPM cycle (beyond performance reporting), including root cause analysis, identification and prioritization of improvements, and evaluation of impacts.
- ☐ Minimum data quality standards are established considering timeliness, accuracy, completeness, consistency, and accessibility.
- ☐ Data requirements are defined collaboratively across business units, including GIS and IT.
- ☐ Data communities of interest (or equivalent) have been established to identify data improvements to support different business needs.

Gaps & Improvement Ideas

Step 2: Obtain Data

Acquire the data needed to support the entire TPM process including data needed to calculate performance measures, understand trends, set realistic targets, and improve performance.

Basic

- ☐ Data collection procedures and protocols are defined and documented.
- ☐ Data collection and processing workflows are mapped to clearly assigned responsibilities and deadlines.
- ☐ Existing agency data sources are reviewed prior to collection of new data.
- ☐ Available external (public and private) data sources are reviewed prior to collection of new data.
- ☐ Quality management procedures are defined and documented, including training and certification for data collection personnel.
- ☐ Requirements are in place that ensure new data collection adheres to agency location referencing standards.
- ☐ Impacts of changes to existing data collection methods are assessed to minimize loss of consistent trend data and disruption to existing reports.
- ☐ Data sources are assessed to understand usage restrictions that may limit value.

Advancing

- ☐ The full cost of new data acquisition is estimated—considering initial collection, ongoing updates, and supporting staff and technology infrastructure.
- ☐ Funding for regular data updates (beyond the initial collection) is planned and committed.
- ☐ There is regular communication with partner agencies to identify opportunities for collaboration on data collection.
- ☐ Periodic scans are conducted to identify ways to improve data quality and collection efficiency.
- ☐ Agency guidance and/or coordination protocols have been established to assist business units wishing to purchase commercial data sources.
- ☐ Specialists with appropriate expertise (in-house or contractors) evaluate use of emerging private data sources.
- ☐ Data requirements are defined with consideration of opportunities to create valuable information through integration of multiple data sources.

Gaps & Improvement Ideas

Step 3: Store & Manage Data

Determine where and how to store the data, how much data to keep, how data can be integrated across repositories, and which best practices should be implemented for QA and documentation.

Basic

- ☐ Data needed for TPM is stored in databases that are managed and regularly backed-up to provide protection from unauthorized access and corruption.
- ☐ Back-ups are tested on a regular, established cycle (e.g., monthly).
- ☐ Quality control procedures are in place to flag records that do not meet established validation criteria.
- ☐ Data dictionary information (metadata) is maintained and stored in a standardized fashion.
- ☐ Annual data snapshots are created for coordinated reporting across data programs.

Advancing

- ☐ Hardware and software requirements for data storage, updating, integration, and access are understood.
- ☐ Central data repositories have been established to integrate data from multiple sources and provide source data for reporting and analysis.
- ☐ Cloud and hosted storage options are considered for larger and more complex data sets.
- ☐ Data retention policies and archiving protocols have been updated to reflect lower storage costs and analysis of TPM business data needs.
- ☐ A range of data storage options are available to support databases with high transaction volumes and memory-intensive calculations as well as archived data retained for future use.
- ☐ Standards have been adopted to enable combining data from different sources.
- ☐ Data from multiple sources are fused to assemble a more complete and accurate data set than would be possible from any single source.
- ☐ Where appropriate, edge computing techniques are used, involving data processing at the source (e.g., at the site of the field sensor) rather than within a centralized repository.

Gaps & Improvement Ideas

Step 4: Share Data

Share transportation performance data across business units within an agency, across agencies, or with the general public.

Basic

- ☐ Employees are aware of key performance data sources within the agency.
- ☐ There are clear agency policies in place that data should be shared unless the need to protect it is demonstrated.
- ☐ There are protocols defined for how to share data to meet different needs that consider use of state and federal open data portals and hosted or cloud solutions.
- ☐ Open data portals are used to share data.
- ☐ Data explanations are provided in "plain English" to help users understand meaning, sources, and limitations.

Advancing

- ☐ Data governance and stewardship structures have been established to facilitate communication about data sharing and identify opportunities for synergies across business units for collaborating or combining data sources.
- ☐ Data sharing agreements are used (internal to an agency and between an agency and its partners) that specify what data will be shared, when and how—and establish a clear understanding of data limitations and expectations for use.
- ☐ Data are shared in formats that are designed to meet the needs of different users, which may include standard reports, data feeds, and dashboards.
- ☐ Data with sensitive elements are sanitized for public distribution.
- ☐ Data contracts and sharing agreements are reviewed to ensure that agency flexibility is retained.

Gaps & Improvement Ideas

Step 5: Analyze & Use Data

Use analysis tools to convert data into information, identify intended uses and users for these tools, designate one or more individuals to develop specialized experience with these tools, and use these tools to support decision making.

Basic

- ☐ Analysts are aware of and taking advantage of existing commercial off-the-shelf, open source, and publicly available tools for analysis, visualization, forecasting, and scenario analysis.
- ☐ Analysts are trained in use of data analysis and visualization tools.
- ☐ Private-sector or university contractors are used to provide data analysis services as alternatives to standing up analysis capabilities in-house.
- ☐ Data are available that are sufficiently accurate to meet analysis requirements.
- ☐ Visualization and analysis tools are used to explore and discover data anomalies and limitations.
- ☐ Data preparation and analysis tasks are well defined and planned to ensure sufficient calendar time and staff resources.
- ☐ Analysts are able to identify trends and causal factors.
- ☐ Data element meanings, data transformations, and analysis assumptions are documented.

Advancing

- ☐ Predictive models for key transportation performance measures are validated based on multiple cycles of application.
- ☐ Targets are established based on predictive analysis relating revenues and programmed work to performance results.
- ☐ Data mining is conducted to support "back-casting"—which involves starting with a future vision and analyzing current and historical data to estimate changes required to move from the current situation to the future vision.
- ☐ Cooperative arrangements across agencies have been established to transform data into information (e.g., the state DOT performs analysis of travel-time reliability, computes measures for each facility, and provides the data for use by MPOs and local agencies).
- ☐ Predictive analytics and machine learning techniques are applied for predicting asset failure probabilities and other performance measures.

Gaps & Improvement Ideas

Step 6: Present & Communicate Data

Interpret performance results, develop effective ways of communicating narratives based on the data, create visualizations in a variety of formats, and identify needs for data improvement or augmentation.

Basic

- ☐ Managers and analysts meet to review and interpret performance results.
- ☐ Story lines for performance results are developed, reviewed, and communicated.
- ☐ Training is offered to internal staff to build skills in data presentation and communication.
- ☐ Staff have capabilities to present data in a variety of formats tailored to the needs of different audiences, including heat maps, thematic maps, timelines, and other infographics.
- ☐ A combination of narrative and graphical presentation is used to communicate performance information.

Advancing

- ☐ Feedback from data consumers is sought and used to improve communication of information to different target audiences.
- ☐ Individuals with expertise in data visualization and communication are available to support development of performance data products.
- ☐ Social media is used to communicate key results or draw people to more detailed communication products.
- ☐ Specialized visualization and analysis environments have been developed—e.g., virtual reality simulators.

<u>Gaps & Improvement Ideas</u>

Abbreviations and acronyms used without definitions in TRB publications:

A4A	Airlines for America
AAAE	American Association of Airport Executives
AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACI–NA	Airports Council International–North America
ACRP	Airport Cooperative Research Program
ADA	Americans with Disabilities Act
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
DHS	Department of Homeland Security
DOE	Department of Energy
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FAST	Fixing America's Surface Transportation Act (2015)
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
HMCRP	Hazardous Materials Cooperative Research Program
IEEE	Institute of Electrical and Electronics Engineers
ISTEA	Intermodal Surface Transportation Efficiency Act of 1991
ITE	Institute of Transportation Engineers
MAP-21	Moving Ahead for Progress in the 21st Century Act (2012)
NASA	National Aeronautics and Space Administration
NASAO	National Association of State Aviation Officials
NCFRP	National Cooperative Freight Research Program
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
PHMSA	Pipeline and Hazardous Materials Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (2005)
TCRP	Transit Cooperative Research Program
TDC	Transit Development Corporation
TEA-21	Transportation Equity Act for the 21st Century (1998)
TRB	Transportation Research Board
TSA	Transportation Security Administration
U.S. DOT	United States Department of Transportation

TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001

ADDRESS SERVICE REQUESTED

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people's lives worldwide.
www.national-academies.org

ISBN 978-0-309-48074-1
90000
9 780309 480741