Final Report

to the

National Cooperative Highway Research Program

(NCHRP)

on Project 14-25

Guide for Selecting Level-of-Service Targets for Maintaining and Operating Highway Assets

University of Wisconsin – Madison

2013

Acknowledgment of Sponsorship

This work was sponsored by the American Association of State Highway and Transportation Officials, in cooperation with the Federal Highway Administration, and was conducted in the National Cooperation Highway Research Program, which is administered by the Transportation Research Board of the National Academies. Dr. Amir Hanna, Senior Program Officer, managed the project.

Disclaimer

Author Acknowledgements

The research reported herein was performed under NCHRP 14-25 by the University of Wisconsin-Madison, the National Center for Pavement Preservation, and Venner Consulting, with the National Cooperative Highway Research Program (NCHRP) serving as Fiscal Administrator. Teresa Adams was Principal Investigator and Ernie Wittwer was co-PI. John O'Doherty, former Maintenance Director of Michigan DOT and currently at NCPP was involved in and assisted with all phases of this project. Marie Venner also contributed to this report.

Contents

Figures

Tables

Abstract

Large organizations such as Departments of Transportation (DOTs) and other road agencies employ thousands of workers and spend millions of dollars to accomplish their missions and achieve their strategic long-term goals. To do this, they undertake hundreds of activities daily across a broad range of disciplines such as administration, accounting, engineering, operations, emergency response, etc. In the past, it has not always been clear how detailed operations activities contributed to a department's strategic goals. For example, could a change or adjustment to the mix and quantities of maintenance activities produce a more thorough accomplishment of the strategic goals, perhaps even at a lower cost?

DOTs utilize high-level policy goals and objectives, performance measures, and analysis of options and tradeoffs for investment strategies to operate, maintain, upgrade, and expand physical assets effectively throughout their lifecycle. A key component of such a strategy is the identification of important maintenance features that make up highway infrastructure and the establishment of a level-of-service (LOS) target for each feature. Such targets have often been established based on expert opinion of individuals and not necessarily following a systematic process. While most DOTs have detailed information on inventories and costs, they do not have ways of aggregating, prioritizing, and making trade-offs among different maintenance activities and levels of service. For many DOTs, the challenge is linking maintenance activities to priorities and using these priorities to set LOS and budgets. This project offers such a quantitative way to develop the prioritization of the maintenance activities according to the policy objectives, included in the attached *Guide for Selecting Level-of-Service Targets for Maintaining and Operating Highway Assets*. A series of steps illustrate a method of streamlining maintenance activities and optimizing their contribution to the accomplishment of an agency's strategic goals.

1. Project Summary

Highway agencies are being constantly challenged to meet the needs of maintaining highway facilities amid growing customer expectations and increasing traffic, with stagnant or decreasing budgets. To meet these challenges many agencies have adopted some form of maintenance performance measurement and management. These tools can assist the maintenance manager by providing the focus of goals. They can allow a better understanding of the condition of the system and the trends in that condition. They can assist in improving the efficiency and effectiveness of maintenance programs and operations. Moreover, they are powerful tools for communicating conditions, directions, constraints, and challenges to staff and policymakers inside the agency and to policymakers, users, and taxpayers outside of the agency.

To be useful in all of the ways outlined, measures and targets must be thoughtfully selected and used. Measures and targets must also reflect the overall direction and priorities of the agency and fit within an overall management framework. To do this they must be structured to respect and complement the culture of the agency. While much has been written about selecting and defining measures, no accepted method exists for establishing targets for those measures. Indeed, even the vocabulary surrounding targets tends to cause some confusion.

The *Guide for Selecting Level-of-Service Targets for Maintaining and Operating Highway Assets*, which is the primary product of this project, draws on a rich literature on performance measurement, the experience of a large number of agencies, and the knowledge and skills of the project team to outline methods for selecting targets that meet the needs outlined above.

1.1. Research Objective, Scope, and Summary

The objective of this research was to develop a *Guide for Selecting Level-of-Service Targets for Maintaining and Operating Highway Assets* (hereafter, the Guide). The Guide contains processes for:

- Assessing the culture of the agency and the resources available within it.
- Selecting the framework, or structure, of measures that best meets the needs of the agency.
- Dealing with data.
- Gaining support.
- Communicating targets and results.
- Defining priorities within maintenance.
- Estimating costs and contributions for alternative actions.
- Optimizing outcomes.
- Analyzing risks.
- Selecting both desirable and attainable targets.
- Linking targets to budgets.
- Managing with targets.

The project was completed in two phases. Phase 1 was largely devoted to the collection of materials and to the analysis of those materials.

- Task 1a was a literature review in which a number of questions were focused upon:

    - How is value, or priority, defined?
    - How are measures connected to strategic, or agency, goals and direction?
    - How are measures used?
    - What is measured?
    - How are targets defined?
    - How are targets and available resources being related?
    - How is risk considered?
    - Are there any processes for target setting?
- Task 1b was a survey of the states and provinces.
- Task 2a was a synthesis of commonly used measures.
- Task 2b was an in-depth series of interviews with 18 selected states that focused on the same questions listed under 1a.
- Task 2c was an evaluation of the processes used for establishing targets.
- Task 3 was a review of the methods of establishing LOS targets.
- Tasks 4 and 5 dealt with developing plans for Phase 2.

Phase 2 was comprised of only two tasks:

- Task 7 was developing the Guide and outreach materials.
- Task 8 was writing the final report.

Accomplishing Task 7, the preparation of the Guide, involved further in depth interviews with states to gather information not only on their processes, but also basic data: inventories, condition assessments, goals, etc., so that the methodologies under development could be applied and tested.

It also involved a review of the issues that we discovered in phase one to develop tools and methods that would address these issues and be adaptable to the many conditions that exist within the transportation field. Finally, it involved writing and honing the Guide to make it understandable and useful.

1.2. Summary of Findings

Through Phase 1, a number of conclusions were drawn. The first is that agencies involved in performance management conceptualize their management processes in many different ways, reflecting the different goals in adopting the practice and the different cultures and constraints of each agency. What is measured and how standards and LOS targets are defined also varies among agencies.

These differences highlighted the need for a flexible approach to structuring a method of target setting and a method that could be used in a wide range of agencies, reflecting their maturity, sophistication, culture, and resources.

In risk management, those agencies that pursued a more formal approach all seemed to follow roughly the same approach of considering the probability of a risk happening and the consequences of it happening.

From the survey of states and provinces, 46 of 47 reported that they used measures in some of their management activities. Of those using measures, 50 percent said they highly used them in maintenance management or maintenance quality assurance.

Measures are commonly used for all of the major elements of maintenance and operations. Measures are used to a greater or lesser degree for several purposes among the 18 states interviewed in depth. Most say they use them to report to agency policymakers and to improve the effectiveness of the maintenance program. A smaller number reported that they were used to ensure that goals were met and to report to external policy makers.

Among the 18 states, historic trends were most commonly used to define targets. Other methods were used to a lesser extent.

The points noted above are but a very small sampling of the information gathered and analyzed over the course of the study. All of this information led to a number of major conclusions that informed the structure of the Guide:

- Given the great range of agencies, their structure, level of sophistication, and resources available to them, the tools and procedures recommended in the Guide must be flexible. They must be adaptable to the range of conditions that exist among agencies.
- Information is a resource that varies most among the states. Therefore, the Guide must offer ideas on how data can be estimated and otherwise manipulated to compensate for data shortcomings.
- Few agencies have made any systematic effort to prioritize the many activities that comprise maintenance. Without such priorities, setting targets is very difficult. Therefore, the Guide must contain tools and processes to established and agreed upon priorities.
- In many agencies maintenance is treated as an expense rather than as an investment. In part this may be because maintenance goals and actions are not tied to agencies strategic actions and the great number of maintenance activities make discussions with policymakers difficult. To meet these problems, the Guide contains ways of rolling up activities into categories that will be more meaningful to policymakers and ways of directly tying maintenance goals to agency strategic goals.
- The actual use of measures to guide and manage maintenance programs is surprisingly rare. While many agencies have measures, few use them to allocate resources, few have integrated them into their budgeting processes, and few conduct regular monitoring of accomplishments. The Guide recommends processes for moving measures into the world of management.
- Risk management is something that maintenance managers do informally and instinctively. Formal procedures to deal with risk are rare. The Guide recognizes this reality and recommends basic, commonsense approaches to dealing with risk.

To meet all of the needs identified, the Guide follows a step-by-step process. The Guide deals with the entire process of managing with measures. It begins with the basics of how standards, measures and thresholds should be selected. It moves the user through a process for assessing the agencies to determine what might work best and to identify obstacles that might be encountered.

It then moves on to outline some of the choices that are available in terms of structuring measures and reports of measures and recommends how decisions on those structures can be made. Extensive coverage is given to gathering, manipulating, and using data. Gaining support within the agency for a performance management program is discussed, and ideas for communicating internally and externally are discussed.

Paired comparison tools are introduced and illustrated to help the user understand how priorities can be defined and how maintenance goals and actions can be tied to agency strategic goals. The use of these tools is then expanded to illustrate how the relative costs and benefits of alternative strategies and actions can be calculated and compared. Linear programming techniques are then offered as a method of estimating the optimal use of limited resources. The risk associated with maintenance actions is considered and processes recommended.

The actual target setting process flows from the paired comparison and optimization process with recommendations on how both desirable, or unconstrained, targets can be set and on how targets can be related to available resources.

Finally, the management process is overviewed with ideas on how targets can be used to inform budgets and how plans can be monitored, refined, and recalibrated.

Data, processes, and experience from four states are used to illustrate how the recommended procedures can be used. The states were selected because they represent a wide range of conditions. Application of the techniques was not smooth in all cases. Effort will be needed on the part of agency maintenance managers to adapt the ideas and some internal procedures to make it work, but the effort should provide benefit.

The Guide may be more ambitious than the research statement would lead one to believe, but the needs revealed in the study tended to expand beyond simply how targets might be defined to how they can be used and how maintenance management processes can be implemented to improve maintenance programs. If the Guide is read in this broader context, it will be of great help to many agencies.

2. Introduction

Large organizations such as DOTs and other road agencies employ thousands of workers and spend millions of dollars to accomplish their missions and achieve their strategic long-term goals. To do this, they undertake hundreds of activities daily across a broad range of disciplines such as administration, accounting, engineering, operations, emergency response, etc. In the past, it has not always been clear how detailed operations activities contributed to a department's strategic goals. For example, could a change or adjustment to the mix and quantities of maintenance activities produce a more thorough accomplishment of the strategic goals, perhaps even at a lower cost?

Increasingly performance measures are used to guide and improve transportation maintenance programs, to inform maintenance budgets, and to communicate the condition of the highway system to policymakers inside and outside of the agency. While a number of studies have been done of processes involved in maintenance performance management, none have looked at the central process of defining targets.

2.1. DOT Motivations for Using Targets in Maintenance Performance Management

Performance management links organization goals and objectives to resources and results. Measures and targets point to the agency's progress toward achievement of those goals and the effectiveness of its investment decisions. Targets "provide transparency and clarity to the resource allocation decision-making process" and when linked to measures and goals, provide "perspective for evaluating the impact of the investment decision *in relation to* the desired end-state, i.e., how significant is a particular investment in helping an agency attain a particular goal" (NCHRP Report 666).

The primary motivations for using targets in maintenance performance management are to prioritize work, reduce life cycle costs for maintenance of public assets, to attain organizational efficiencies in times of tight/diminishing budgets, and to practice proactive rather than reactive modes of maintenance. Targets communicate priorities, expectations, and work levels both internally and externally. Further, they allow an agency to continually hone budget allocations in response to both available funds and maintenance needs in pursuit of continuous improvement.

Constrained budgets and reduced funding are causing state transportation agencies to re-evaluate spending and allocations for maintenance. The National Surface Transportation Policy and Revenue Study Commission notes that the bulk of the US highway network was built prior to World War II, 13 percent US bridges are now structurally deficient and 15 percent of miles traveled occurs on pavement ranked unacceptable in the HPMS system, a level that could reach more than 40 percent of miles traveled by mid-century, under current funding levels (NTSB, 2007). Budgets have not shown signs of increasing much in the current funding environment. Many DOTs have been pressed to create budget scenarios or actual plans to accommodate decreases ranging from a few percent to more than 10 percent (Venner, 2012; unpublished data). Meanwhile, travel, especially freight travel, continues to increase.

A DOT's maintenance actions are designed to minimize asset deterioration and enable the DOT to attain acceptable condition levels. In the process of their design and application, LOS targets help DOTs to communicate asset needs, performance, and agency costs and risks. This greater knowledge and understanding can help avoid some negative outcomes. Targets and data-driven exploration of what is attainable can provide transportation staff the support they need to select actions and projects with proven ability to improve performance in agency priority areas cost-

effectively. For example, Washington State DOT found that increasing preventive maintenance on traffic signals from executing 34 percent to 79 percent reduced overall maintenance costs by cutting emergency call-outs in half, while increasing safety and decreasing user costs (Baroga, 2012; personal communication).

The initial survey and interviews for NCHRP 14-25 documented the many purposes for which DOTs used maintenance performance measures. These are listed below, in order of frequency of use. DOTs relied on measures for continuous improvement and program communication and reporting and also to guide budget allocation and re-allocation.

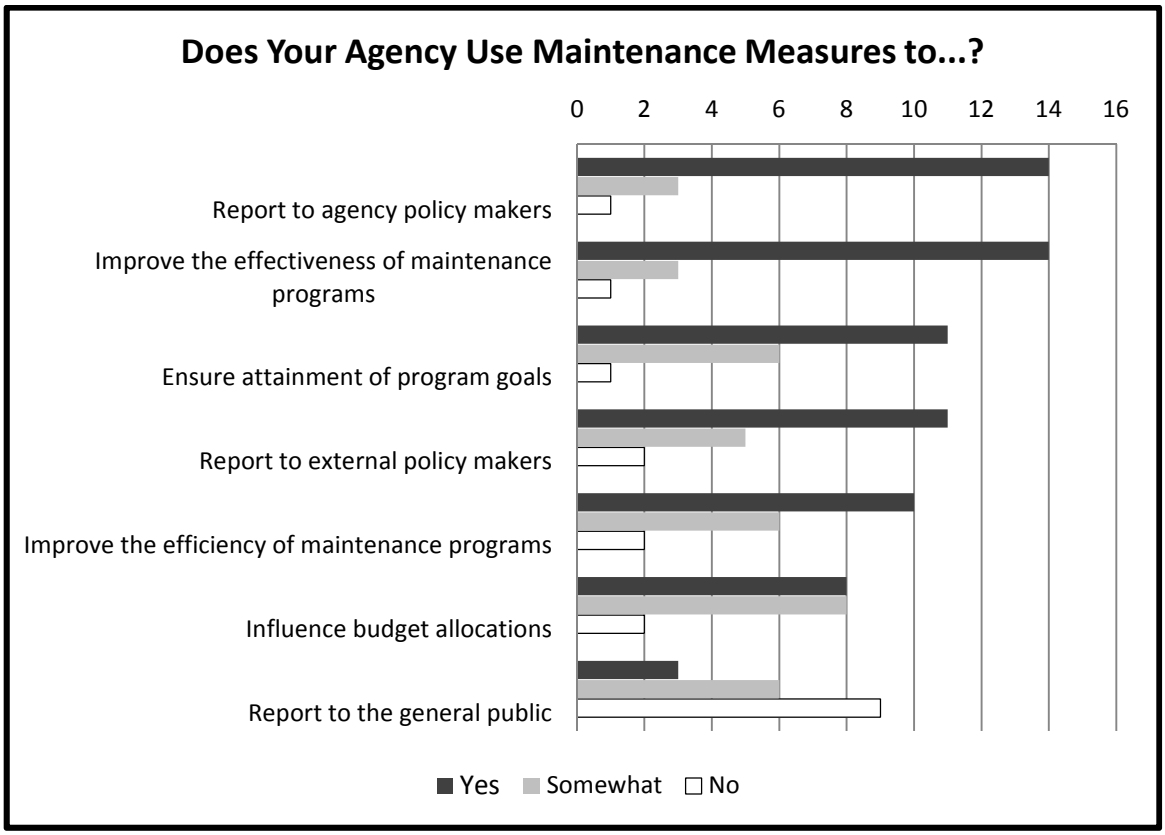


Figure 1. Agency Uses of Measures

Limited budgets constrain a DOT's abilities to act on the results of condition assessments. DOTs vary in their ability to influence the structure of the agency's budget or move funds from one program to another. A rigid budget structure can make it difficult to align agency budgets with priorities and performance measures, but increasingly DOTs are implementing enterprise risk management approaches that provide the analytics and justification for transferring resources to where they will do the most for the public, minimizing risks and producing benefit.

Continual evaluation or measurement and demonstration of progress can demonstrably build trust and respect among the DOT, elected officials, the media, and stakeholder groups (WSDOT, 2012). DOTs can take such credibility and transparency to the bank, as Washington State DOT has demonstrated with its successful bond measures and tax increases in the years since the agency implemented is Gray Notebook and Strategic Communication effort. WSDOT's efforts to assess and communicate system and agency performance helped raise support for two recent funding increases, a five-cent gas tax increase in 2003 and a nine-cent gas tax increase in 2005.

2.2. Benefits of Setting Targets for Maintenance Performance Management

State agencies report a number of benefits from the use of measures. Some of the benefits they have received from measures and targets are outlined below.

2.2.1. Alignment and Re-Prioritization

Targets help DOTs align goals and performance areas. As the Kentucky Transportation Cabinet indicated, LOS and targets "allow districts to sit down with people and find areas of improvement or focus and to identify problem areas." The agency highlights achievement or non-achievement of LOS thresholds and "money is allocated to specific areas and must be used as intended." Kansas DOT concurred with benefits of aligning agency action with agency priorities; KDOT relayed that:

> The manner of reallocating funds within the agency has improved. Measures help to reallocate to needs. They force people to pay more attention to what they're doing. KDOT has always measured outputs, but now they can ask what they have gotten for the output. The fact of measuring and inspecting has given front line employees an understanding of the bigger picture.

Wisconsin DOT shared how they have used targets and areas in which targets were surpassed to identify areas where effort may be diverted to help from other needs:

> It's given better allocation of funding and focus in the right areas…Quite frankly, you look at maintenance foremen that are out there. They work off of their own paradigms of what is important. (LOS targets allow the agency) to show from the maintenance standpoint (that) the public doesn't expect this level of performance in some areas, but more is needed in other areas. This gives the ability to have that conversation…better ability to coach people who are doing the best job they can, have them focus where the public wants.

WSDOT accounts for the contribution of maintenance activity to agency goals and objectives, by weighting of policy objectives and then weighting each maintenance activities relative contribution to each policy objective. Multiplied together, the weighting of policy objectives (e.g., Safety and Operations rank highest) and the level to which each maintenance activity contributes to that objective (0-9) produce a relative importance ranking for each maintenance activity that helps the agency ensure alignment with its highest priorities and the tradeoffs it will make among LOS levels to meet budget exigencies. WSDOT reports that the greatest benefit they've had from measures and targets in maintenance are that the DOT has been able to better communicate their program with executives and managers. The agency's credibility has improved and the DOT has more tangible goals for maintenance managers to manage towards.

2.2.2. Awareness

Arizona DOT's State Maintenance Engineer said the greatest benefit he saw was the "awareness that it brings to folks in the field. When they walk these sections and see the actual condition of some of these assets." The LOS and MQA information becomes more feedback to the districts, for their use and planning. Tennessee DOT concurred that use of targets raises awareness, simply because "county maintenance supervisors now see more of the system, since they have to evaluate it."

Minnesota DOT's State Maintenance Engineer said, "with measures and targets, they are more cognizant of performance challenges and successes. Performance measures and targets also enable MnDOT to identify whether the agency is meeting customer requirements." In addition to "making people pay attention to what is going on," Missouri DOT (MoDOT) said with targets, staff "saw

places of need and places where too much was being done" enabling them to make adjustments that could bring them into alignment with agency priorities.

2.2.3. Self-evaluation and Continuous Improvement

Colorado DOT's Maintenance and Operations Branch Manager said that measures and targets "help the agency strive to do the best they can." Florida DOT noted that the agency has been performing (more) consistently since it (the measurement and target-setting process) began. They've trended upward from the start. The District report shows the areas in which they can make changes and the annual quality assurance review looks at that plan versus what is completed in each district." The Idaho Transportation Department and many others emphasized their use of targets and measures to "improve processes" and "identify things they'll be able to do better in the future."

2.2.4. Communication

As MnDOT stated, "performance measures and targets give MnDOT a formal method of telling their story to both citizens and stakeholders;" the measures and targets "are often the public face of the department."

Utah DOT said TAM LOS targets and measures have enabled them "to report accomplishments to the legislature in an understandable and forthright manner," sometimes helping them attain needed funding. The same has occurred in Washington State; WSDOT said they've "been able to better communicate their program with executives and managers, which has improved credibility."

Colorado DOT's Maintenance and Operations Branch Manager also said that measures and targets "make it very easy to talk to the Commission. Everyone knows a C is so-so and would like to have A's. One of the Commission members had just seen his grandson's report card, and said they were going to pony up the funds to get to a B." With Bridge and Pavements, Pavements have always done Good-Fair-Poor, as has Bridge. The Commission and general public have had a hard time correlating (good, fair, poor) to A,B,C – shifting to letter grade world right now. This argument was ultimately convincing."

Agencies also check their alignment, progress, and anticipated rate of progress by evaluating targets in comparison to plans and units accomplished work activities completed, looking at costs per lane mile or other outputs, materials tracking, and maintenance backlog. MnDOT uses intermediate outcomes such as smooth pavements and clean roads to set investment priorities. At MnDOT overall outcomes and performance to targets are used to gauge customer satisfaction and set investment priorities. Nevada DOT is producing a baseline to communicate with the public about what they get for their money.

3. Research Objective, Scope, and Summary

The objective of NCHRP Project 14-25 was to develop a guide for selecting level-of-service (LOS) targets for maintaining and operating highway assets. The Guide includes a self-assessment, choices for selecting a framework and communicating targets, a discussion of the context for LOS target-setting, guidance on setting targets congruent with agency priorities, and how to manage the maintenance program with targets.

Phase 1 included (1) collecting and reviewing information on identification of maintenance features or attributes and the establishment of level-of-service targets for maintenance and operations of highway assets, (2) identification and evaluation of maintenance features or attributes and processes for establishing LOS targets, and (3) recommending methods for establishing LOS targets. Phase 2 used the findings of Phase 1 to develop, test, and illustrate processes outlined in the Guide associated with this Report.

Information on current maintenance asset management and target-setting processes was reviewed in the literature and a survey of DOTs. Phase 1 collected and reviewed relevant domestic and foreign literature and research findings and information on identification of maintenance features or attributes and the establishment of level-of-service targets for maintenance and operations of highway assets. This literature review addressed maintenance features of highway assets (e.g., pavements, bridges, and roadside appurtenances) and maintenance operations (e.g., snow and ice control) commonly included in highway agencies' maintenance quality assurance programs as well as materials related to maintenance management, maintenance quality assurance, performance management, public private partnerships, risk, and contracted maintenance.

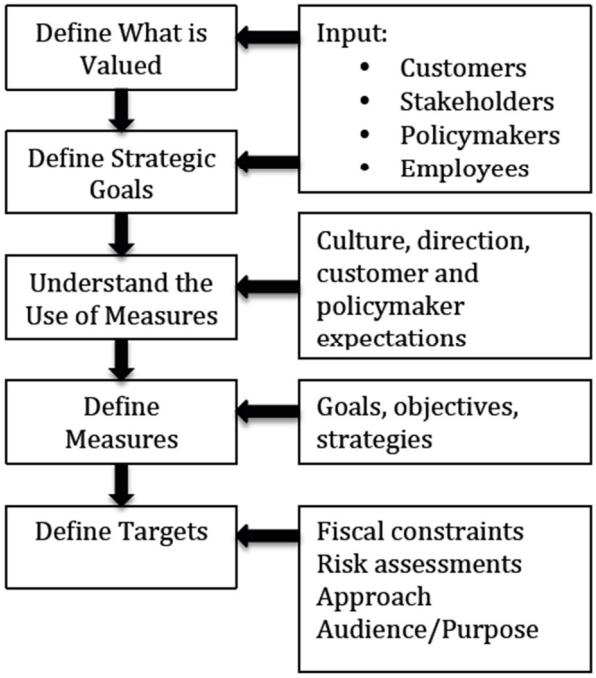


Figure 2. Content Scope for the Literature Review

In addition, maintenance and asset managers at 28 state DOTs and one Canadian province responded to a questionnaire, yielding 43 responses, some provided as a group, most provided by individuals. The survey acquired unpublished information on how agencies use performance measures and to help identify candidate agencies for subsequent in-depth interviews used to complete the guide. It covered maintenance features and attributes, methods for establishing thresholds and LOS targets, practices for relating measures and targets to strategic goals, and methods for dealing with uncertainty and risk.

Based on results of the survey conducted in Task 1, the research team identified agencies for in-depth interviews. Eighteen state and provincial transportation agencies were interviewed, including Arizona, Colorado, Florida, Idaho, Kansas, Kentucky, Minnesota, Missouri, North Carolina, Nova Scotia, Nevada, New York, South Carolina, Tennessee, Utah, Washington State, Wisconsin, and Wyoming. When possible, the research team interviewed multiple individuals within each agency, as a group or separately, for the purpose of capturing both the strategic and tactical levels of

information. Promising reports or other documents demonstrating the use of measures and targets in maintenance programs were also collected. Those reports and documents and other newly discovered references were included in the literature review. The research team synthesized the interview data and prepared a narrative summary of the findings. Most of the interviewed states have been using some form of formal maintenance targets for many years. All but four have at least ten years of experience using one or more performance measurement system.

These interview results provided the data for a synthesis of findings. Also, based on the in-depth interviews with agencies and the purpose of the targets, the research team developed a set of criteria for evaluating the applicability of various processes. Based on this evaluation, the team recommended methods for establishing LOS targets for common highway maintenance features or attributes.

The research team assembled results of the literature review, survey, and in-depth interviews to prepare a systematic process for defining methods for establishing LOS targets. The process description identifies many of the key steps, decisions, perspectives, participants, issues, tools, and techniques that must be considered in establishing LOS targets for highway maintenance. For example, the process notes the difference between attainable and desirable targets as they pertain to the level of investment, as well as the ramifications of lower LOS. Overall, the target setting method has two important components: the conceptual structures for targets and the approach used to define specific targets within that structure.

With this information, a Synthesis of Commonly used Measures for Highway Maintenance and Operations was developed in matrix form, detailing the following for each asset type:

- Feature / Physical Asset / Activity
- Attribute / Condition
- Threshold / Criterion / Tolerance Level / Standard
- Measure / Deficiency Quantity

The information gathered was reviewed, evaluated, and summarized. The initial work plan for developing the Guide updated based on this information and implemented in Phase 2. The in-depth interview results also provided the data for development of an evaluation process, a set of criteria or evaluating the applicability of various processes. The project team organized the criteria to be consistent with the structure of the interview guide and applied the evaluation criteria to the raw interview results for each of the 18 agencies interviewed. In Phase 2, the team performed in-person visits to Colorado, Michigan, North Carolina, and Wisconsin DOTs.

In Phase 2, a guide for maintenance program managers, asset managers, and maintenance quality assurance (MQA) specialists, and senior and policy managers with authority for budget-setting was developed.

4. Summary of Findings

4.1. Phase 1 and Survey Results

The Phase 1 research revealed that maintenance LOS target-setting is often an iterative process involving a range of staff throughout the agency. Some states rarely change their targets, which serve as the system-level results they expect. Others modify them with rising and falling budgets or the emphasis of Transportation Commissions, recent failures, or customer service surveys. DOTs do sometimes change the standards used to distinguish pass-fail or levels of service. States utilize a range of measure types—input, output, and outcome—to manage their programs. They do not rely on any one type of measure, but understand the importance of each for specific purposes in the management process. While some states have connected maintenance measures to the overall strategic direction of the agency, most linkages are informal. The tools used to estimate the costs of maintenance activities and the impact of activities on the goals of the program or the agency are also informal in many states. While many agree on the desirability of tools to support the agency's ability to evaluate the alternative investments that might be made within overall budget constraints, currently many of these trade-off tools are informal. Risk management is also just beginning to be incorporated into LOS target setting.

Nearly all (98 percent) responding agencies of the 28 DOTs and single Canadian province reported using management approaches that require qualitative and/or quantitative performance measures. These responding agencies made plentiful use of performance measures, across agency functions. More than 80 percent of respondents in both maintenance management (82 percent) and strategic or business planning (87 percent) made moderate or high use performance measures. Asset management and organizational efficiency were close on their heels, with 72 percent and 76 percent respectively, making moderate or high use of performance measures.

Agency or gubernatorial executives were found to play the largest role in determining what is valued enough to be measured. Agency employees also play a significant role. Legislatures, industry experts, and customers/motorists contribute significantly, while financial partners, local governments, and the general public (which includes non-motorists) contribute significantly less to identification of what is measured. Surveys and public meetings were the most commonly used tools for gathering public input, followed by focus groups, telephone hotlines, and social media tools.

Table 1 summarizes how agencies use four broad types of measures: input (or consumption), output, intermediate outcome, and outcome (or impact) measures. Input or consumption measures include such budgeted items as funds, building materials, equipment, and personnel hours. Output measures include easily measured units of production such as miles of pavement maintained and acres of grass mowed. Intermediate outcomes include measures of output quality such as smooth pavement, improved safety/accident reduction, and animal collisions. Finally, outcome measures include various ways of ascertaining driver or customer satisfaction such as reduced injuries and death, travel times, improved economic competiveness. The latter are often qualitative and indirect measures.

At the strategic decision making level, agencies most commonly use outcome or impact measures, followed by intermediate outcome measures. Both of these techniques are more useful measurements from the perspective of the intended end user than input and output measures. At the network (tactical) level, responding agencies primarily use measurements of output production and intermediate outcomes with equal frequency. As with strategic level decisions, agencies relied more on these types than simpler measures of input and output. At the project (operational)

decision level however, the choice of predominant measures shifts to input and output measures. Half of the responding agencies relied most heavily on measures of output production, while another third primarily tracked inputs. Few agencies used performance measurement of intermediate outcomes or end-user outcome and impacts at the project or operational level of planning.

Table 1. Types of Performance Measures used for Strategic, Network, and Project Level Decision-making

Type of Measure	Strategic Level		Network Level		Project Level	
	Responses	%	Responses	%	Responses	%
Input Consumption	3	8%	3	8%	12	32%
Output Production	3	8%	16	42%	19	50%
Intermediate Outcome	14	36%	16	42%	4	11%
Outcome/Impact	19	49%	3	8%	3	8%
Total	**39**	**100%**	**38**	**100%**	**38**	**100%**

Of the DOTs responding that they use performance measures to communicate the condition or needs of the system to external parties and guide agency actions and communicate directions to agency staff; 90 percent used performance measures "moderately" or "highly" toward these ends. DOTs also use performance measures to define problems and find solutions to those problems (72 percent "moderately" or "highly").

Within responding agencies, 95 percent report using performance measures at the program and asset-class levels. Use of measures at the division, bureau, region, or other major organization level and use at the agency level is nearly as widespread (87 percent). Approximately half of responding agencies use measures to convey the relationship between employee roles and broader agency goals. All responding agencies use pre-defined performance targets or standards for measuring successful outcomes. Historic trends are also widely used (62 percent of respondents). Few agencies (18 percent) directly compare measures or results to those of other agencies or jurisdictions.

When asked how they set performance targets, a majority of responding agencies reported setting performance targets based on desirable or aspirational outcomes. Of these responding agencies were evenly divided between those that set performance without regard to the budget (36 percent) and those that balance aspirational with attainable targets (38 percent). Few agencies limit targets to only those that are attainable with available resources or which are derived from historic trends independent from current budget realities.

Table 2. Methods of Setting Performance Targets

Method for setting performance targets	Responses	%
Both attainable & aspirational targets are defined to illustrate current budget shortcomings	15	38%
Targets are defined that illustrate what would be a desirable outcome, without regard to budget	14	36%
Targets are selected that can be attained within available resources	5	13%

Method for setting performance targets	Responses	%
Targets are selected based on historic trends and may not reflect current budget reality	4	10%
Target reflect the type and purpose of the measure	1	3%
Total	**39**	**100%**

Table 3 summarizes the relative importance of various methods of evaluating and prioritizing maintenance activities. Management systems top the list of most highly used approach, but this, expert judgment of asset or program managers, and life cycle cost or other economic analysis are "moderately" or "highly" used by more than 4 out of 5 respondents. Other methods included iterative processes between program managers and the department's commission, use of enterprise risk management tools, and collaborative processes with affected stakeholders. Strictly political decisions were least common, indicating that evidence-based decision-making is squeezing out politically influenced decision-making. Still, infrastructure failures and other issues that hit the news factor into maintenance decision-making and demands by the executive and legislative branches, and would be considered "political" input.

Table 3. Agency Approaches to Prioritizing Investments

Approach for prioritizing investments	Not Used	Little Used	Moderately Used	Highly Used	Total Responses
Management systems (pavement, bridge, etc.) are used	0%	18%	28%	54%	39
The expert judgment of asset or program managers is used	8%	8%	54%	31%	39
Life cycle cost or other economic analysis is used	5%	18%	54%	23%	39
Decisions on resource allocation are made politically	8%	38%	46%	8%	39
Other	29%	14%	29%	29%	7

As indicated in Table 4, most of the survey respondents report using one or more techniques to evaluate different categories of risk. Agencies relied most heavily on expert judgment to assess most categories of risk. One notable exception was widespread use of formal modeling processes to calculate risks and costs associated with deferring maintenance. One agency reported that risk evaluation was important to operations, but not for performance measurement.

Table 4. Methods for Evaluating Risks

Type of Risk	Formal Modeling Process	Expert Judgment	Probability of Failure	Other	Risk is not considered	Total Responses
Risk of legal liability for actions taken or not taken	5%	55%	13%	5%	21%	38

Type of Risk	Formal Modeling Process	Expert Judgment	Probability of Failure	Other	Risk is not considered	Total Responses
Risk associated with the uncertainty in the many estimates made in planning and management processes	14%	61%	6%	0%	19%	36
Risk of higher costs due to deferred maintenance	41%	32%	14%	3%	11%	37
The consequences of budget uncertainty	22%	54%	8%	3%	14%	37
The consequences of a facility failure	8%	61%	17%	6%	8%	36
Other	25%	0%	0%	50%	25%	4

Nearly four of five responding agencies use defined maintenance management programs or maintenance quality assurance programs. Table 5 illustrates the variety of specific asset categories for which performance measures are used. Use of performance measures for roadway pavements was nearly universal. With the exception of traffic control devices, use of maintenance and operations performance measures for other transportation related asset categories is also widespread. Categories of maintenance for which performance measures are used but not listed in Table 5 include tunnels, fences, guard rails, litter, sweeping, road clearing, and the condition of maintenance equipment.

Table 5. Use of Performance Measures by Specific Maintenance or Operation Category

Category		Responses	%
Roadway pavements		30	97%
Traffic Management (signs and pavement markings)		26	84%
Vegetation		26	84%
Roadsides		24	77%
Bridges and structures		24	77%
Culverts		24	77%
Drainage		23	74%
Roadway shoulders		22	71%
Snow and Ice removal		21	68%
Rest Areas		17	55%
Traffic Control Devices (ITS technologies)		13	42%
Others		6	19%

Several agencies emphasized the need to work within budget constraints. For example, Wyoming has established a formal sampling approach to inspecting pavement conditions at regular intervals because resources do not allow exhaustive inspections. The data is then used to prioritize repairs where they are most needed. Typical of many responding agencies, another agency reported using a

variety of individual performance measures, but ultimately relying on the expertise and iterative decision making to evaluate, prioritize, and set maintenance targets.

Several state agencies commented on their attempts to develop more sophisticated performance management approaches and analytic tools that are driven by performance measures. Washington State has been utilizing the Maintenance Accountability Process (MAP) to link the tools to strategic planning, the budget and maintenance service delivery since 1996. Colorado uses Maintenance Levels of Service (MLOS), a performance-based budgeting system. It evaluates money spent, production outputs, and actual level-of-service attained versus targets. The tool can use the performance measurement data to model different program scenarios for the next budget. The state transportation commission can then use the modeled scenarios to establish goal and finalize the budget. After ten years of experience, another agency is currently undergoing a comprehensive review of their Maintenance Quality Assurance (MQA) program. At the other end of the spectrum, some state agencies are initiating defined performance measurement and performance management systems for the first time, drawing from the experience from other agencies.

4.2. Synthesis of Commonly Used Highway Maintenance & Operations Measures

After collecting information from the literature, interviewing agencies, and drawing on personal experience in state highway agencies, it was possible to construct a synthesis of commonly used measures for highway maintenance and operations in sufficient detail to discern a number of patterns and trends. The following table summarizes some of the principal characteristics observed in the data. See Appendix 4 of the Guide for the full synthesis.

Table 6. Summary of Commonly Used Measures

Major Element	Characteristics	Measures	Targets
Network	General	• Satisfaction Index, e.g., 7.0/10.0	• Percent meeting qualitative condition, e.g., Level of Service (LOS) A, B, etc.
Pavements & Shoulders	General	LOSs, e.g., A, B, C, etc.	
	Defects	• Physical measures, e.g., rut depth, areas of potholes, etc. • Percentage affected • Total number of defects • International Roughness Index (IRI)	• Speed of repair, e.g., within 2 days • Cyclical, e.g., 4-year cycle • Percent in various condition by sub-system, e.g., Interstates, arterials, etc.
Drainage	Defects, blockages, missing / broken components	• Physical measures, e.g., length of blocked ditches, etc. • Percent blocked	• Cyclical • LOS targets
Traffic Control Devices	Defects	• Physical measures, e.g., legibility, night visibility, etc.	• Cyclical • Repair / replace within a time frame, e.g., 24 hours • LOS targets • Cyclical – replace a percentage annually
Winter Maintenance Ice & Snow Removal		• Customer satisfaction rating (index) • Hours to achieve bare pavement by types of	• Achieve bare pavement for various types of facility within a time frame • Percent bare pavement

Major Element	Characteristics	Measures	Targets
		facility • Time to restore traffic flow / LOS • Time for initial response	within a certain time • Percent of initial responses within a target time
Bridges & Structures	General	• Percent structures in qualitative condition, e.g., good, fair, etc.	• Percent in various qualitative condition
	Defects	• Physical measures of deficiencies	• Cyclical • Programmed maintenance, e.g., bearing lubrication • LOS targets • Programmed painting, washing, etc.
Rest Areas	General	• Qualitative measures of adequacy	• Qualitative levels of adequacy, cleanliness, etc.

Quantitative vs. Qualitative Measures. Some measures are quantitative while others can only be expressed qualitatively. Output measures tend to be quantitative such as lineal feet of cracks sealed or man-hours spent in snow removal. Outcome measures are often expressed qualitatively as indices such as graduated levels of service (LOSs), e.g., A, B, C, etc. Other outcome measures may be quantitative such as mean response times for the network or the percentage of response times falling below certain thresholds.

Physical vs. Index Metrics. Defects tend to be based on physical measurements such as areas of distress, percent of surface affected, etc. For example, drainage measures focus on blockages and water accumulation. Targets are frequently cyclical, e.g., replacement of a certain percentage of the feature annually. Other targets may be expressed as LOSs which reflect quantitative ranges of some variable such as specifying that the time to clear snow on roads carrying ADT ranges (e.g., 10,000 to 30,000) should be within the range of 2 to 5 hours.

Level of Aggregation. At the network level, measures tend to be percentages meeting certain qualitative thresholds. For example, MnDOT has a target of 7.7/10.0 for customer satisfaction with state highway maintenance, while VDOT's target for lane-miles in sufficient condition on Interstate and Primary routes is at least 82 percent.

Relationship to User. When highway users are directly affected, performance targets for surface defects tend to focus on the speed of repairs or are cyclical, e.g., crack sealing each *n* years. For traffic control devices, the focus is on functionality, hence safety, e.g., legibility, night visibility, reflectivity. On the other hand, most users are not interested in segment details such as the number of ruts or their average depth.

Reactive vs. Anticipatory. Another way of looking at measures is to distinguish between repairs and preventive maintenance. Repairs merely restore functionality to an asset—they do *not* extend life. Performance of repairs is often judged by speed of response, e.g., repairing potholes within two days. On the other hand, performance targets for surface defects may be cyclical, e.g., crack sealing each *n* years, or cleaning ditches on a 10-year cycle (NYSDOT), or replacing guardrail on a fixed cycle (NYSDOT). These are examples of preventive maintenance.

4.3. Processes Used for Establishing Targets for Highway Maintenance and Operations

The research team conducted multi-hour (in some cases multi-day) in-depth interviews with 18 states and provinces including Arizona, Colorado, Florida, Idaho, Kansas, Kentucky, Minnesota, Missouri, North Carolina, Nova Scotia, Nevada, New York, South Carolina, Tennessee, Utah, Washington State, Wisconsin, and Wyoming. Most of the interviewed states have been using some form of formal maintenance targets for many years. All but four have at least ten years of experience using one or more performance measurement system.

The impetus to develop measures and targets for maintenance varied. In many cases, the requirements were mandated by state legislatures as part of a wider movement toward performance-based budgeting or similar efforts to link budgets to outcomes. Some agencies reported being driven more by the need to compete for increasingly scarce public funds, others as a useful tool to improve and demonstrate greater agency effectiveness and efficiency.

The greatest reported benefits from adoption of road infrastructure maintenance measures and targets include the assembly of better data to assess the overall state of the transportation system and the prioritization of maintenance on a more objective level, rather than a more subjective, anecdotal, and *ad hoc* basis. The measurement information is useful for communicating priorities to agency heads and legislatures so that they can make the best use of available resources and support the case for additional budget resources (or for protecting existing budget resources). Targets and measures also show administrators and policy makers how well an agency is meeting its strategic goals and objectives. The information can also be used across subunits and downward to regional managers and road crews to illustrate how local priorities fit in with or compare to maintenance needs elsewhere in the state.

Interviewed officials reported several types of challenges to successful implementation of maintenance measures and targets. The first is cultural—overcoming initial resistance to implementation, and convincing everyone involved of the benefits. Long-time line personnel may be suspicious of the purposes of measurement, lack the expertise or resources to collect and report data, and be skeptical of its utility, especially for the effort required. Unless these concerns are addressed, line personnel might actively, passively, or unintentionally undermine efforts to implement maintenance measures and targets. Supervisors and budget staff often need to be convinced that the additional resources needed for effective implementation are worthwhile. The second challenge is technical—developing an objective and consistent methodology for collecting data. Data collected by local crews can be biased and/or inconsistently gathered and reported over space and time, making the data less useful. Cross-training of staff or hiring of dedicated measurement inspectors can improve the quality and consistency of data collection, but can be costly. Staff turnover is an ongoing challenge to data consistency. Finding the best balance between funding for performance measurement and funding for actual system improvements is difficult, funding for performance measurement is often constrained. Finally, agencies struggle with effectively integrating measurement data and truly cost-effective strategic planning. Competing political priorities can further complicate matters.

4.3.1. How Does an Agency Choose What it Values?

In a world of limited resources, state transportation agencies must determine the relative value of services and assets when allocating maintenance resources. How an agency goes about defining the relative importance or value of assets is crucial in developing performance measures and rating systems that inform and reflect the strategic mission of the agency.

Maintenance priorities and activities are potentially affected by multiple stakeholders. Each group of stakeholders has its own values and priorities, some of which complement and some of which compete with those of others. Categories of stakeholder include the end users (commercial and non-commercial users); other stakeholders indirectly affected by transportation decisions (environmental groups, etc.); elected officials who must balance demand for transportation with demands for other services and restrained spending; allied fields such as contractors; suppliers; consultants; local public works agencies; and internal staff (senior management, mid-level management, and line crews).

Figure 3 shows how agency representatives viewed the relative significance of each of these stakeholders when defining maintenance targets and priorities. While responses varied somewhat from state to state, interviews with state transportation agencies reveal that input from within agencies typically plays a more significant role in defining maintenance targets and priorities than input from external stakeholders, including state executives and legislators. To varying degrees, state elected officials appear to have some influence on the relative importance of establishing maintenance measures, standards, and reporting methods. In addition, they may on occasion exert their authority for specific maintenance targets or projects. But agency staff—and to a lesser extent, partners—in local government and allied fields, are those most responsible for identifying and prioritizing targets on an ongoing basis. They are also responsible for generating much of the data used by elected officials. In most states, highway users (commercial and non-commercial) and other advocacy groups played less significant or even insignificant roles in understanding the relative value of services and assets, at least directly.

The tools used for gathering and disseminating input on values depend on who is providing the input. For internal agency input, the agency head and senior staff usually establish the type of information and manner in which it is to be reported by and to agency subunits. Interviews, meetings, and documents are among the specific tools most frequently used. Collected information is reported up the organizational hierarchy, usually through mid-level and regional subunit staff. For states which do make an effort to gather outside stakeholder input, various forms of survey or comment collection are used. State-level input by elected officials usually occurs once a year at most. Agencies assemble data and develop budget recommendations; state officials decide.

Figure 3. Role of Stakeholders in Defining Maintenance Priorities

Most states reported that a significant dichotomy exists between engineering-oriented priorities geared toward system maintenance (driven from within the agency) and a customer- or policy-driven orientation with value inputs from the public. Different states address this dichotomy in different ways, and the states reported different levels of success in reconciling the two. In states where elected officials tend not to interfere, the balance skews toward the engineering and system maintenance orientation. In states with more political involvement, public demands often compete with mundane but essential capital projects such as the rehabilitation of bridges. Agencies most often attempt to reconcile this dichotomy through the adoption of long-range capital budgets, which in turn minimize subsequent political interference. States such as Washington and Colorado have or are considering dual measures or a tiered approach to developing maintenance plans where competing values are scored and balanced in a systematic way.

In order to better balance competing priorities, about half of the states reported attempting to improve their approach to gathering stakeholder input in recent years. These changes have mostly been incremental, have met with mixed success, and in at least one case the quality of input from internal agency employees actually decreased in response to complaints over the effort required to fill out detailed information and to demands by a legislature to cut budgets. Several states have

expressed frustration with the use of letter grades (A-F) at expressing level of service (LOS) because they are too general to be used for prioritizing of projects or communicating with legislators and the public. Efforts to obtain policy input from the public through surveys and focus groups has mostly been sporadic and with disappointing results. Those agencies that have tried have struggled to translate this input into a scoring system that can be compared with engineering-oriented objectives.

Agencies stated several pieces of advice for implementing a formal approach to understanding and communicating the competing values. Two-way communication is essential in obtaining buy-in from internal stakeholders. The resource demands of implementing a formal approach to obtaining and measuring input should not be ignored. Dedicating reliable, expert staff to obtain uniform, quality information is ideal, but can be costly and may create problems with line staff and budgets. Trade-offs between quality and cost of gathering input must be made. Poor data can be of little or no value and too much data can yield diminishing even negative returns on investment. Trade-offs must also be made between the scope of the input desired and ability to deliver on outcomes. Focusing on core missions (i.e., bridges, roads, etc.) and measures makes for better management, especially when work effort/focus and measures align with missions. The effective communication of the value of system maintenance to officials and general public in a cost-effective manner can improve buy-in from those parties.

4.3.2. Connecting Maintenance Measures and Targets to Agency Strategic Goals and Direction

Figure 4 summarizes the self-reported relationship between maintenance measures and agency strategic goals. All of the interviewed agencies reported that maintenance measures and targets are at least somewhat related to their strategic goals, and most reported that they work toward the same underlying philosophy. The degree to which maintenance measures are mapped to an agency's strategic vision and the degree to which they are integrated with analytic tools and reporting structure varies from state to state. Most agencies reported that agency executives are involved with defining maintenance measures and targets in a manner that makes them compatible with strategic directions. Only half of interviewed agencies indicated that maintenance measures and strategic directions flow from the same planning process. Strategic directions are also derived from other criteria and processes, such budgeting and capital spending.

Most agencies reported that their maintenance field workers understand how their efforts contribute to the overall agency's strategic direction, at least within the worker's particular geographic area or specialization. Two-way communication is the key to this understanding. Workers react and presumably perform better when asked for input. These workers know when they have been heard, in particular if they see how their data collection efforts are translating to resources and ability to execute work. Proper communication of measurement tools and their uses can therefore serve multiple productive purposes.

Frequency of communication is also important. Field workers and line supervisors are often too busy to dwell on the bigger picture. Mid-level managers are often more in tune with the agency direction than field workers. Therefore, senior management must make ongoing efforts to explain the importance of asset management and make the connections between data-gathering, strategic planning, and funding decisions, thereby nurturing the organizational culture at all levels. For example, Colorado has an innovative program that directly assigns specific roads to specific crews, making them the primary point of contact with the public. This system fosters a sense of ownership, accountability, and pride.

About half of the agencies reported that maintenance funds are isolated or mostly isolated from funding for the agency's capital improvement and other programs. Transfers from one type of fund

to the other can go in either direction. Some states have made dramatic cuts to maintenance and preservation, including Utah DOT, which recently dropped all preservation work on its secondary, non-urban roads. Most states reported that they based their maintenance budgets on historical trends or a mixture of historical budgets and maintenance projections. Often maintenance needs drive the distribution of funds, not the level of funds.

Agency representatives provided several pieces of advice for linking performance measures and targets to strategic direction. One warned that establishing and maintaining a performance management culture takes perseverance and a large commitment of time and resources. Noted benefits include the possession of a tool that can be used to show decision makers the relationship between outcomes and specific maintenance decisions. Strategic direction should be tied to broad but achievable customer-based outcomes (e.g., smooth pavement or system performance) and specific measures should be identified and collected accordingly. Once in place, the objects measured (and perhaps the measures themselves) should be reviewed frequently to see if they are working.

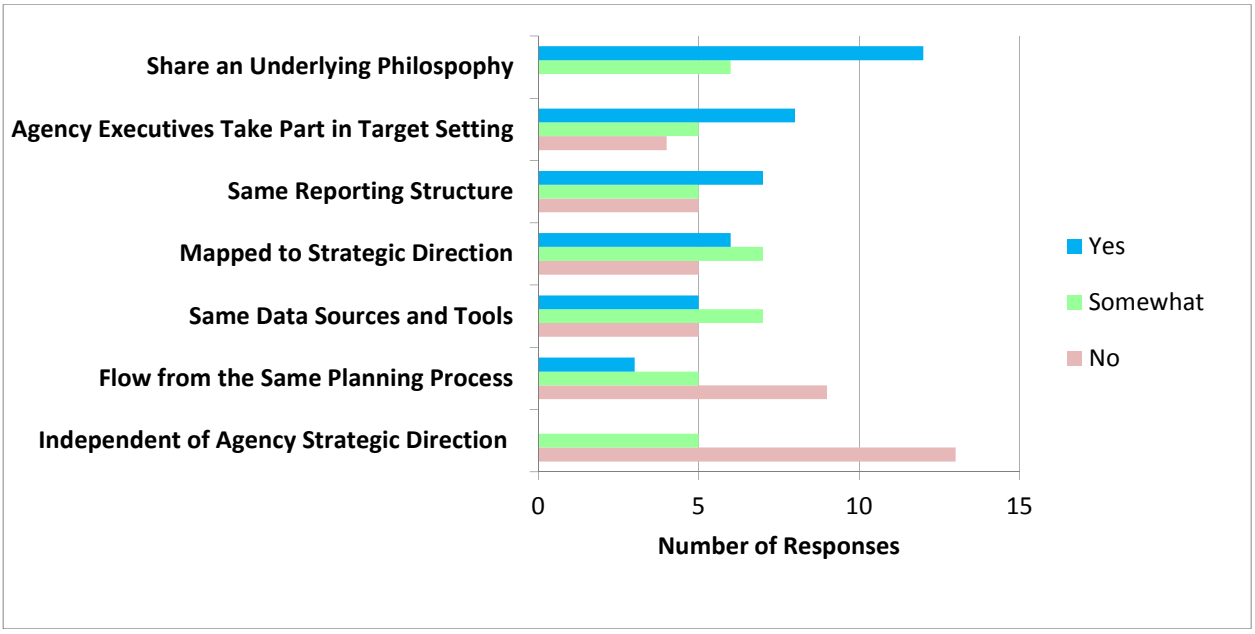


Figure 4. Relationship between Agency Maintenance Measures and Targets and Strategic Direction

4.3.3. Using Maintenance Measures and Targets

How maintenance performance measures and targets are used often influences how they are defined and structured. Figure 5 shows that most agencies interviewed use their maintenance program for internal uses: to improve program effectiveness, increase efficiency, and measure outcomes of the maintenance program. Other common uses include reporting results to policy makers and administrators within the agency. Reporting results to external policy makers, budget officials, and the general public is emphasized less on average. The relatively weak role of maintenance measures and targets in determining budget allocations in half of the states is noteworthy, though some states shared anecdotes of how Transportation Commissioners sought to raise LOS target grades.

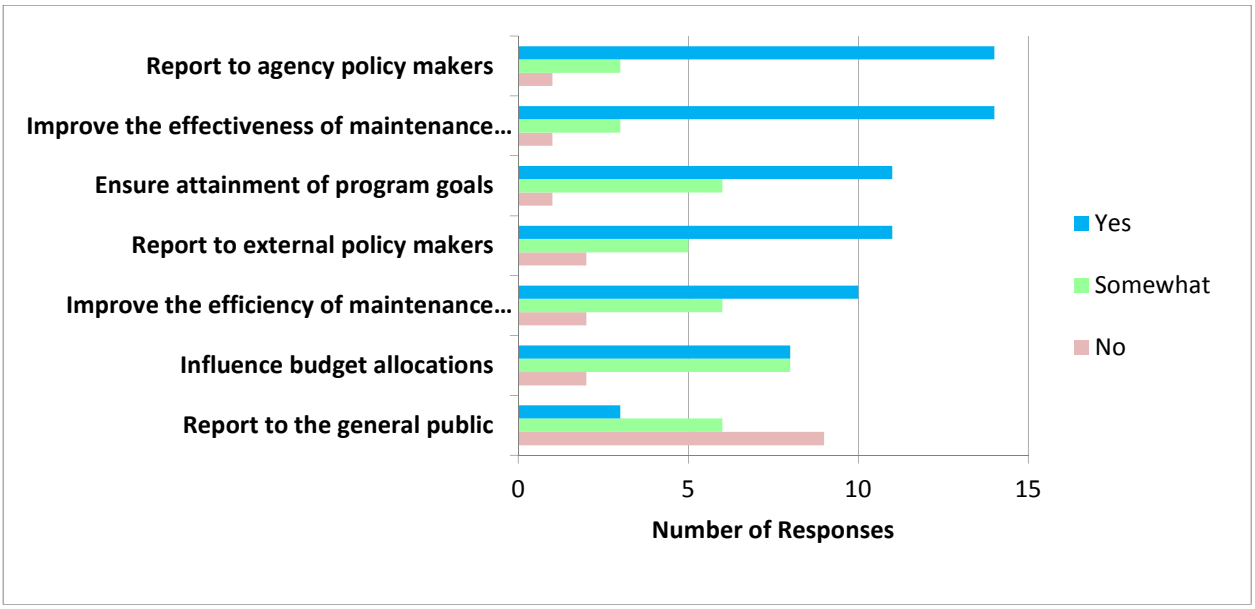


Figure 5. Use of Measures and Targets in Maintenance Programs

Approximately half of the interviewees stated they used mostly the same measures and data to serve multiple purposes, with some tailoring of the level of detail for specific needs. Some use systems such as Maintenance levels-of-services (MLOS) or Maintenance Rating Programs (MRPs) that have better levels of detail. Of those that use a single system, some reported that single simplified measures such as A-F or pass-fail grading are not detailed (or objective) enough to guide maintenance action. As a result, many are moving toward new systems, complete inventories tied with maintenance data, or ensuring that they retain more specific information used for the overall grade. Condition assessments of a representative sample of assets are often seen or criticized as unrepresentative or otherwise inapplicable, by frontline staff. Changing or tweaking systems must be done carefully insofar as it can cause confusion, but states maintained that the pay-off is worth the effort.

Respondents emphasized the importance of tailoring measures to their intended end uses. While overly simple measures are not useful to anyone, using all potential measures and collecting the necessary data may require more resources than agencies can feasibly maintain and may exceed what is necessary or even desirable to answer the most pressing questions. Some DOTs are taking a careful look at this question, and the overall trend in recent years has been to scale back the number of asset condition assessment measures used. Limiting information gathering to what is strictly needed also reduces internal opposition to its collection. Essential questions for agencies to consider when deciding what maintenance measures to use include: What data are most effective or needed to communicate to legislatures and other decision makers? What measures are needed to inform the strategic goals? How efficiently can they be collected?

Most states rely on a mix of traditional and outcome measures. Nine interviewed agencies reported using all four types of measure. Four interviewed agencies reported combining some use of one or more outcome measures with both traditional input/output measures. The remaining four agencies reported relying much more heavily on one or both types of outcome measure than on input/output measures. In total, use of all types of measure was approximately equal, though slightly more agencies reported using outcome and intermediate outcome measures than traditional input/output measures. In general, outcome measures are more difficult to measure because they lack widely accepted, objective standards, require more resources, or both. A majority of interviewed agencies reported using formal review systems with varying degrees of hierarchal

or rating schemes. Systems for scoring the conditions of pavements and bridges are the most widely established and used. Several agencies are attempting to implement or would like to implement broader, more comprehensive measures for network performance and LOS but few have been successful. One potential motivation for developing comprehensive measures of all types would be for the administration and oversight of private contractors for states that outsource maintenance, but few of the interviewed agencies contract out their maintenance.

Agencies stressed that they review maintenance data on a fairly regular basis. Re-evaluation of the maintenance measures themselves is less frequent and less regular, with some states reporting that they review their program on a five-year basis. States use both formal and informal review processes. Several agencies with established and comprehensive measurement systems review their systems for data collection the least often. Owing to their experience, these agencies advise being selective in adopting only those measures that are most useful and can be consistently and reliably collected. Realism with regard to the resources needed to collect and analyze data is a key to success.

4.3.4. Defining Targets

How agencies establish performance targets is critical to the success of maintenance programs. Figure 6 summarizes the various approaches to defining maintenance performance targets. Most agencies take an incremental approach: defining targets in terms of straightforward, readily available measures and comparing current and target performance to the past. Defining targets by comparing operating subunits is also fairly common. Fewer agencies actively compare their targets with those of other agencies or develop optimal targets independent of available or identifiable resources. However, a couple of agencies identified the benefits of less frequently used approaches. Utah recently began encouraging subunits to develop more ambitious targets, which are then compared with those of other subunits. Though no subunit gets all they ask for, it is seen as encouraging subunits to strive for better results. Wisconsin develops optimal targets in addition to attainable ones at the state level to demonstrate the need for state transportation funding with respect to existing investments.

Most agencies stated that the process of defining maintenance targets was both constrained by agency budgets and used to define them. More agencies stated that operating within budget constraints dominated the positive role that defining maintenance targets could play in defining the budget and building the case for additional funds. Although most states use targets to communicate maintenance needs to state elected officials, other state revenue and spending priorities act as constraints. Consequently, defining maintenance targets plays less of a role in establishing the state's budget allocation than it does in dividing up the resources after the state budget is completed.

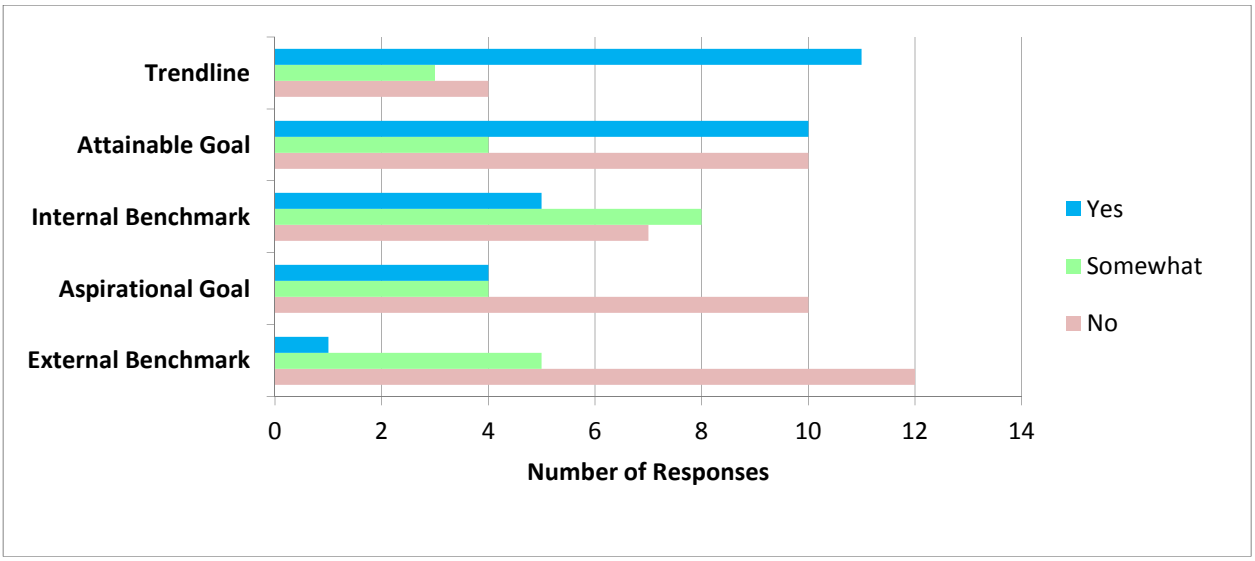


Figure 6. Guiding Parameters for Setting Maintenance Performance Targets

Figure 7 summarizes the various approaches that agencies use to define maintenance targets. In general, agencies exercise a great deal of autonomy in defining maintenance targets. For the most part, external policy makers, budget negotiators, and service providers (public or private) play little or no role, although as noted above, external actors and constraints can shape agency planning and asset management processes.

There were large variations in how often states defined or updated their maintenance targets. Approximately one-third do so annually, and several do so biennially. A few states that do so regularly revisit them at different intervals depending upon the asset or circumstances. Some states do not review maintenance targets on a regular basis or established schedule. This is more typical of states that have only recently implemented or are considering implementing more comprehensive approaches but it is also true of some agencies with established programs.

Agencies use a wide variety of methods to monitor their targets. Some have formalized reporting requirements, the frequency of which varies from state to state, level of subunit within the agency hierarchy, asset type, and other factors. Some states do not have a formal protocol. Agency policies also vary for corrective action when targets are not met. Most hold district or other subunits accountable through audits and review of required reports. If failure is the result of budget shortfalls or unforeseen events (e.g., reallocation of funds to meet emergencies), little is done other than adjustments within agency. If not a result of external factors, responses range from the informal to formal administrative actions such as employee evaluations.

The specific processes used by agencies to set targets depend somewhat on the tools and expertise available. DOT maintenance and asset managers tend to utilize the output of the various systems they use, in dialogue with Region/District maintenance personnel. States draw on their Maintenance Management (or Activity Tracking) Systems (MMS), Maintenance Accountability Process (MAP) as in Washington State, output data from Automated Vehicle Location (AVL) and mobile data collection systems, Geographic Information Systems (GIS), Emergency Road Maintenance (ERM) systems, and various business and financial analysis systems. Most still rely on narrower silo systems that address specific assets but cannot be used to assess the relative importance of a given asset (for example, a specific route or bridge) to the transportation system as a whole. Other agencies reported efforts to define targets that are more closely linked to strategic

outcomes. There are significant technical, political, and resource challenges to implementing these approaches.

When asked how maintenance targets *should* be defined, comments varied. Agencies emphasized that they should align with stated top priorities and efficient use/expenditure of public funds. In maintenance, some states such as Washington are devoting increased attention to task completion and stressing agency-wide ownership of asset condition, given the influence of preservation and improvement investments. Most interviewees conceded that target definition must strike a balance between the competing factors. Regardless of priorities, targets should be as simple, understandable, and as appropriate to the objective as possible.

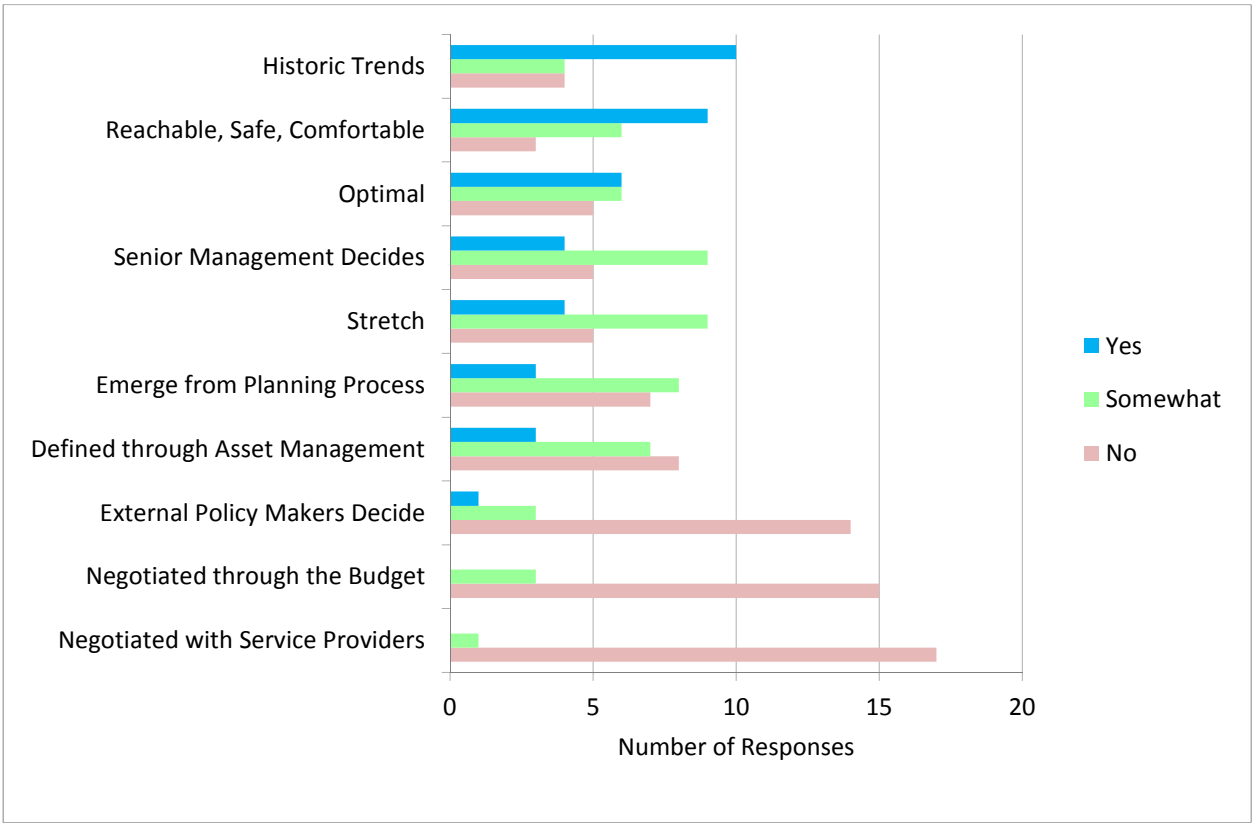


Figure 7. Agency Approaches to Defining Maintenance Targets

4.3.5. Optimizing Attainable Outcomes and Allocating Revenues to Maintenance Targets

In order to optimize attainable maintenance outcomes, agencies must develop systems for allocating expected revenues and resources to maintenance targets. Interviewed agency representatives relied on a number of approaches to ranking and programming maintenance activities within their budget constraints. Most of the agencies interviewed relied at least somewhat on all of the methods of matching maintenance targets to resources listed in Figure 8.

Benefit-cost and capital-maintenance trade-offs are among the broad economic approaches used by agencies to match targets to resources. Most agencies reported using benefit-cost analysis or similar analysis in setting performance goals. For example, Utah uses Maintenance Management Quality Assurance (MMQA) and Washington uses fairly detailed analyses for pavement preservation and bridges. However, Washington noted that for many agency activities, the level of predictability and precision necessary for effective benefit-cost analysis was unavailable, making any analysis less predictable and useful.

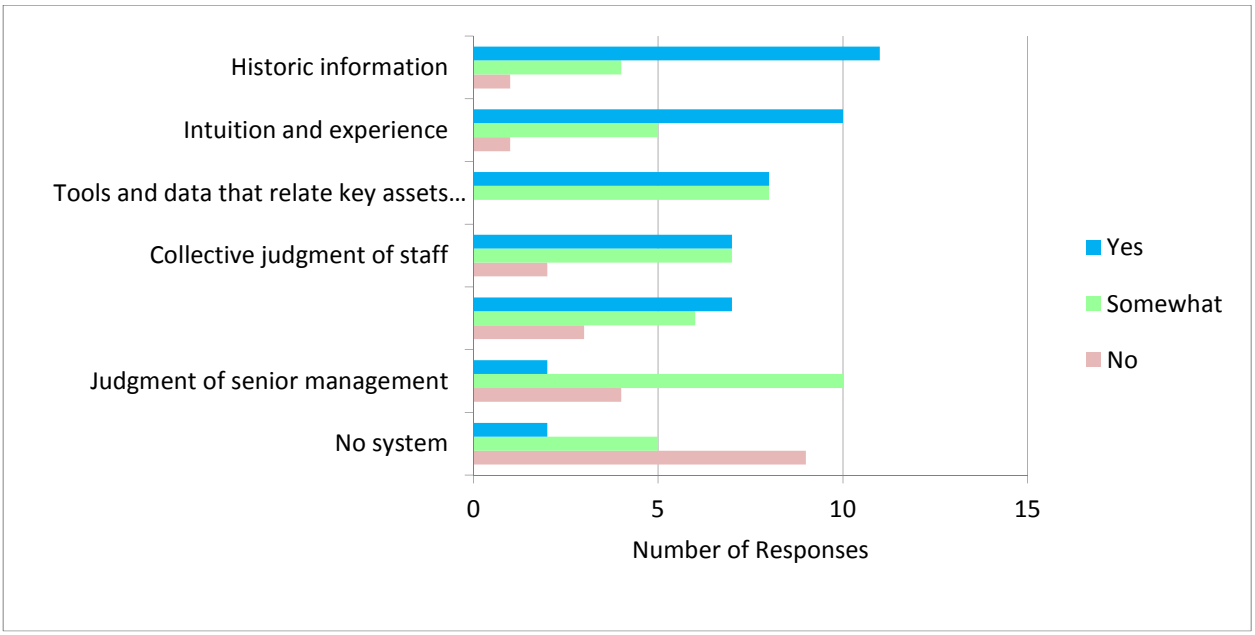


Figure 8. Methods of Matching Agency Maintenance Targets to Resources

Several states use tools to help determine trade-offs between capital and maintenance expenditures. Pavement Management Systems (PMS) are the most commonly used tool for this purpose. Other specific tools noted by interviewees included Maintenance Levels of Service (MLOS) (Colorado), Maintenance Budgeting System–Level of Service (MBS-LOS) (Arizona), Agile Assets (Wyoming), various GIS systems, and–in cold weather states–weather prediction tools to assist with snow and ice management. Long-term levels of service are assessed more at the conceptual level than by means of a more specific analytic tool, and long-range outcomes generally are neglected due to more immediate concerns, lack of available analytic tools, or both.

4.3.6. Understanding Risk and Reliability in Relation to Targets and Measures

Risk and reliability are aspects of highway maintenance with important consequences to community health, agency effectiveness, and responsible governance. Figure 9 summarizes the extent to which interviewed agencies incorporate considerations of risk and reliability into their maintenance programs, metrics, and target-setting. To some extent, almost every agency interviewed has integrated these considerations into the various specific activities, albeit sometimes informally or incrementally. The level of formality and emphasis can vary greatly depending on agency, agency subunit, nature of the risk, type of asset, and who is affected. For example, Utah has developed an approach to using accident data and asset damage to not only identify and correct specific problems (e.g., guard rail accidents), but also to investigate root causes and apply preventive future actions. Most agencies conduct employee training on the safety, medical, and legal risks associated with unsafe maintenances activities. Except for traffic safety data, few agencies formally measure the severity of risk specifically associated with maintenance programs. Some notable exceptions included Colorado's attempts to predict avalanches, Minnesota's and Idaho's winter hazards, and Washington's efforts to anticipate the impact of increasingly extreme weather events and accordingly, change design standards for assets such as drainage culverts.

The reliability of highway maintenance data is another potential area of public concern. Only a few agency representatives provided responses to questions regarding data reliability. Of those responding, the collection of large samples, redundant samples, and oversight and shadowing of

data gatherers and inspectors were among the methods used to ensure data reliability. Washington reported using duplicate surveys for their MAP field surveys, with at least one of the surveys completed by crews from outside of the area in question. Most responding agencies systematically divide assets into smaller units of analysis (e.g., specific segments of road) which helps ensure consistent and complete data collection. However, the sheer number and scale of some assets (e.g., culverts) means that data are not always complete or current.

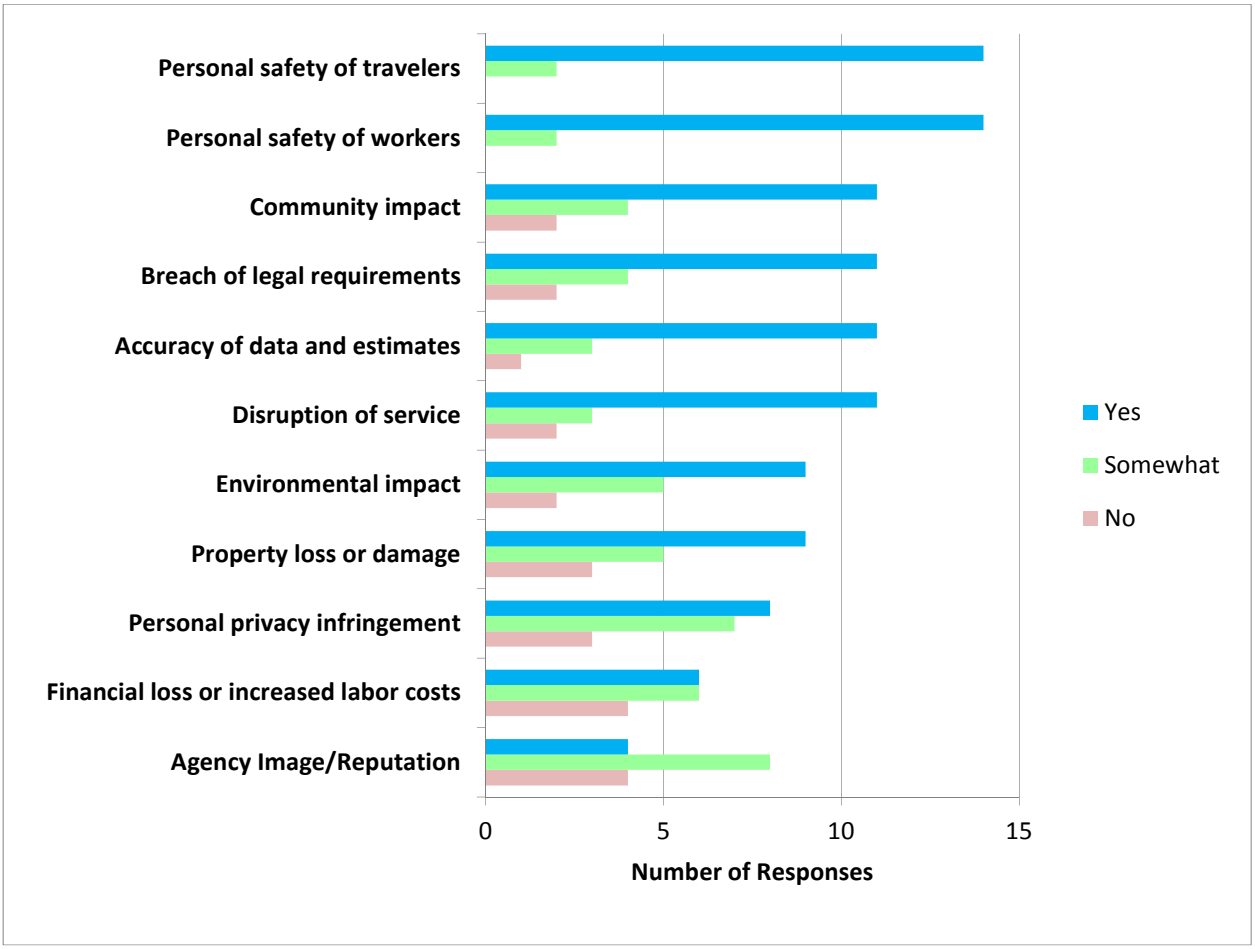


Figure 9. The Role of Risk and Reliability in Agency Maintenance Programs

4.3.7. Tools Recommended by DOTs

An agency's ability to define maintenance measures and prioritize targets in the manner desired will depend on the selection of tools, management structures, management systems, data management systems, or analytic tools. Figure 10 lists categories of tools available to highway maintenance agencies and their relative importance to the maintenance activities of interviewed agencies. Formal performance, personnel, and data management tools are the most important. Tools that provide customer feedback and executive information tools are generally valued as the least important but still used somewhat by most agencies.

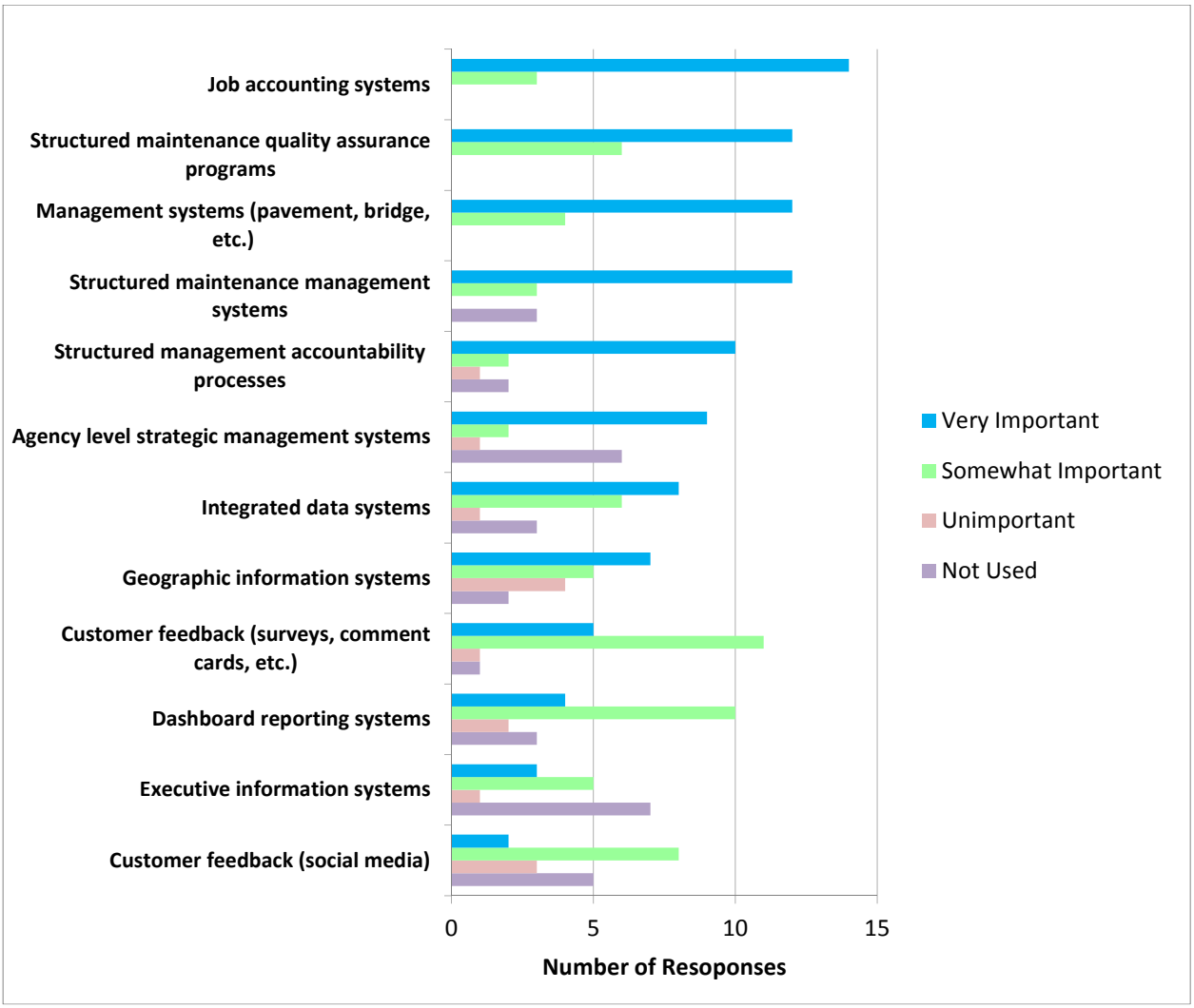


Figure 10. Relative Importance of Management Tools for Road Maintenance

As noted in earlier sections, specific tools and approaches include AVL systems, mobile data collection, GIS, MMS/Agile, PeopleSoft, and Dashboard systems, and various integrated business and financial planning tools. Colorado employs dedicated maintenance LOS coordinators. The degree to which all of these systems are integrated varies. Agency representatives emphasized the desirability of selecting tools, data measures, and systems capable of working well with one another, and preferably which serve more than one purpose. Full integration is more aspirational than realistic, but adopting this outlook can assist in making maintenance programs as cost-effective as possible.

5. Methods of Establishing Level of Service Targets for Commonly Used Highway Maintenance Features and Attributes

The public sector is under intense pressure to improve its operations and deliver its products and services more efficiently and at the least cost to the taxpayer. Performance measurement is a useful tool in this regard, since it formalizes the process of tracking progress toward established goals and provides objective justifications for organizational and management decisions. Thus, performance measurement can help improve the quality and cost of government activities as well as determine whether they are fulfilling their vision and meeting their strategic goals. The team identified central considerations for methods of establishing level of service (LOS) targets; to establish plausible and realistic targets, an exploration of the performance management process should give serious consideration to several dimensions that characterize every target setting process. The team also presented five alternative LOS target-setting frameworks.

As Figure 11 illustrates, a single strategic goal, *reduce crashes*, may have more than one strategy. Each strategy is supported by at least one action. Each action produces measurable outputs. Each output requires some inputs. Inputs can be measured by input measures—how many dollars, staff hours, etc. Each output can be measured be output measures—miles of shoulder blading, etc. The combination of input and output measures yields an understanding of efficiency.

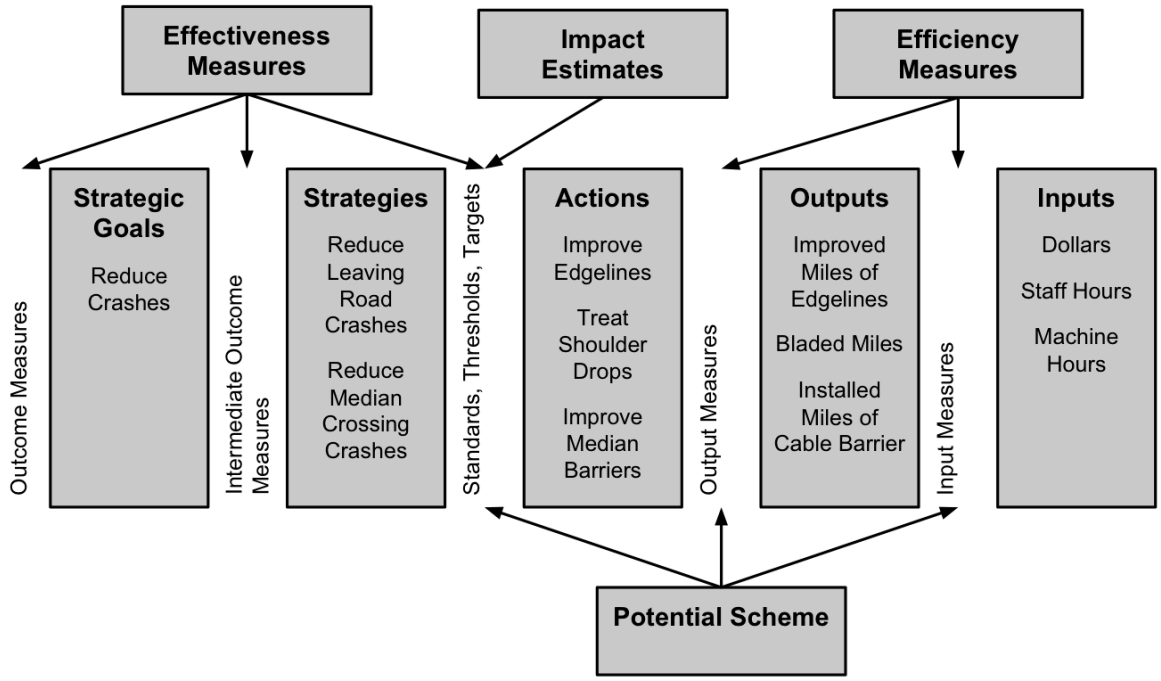


Figure 11. Strategic Context for Maintenance LOS Target Setting

All three of these concepts are important management and diagnostic tools to be used in monitoring the program. Actions are defined and measured by standards, what is acceptable or unacceptable; thresholds, how widespread are those acceptable or unacceptable conditions; and targets, what the new threshold should be. Intermediate outcome measures define whether the strategies have produced the desired results—reduced leaving-the-road and crossing-the-median crashes. Outcome measures define whether the strategic goal of reducing crashes has been attained. Effectiveness is defined by the outcome, intermediate outcome, and threshold measures. Impact estimates should be made for each of the proposed actions, which, combined with input

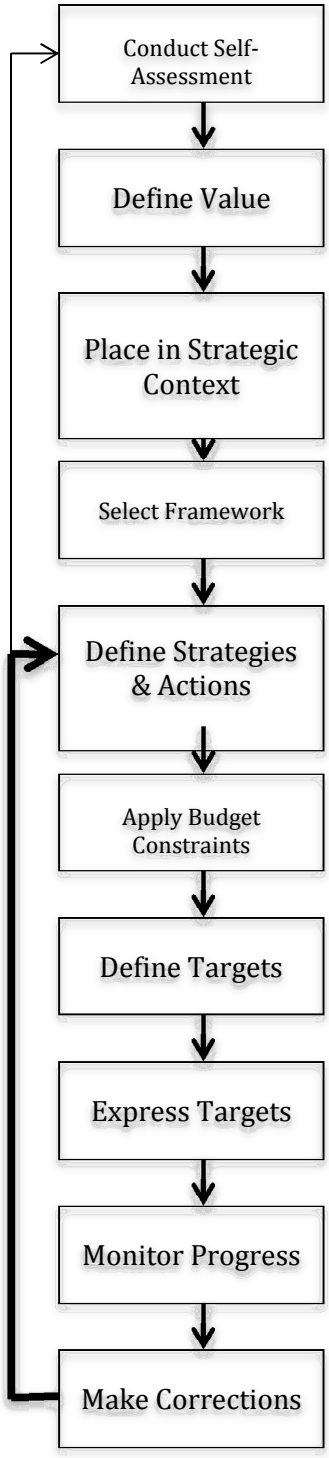


Figure 12. A Linear View of Goal Setting

estimates, will produce an understanding of the cost-benefit relationship. Finally, the combination of strategies, actions, and outputs might constitute a scheme that could be evaluated for its relative cost and benefit. These relationships must be understood when targets are defined because targets may be involved in five different locations. In the location that is of the greatest interest to maintenance managers, targets and standards interact in a manner such that changing either can have the effect of changing the program outcome. Reduce the standard and less of the system will be found acceptable. Increase the target and more of the system will require attention.

The process of target setting is not really top-down or bottom-up, because it involves reconciling strategic requirements with the reality of available resources and capabilities. Ultimately, a public entity such as a state transportation agency is held accountable by the taxpayers. The target setting targets process must be understood and coordinated at many levels, all of which can influence the outcome. As a starting point, the major steps of the process can be conceptualized in a linear fashion as shown in Figure 12. The steps follow from a rational target setting and management process. These are the basis for preparing a guide for LOS target setting. This document discusses each of the ten steps and the key factors that influence maintenance target setting.

5.1. Conduct a Self-Assessment

Each transportation agency is unique in some manner. A manager, whether the agency head or a maintenance chief, who would establish a management system using measures and targets must do so in a manner that reflects the reality of the agency, the resources that it possesses, its culture, and the environment within which it exists. Conducting a self-assessment may suggest a process that is more formal than necessary, but the questions are useful nonetheless to explain what the agency is willing to do or is capable of doing.

Will senior leadership champion the effort? An agency must have a culture of interacting across organizational units and of trust of its leaders. Just as clear, consistent, and visible involvement by senior executives and managers is a necessary part of successful performance measurement and management systems, senior leadership should be actively involved in both the creation and implementation of the target setting process. Managers should understand the challenges in contradicting an aspect of the agency culture and be prepared to take steps to counter the resistance when undertaking new, worthwhile approaches.

Is there a sense of urgency? Is there a business imperative, mandate, or impetus to move or move more aggressively to enhance the maintenance performance measurement system? Everyone involved should commonly aware of the driving motivation for the

target-setting initiative.

Does the agency have a maintenance performance measurement and management system? Successful target setting requires the agency have in place a clear and cohesive maintenance performance measurement framework that is understood by all levels of the organization and that supports objectives and the collection of results. The performance measurement system must have a uniform and well-understood structure, procedures for interacting across organization levels, and a clear calendar of events for what is expected from each organizational level and when.

Do the maintenance performance measures align with the agency's strategic direction and can they reveal trends? The established performance measures should enable employees and managers to understand and work toward the desired maintenance goals. The targeted measures should reflect whether the agency is meeting customer expectations. For this reason, executive managers will prefer targeted measures that reveal trends rather than day-to-day details. If a measure and its corresponding data requirements cannot be linked back to strategic direction, they should be considered for elimination from the target-setting process. This frees the agency from unnecessary data collection and maintenance initiatives in areas that produce little value.

What data resources exist? Measures are only as good as the data that support them. If key pieces of data are not available or cannot be modeled or merged with other data in a way to produce useful information, the shortfall will have to be overcome or the items measured will have to change.

Can the agency establish a baseline for each of the measures selected? The baseline value for targeted measures should be stable and consistent with repeated measurements. Measures that fluctuate over a wide range of values within short periods of time are not suitable for target setting.

How sophisticated are the analytic tools of the agency? Integrated data systems, GIS, and statistical analysis programs can greatly expand the range of measures that can be used and the manner in which they are displayed.

What staff resources are available? New data may have to be collected, stored, and analyzed. Existing data may be analyzed in different ways. Are the staff time and skill sets needed to do these available?

How media savvy is the agency? This question points to two different sets of issues. First of all, does the agency use the full range of tools to collect information from customers? Can it use social media tools or electronic surveys? Can it conduct virtual public meetings? Can it collect comments and complaints in a systematic manner and mine the data that are produced? While not all of these tools are needed for a successful program, their availability greatly increases the range of options available. The second issue deals with how well the agency interacts with news outlets. If staff is comfortable with those outlets and have a solid relationship with them, the process of establishing some accountability and transparency will be greatly enhanced.

Agency environment refers to the circumstances and conditions that interact with and affect an organization. These can include economic, political, cultural, and physical conditions inside or outside of the organization. Fiscal climate may dictate a reduction in funding and, therefore, a need for closer management of agency programs. Similarly, a high distrust of government or legislative interests may suggest the need for greater transparency. Finally, different administrations, both at the agency level and the executive level have varying degrees of openness to the free flow of information. The measures and framework should be sensitive to these contextual factors and respond to them.

5.2. Define Value

Inevitable limits on resources imply that some things must be valued more than others. The result of this step in the target setting process is to make the tough decisions as to how the targets will line up across various measures. Agencies must make choices that prioritize some assets, facilities, or services over others. Deciding what to prioritize as more or less important is a difficult and critical undertaking. Failure to align maintenance targets with the priorities of the stakeholders will have major impact on the credibility of the measurement process. One strategy is to recognize a balanced scorecard of performance measures for target setting. The scorecard measures performance across several perspectives: financial, customers, internal business processes, and learning and growth.

A survey of transportation agencies revealed that five basic constituent groups provide input for defining value. These groups are highly or moderately used by over half of the 40 agencies surveyed (Table 7).

Table 7. Stakeholder Groups that Contribute to Determining What is Valued Enough to be Measured by Transportation Agencies

Stakeholder group	Not Used	Little Used	Moderately Used	Highly Used
Executive, agency head or gubernatorial	0%	15%	28%	58%
Legislators, state or provincial	13%	20%	45%	23%
Employee	3%	20%	50%	28%
Industry experts and practitioners (contractors, consultants, transit operators, etc.)	13%	23%	50%	15%
Customers (motorists)	16%	24%	53%	8%
General public (motorists and non-motorists)	23%	33%	35%	10%
Local governments	23%	43%	28%	8%
Partner, financial partners P3 application	49%	28%	21%	3%

The tools used for identifying priorities are also diverse. Table 8 shows that public meetings and surveys are the primary mechanisms for more than half of the 40 agencies surveyed. Data used to measure customer satisfaction, dissatisfaction, or indifference are usually collected via customer surveys administered by a third party or an in-house office.

Table 8. Tools Used to Gather Input from Stakeholders

Tool	Not Used	Little Used	Moderately Used	Highly Used
Public meetings	8%	10%	53%	30%
Surveys, electronic, paper, telephone	5%	23%	55%	18%
Focus groups	20%	35%	38%	8%
Comment cards	23%	33%	43%	3%
Social Media tools or electronic communities	18%	45%	28%	10%
Telephone Hotline	23%	43%	20%	15%

5.3. Place in Strategic Context

An agency's highway maintenance function shares goals, strategies, and even resources with other functional areas within a transportation agency. Therefore, actions related to measures and targets must be placed within the context of the overall agency. Three issues are significant. The first is the agency's strategic plan. Not all agencies have formal plans, but most have some notion of which are the most important of their many programs and goals.

Ideally, an agency will have some measures and targets for the goals highlighted in its strategic plan. Maintenance measures and targets should be defined in relation to these agency goals, and they should share analytic tools, data sources, and reporting structures. Not all of the maintenance measures may be reported at the agency level, but some summarized version should be reported to define how maintenance activities contribute to agency goals.

Also ideally, the agency would have taken the steps necessary for internal alignment so that all agency staff have some understanding of its strategic direction and of how their efforts contribute to those goals. In reality, agencies find the connection between maintenance targets and strategic direction to take many forms. The 18 states interviewed answered this question: How would you assess the relationship between measures used in maintenance with the strategic direction and goals of the agency?

5.4. Select Framework

Measures and targets can be conceptualized and presented in many ways. There are some common target setting philosophies that should be considered when selecting a framework.

Stretch. This practice is to set targets that will force the agency to stretch to exceed past performance. By benchmarking, the agency can be aware of performance among peers to ensure the target is attainable. Setting impossible targets can undermine employee morale. An agency following the stretch philosophy should set attainable goals that motivate employees. If the target is not met then corrective action is implemented. Conversely, if targets are exceeded, the target can be reset to continue the stretch.

Empowerment. Responsibility goes with authority resulting in accountability. Maintenance workers and managers are likely to meet or exceed performance targets when they are empowered with authority to make decisions and solve problems related to the results for which they are accountable. This philosophy engages those who are responsible for achieving the target in a negotiation process for setting the target.

Cross-agency consistency. This increases employee understanding of the agency's maintenance mission and goals and unifies the workforce behind them. The targets are set and monitored as a way to communicate and regulate consistency across regions, counties, or districts. The targets may be used to identify opportunities for reengineering and resource allocation.

Accountability and transparency. For the agency as a whole, the essential top-level responsibility is to provide services to citizens and businesses. In this philosophy, target setting is used to conduct a gap analysis and manage expectations. The targets portray the facts about annual objectives, year-to-date performance, and the relationship between performance and resource allocation.

Continuous improvement. In this case, the targets are markers for recognizing when corrective action is necessary. Whether applied for longer term or for short-term corrective actions, the target setting process is the basis for creating a learning organization. A popular continuous improvement model is the Deming cycle (Vector Study, 2008): plan, do, study, and act. The *plan* is what is expected to happen for any selected action. The *do* is the execution of what was planned. The *study*

compares the result of what actually happened to the expected results. The *act* is taking action according to the results.

The different framework approaches each has its positive and less positive attributes. These frameworks are not always mutually exclusive. They can be and are often combined, but the choice of framework will tend to influence how measures can be used and how targets are selected.

Pass-fail is a widely used method of evaluating condition. A criterion or standard is set and it is either met or not met. The pass-fail approach notes only the extent of a condition, not its severity. For example, in the extreme, if the standard for ditches allows a passing score if it is less than 50 percent obstructed, a perfectly unobstructed facility would be reported as the same as one that is just less than half obstructed.

Level of Service (LOS) is a widely used approach. It is essentially a capacity-based construct that has been adapted to describe other service characteristics. NCHRP Project 20-07-206 (Zimmerman & Stivers, 2007) illustrates how two states have defined a pass-fail performance threshold: greater than 50 percent of the ditch filled with sediment. This pass-fail grading is then transformed into a level of service rating based on the percent of samples that pass the standard. If 2 percent or less of the samples on a route or within a geographic area fail this standard, that route or geographic area is said to have a LOS rating of A. The LOS approach has two major drawbacks: 1) In the extreme, a LOS rating of A could be attached to a system where all sampled ditches are just under half full of sediment (speaking to an adequacy of condition at a given moment in time, but not necessarily to how quickly that situation might change); 2) one of the segments sampled could contain most or all of the failures, making it in very poor condition, while the larger sample of segments would have a more positive rating.

A more robust application of the LOS differs from the previous application in that it requires the raters to rate each asset in the entire inventory according to the LOS scale. NCHRP Report 677 (Dye Management et al., 2010), which deals with standards for the Interstate system, suggests this approach. It works best when applied to a system in which 100 percent of the system is sampled, which is more likely with the Interstate. It is frequently used when dealing with bridges, using the National Bridge Inventory System (NBIS). This has the advantage of providing more detailed information for the maintenance manager on the condition of specific segments of the system. It also would have the advantage of unmasking some of the potential deviations of the pass-fail ratings. It has the potentially serious drawback of introducing more judgment into the condition rating process. It also makes the roll-up to route, system, or region ratings more complex when the sample is less than 100 percent.

Trend lines are also widely used. Advocates for this approach argue that it produces a greater incentive for improvement since the trend can always be improved, while a specific target can also be seen as a cap. However, once the target is reached, no incentive exists to strive for further improvement. Trend lines are also easy to read and understand.

Trend lines have a major drawback: In a time of scarce resources, trend lines will not assist an agency to allocate its resources consistently with its priorities. Even within the example above, if resources are tight, the agency might want to sacrifice some convenience and cleanliness for more safety. Since trend line imply targets that value all things equally, they would be difficult to use to bring about that change.

Benchmarks are another often used approach. An organization compares itself to another peer organization. This is often easiest to do within an organization.

Even in this internal comparison, one could question whether the regions are really peers. Traffic volumes, as well as terrain and soil types, vary widely. Even the values of the residents of the region

may vary. With these variations, the fact that northwest Wisconsin has only a 2 percent backlog on hazardous debris may or may not be a good thing when compared to the 12 percent backlog in southeast Wisconsin.

Another benchmarking approach is to find other agencies that might be considered comparable. Finding truly comparable agencies with comparable data collection and evaluation processes, as well as timely information, is a challenge.

Tiered approaches define more than one standard or set of standards. For example, gap analyses and tradeoffs may be defined to provide contrasts to those that are attainable. If the gap is significant, decision makers may be spurred to change policies or to increase funding.

System type is often used as a basis for defining different tiers of goals. Typically, the Interstate system, the National Highway System or the primary system is expected to be in a higher condition and service level than minor systems.

Short- and long-term is another distinction that can be made in a tiered framework. Trend lines and resources allow for the improvement of the system over time. An asset may have an LOS rating of D now, but is expected to improve to B over five years.

5.5. Define Strategies and Actions to Achieve Goals

This step in the target setting process draws from the agency information resources to define strategies for achieving its goals. In other words, with an understanding of agency goals, what strategies or actions can be taken in maintenance to further those goals? For example, reducing crashes is often a strategic goal of an agency. A review of a safety management or information system might demonstrate that a significant number of crashes are the result of vehicles leaving the road. Condition information and safety management might suggest two contributing factors: low shoulders and worn edge lines. The expert judgment of staff could suggest actions that could be taken to improve the quality of each of these features. Historic data or more formal cost-estimating tools could assign a cost to each action that might be taken. Agency information systems and expert judgment should also allow an estimate of the impact of each action to be made. Such information will be key for determining what, if any, actions can be taken by maintenance forces to further the agency goal.

5.6. Apply Budget Constraints

This step in the process is to recognize fiscal realities. Typically, funding is established by a legislative body or in some cases by a commission. Within those externally imposed constraints, the agency usually has some flexibility to make allocations between regions, programs, or elements within programs. Making those final agency allocations, which is where final targets are usually set, involves three key elements.

Fiscal risk is a real consideration in most of the states interviewed. If transportation revenues are less than planned, the budget may be subject to reduction. If unplanned costs are great as a result of winter weather or tropical storms, those costs may force a reallocation. If bid prices are higher than expected, the planned work may be reduced. The list of potential risks to funding levels is much longer. All of them must be considered and some provisions made for reversals. This is usually not an elaborate process, but rather an understanding of what can be deferred if the need arises.

Cost/service models vary greatly among states. Some states have sophisticated models that look at past costs and outcomes along with current prices to predict the costs and impacts of alternative investment schemes. Others rely more heavily on the experience and judgment of staff.

Program trade-offs is the process of considering the funding available, the strategies that might be employed (as defined in the previous step), and the impact of reallocations on other activities or other organizations (one region versus another) to make final allocations. Except in increasingly rare circumstances, when added funding is available, these decisions are made in a zero-sum environment. If one area increases, another—or many others—must contract. The impact of these decisions must be understood. The tools available to assist in this understanding include management systems, cost/service models and the expert judgment of staff.

The result of all of these efforts should be a program or operating budget for the maintenance program.

5.7. Define Targets

In this step, the agency has the framework, values, strategic actions, and budget allocations for setting attainable goals. Here the actions include a final management assessment of the budget plan and associated goals to ensure that risks have been adequately addressed and professional standards met. The targets can then be articulated in a manner that has meaning to those inside the agency who will implement the program and those outside who may want to understand the actions the agency is taking.

Another type of target must also be considered. More than 70 percent of those who responded to our survey said that they define desirable, but not necessarily attainable, targets in some manner (Table 9).

Table 9. How Agencies Set Performance Targets

Type of Target	Responses	%
Both attainable and aspired targets are defined to illustrate the shortcomings of current budgets	15	38%
Targets are defined that illustrate what would be a desirable outcome, without regard to budget	14	36%
Targets are selected that can be attained within available resources	5	13%
Targets are selected based on historic trends and may not reflect current budget reality	4	10%
Other	1	3%
Our agency does not set targets	0	0%
Total	39	100%

The state interviews also confirmed the use of aspirational goals, with eight of the 18 saying they used or somewhat used such goals. Aspired targets are defined using the best professional information available on the maintenance needs of the asset and how to attain it to achieve best long-term value. This is supplemented by the best understanding of what the public really values, as described above. Both are important. For example, if surveys or other tools have determined that the public ranks smooth pavements very highly, that conclusion should be further evaluated to determine exactly what the public understands to be smooth. The chances are very good that they will accept a lower standard than the typical professional measures would define. Assuming that the two are really the same could cause an overly high target to be set. At the same time, the public's concern with smooth pavement may not comprehend the benefit of preventive maintenance. Again, relying only on the single source might cause preservation to be undervalued.

5.8. Express Targets

In this step of the process, the agency communicates targets in a manner appropriate to the audience. The message may be different depending on the purpose and the perspective, but the messages should be consistent. Professionals within the agency expect to work with measures defined in technical terms that have meaning to them and speak to asset preservation or failure. The agency may introduce a new internal document on key performance indicators that uses color-coded graphs or markers to show progress or trouble.

A picture is worth a thousand words. Performance data can be displayed in a wide variety of ways, including graphic presentations such as histograms, bar charts, pie charts, and scatter diagrams. Run charts can be used to identify meaningful trends. The Pareto chart is another useful tool to show the relative importance of different categories.

Most agencies use some form of spreadsheet to organize and categorize their reporting data. Spreadsheets have useful analytical tools. A strategy for managing uncertainty and routine variations is statistical control process (SPC). The most common tool is the control chart, which is used to detect differences in variation among numerical results. The control chart filters out routine variation so that exceptional or problem situations can be revealed.

The information should not only be shared internally, but also externally with customers and stakeholders. Policy makers or members of the general public tend to expect simple or roll-up metrics such as: excellent, fair, or poor. The key, as targets are expressed in different ways, is to maintain a clear understanding of how alternative expressions relate to the basic measure. If that relationship becomes weak, the system will lose integrity.

5.9. Monitor Progress

An important aspect of performance measurement is it iterative quality. Agencies should continually assess whether their current targets are sufficient or excessive, are proving to be useful, and are driving the agency toward the desirable outcomes. Monitoring progress should involve three separate but related questions:

Has the goal been attained or was the purpose achieved? This should be considered at more than one level. Using the example from above of reducing crashes by focusing on leaving-the-road crashes with strategies of improving pavement drop-offs or improving edge lines, the first question would be: Have crashes been reduced? The second: Have leaving-the-road crashes been reduced? In other words has the agency goal been attained and did the maintenance strategy work as anticipated?

For example, if the established target is to increase from 85 percent to 95 percent the miles of roadway meeting the criterion of having a shoulder drop off of less than 1.5 inch, was the result accomplished? This may also be estimated or reported as a percentage of needed preventive maintenance implemented related to outcome. For example, Washington State DOT found that by raising their amount of preventive maintenance performed by 25 percent for traffic signals, they could reduce signal failures by half, saving money in the maintenance process over and above the traveling public's time and money and other costs.

Was the strategy implemented? This is really an output measure. If the strategy was to improve pavement drop-offs by increasing the cycle of shoulder maintenance from *X* to *Y*, has that occurred?

Were the budget allocation and costs estimates correct? Additional funds would have been allocated to shoulder maintenance to bring about the increase required. Did the amount estimated

correspond to the work performed? Thus, some DOTs such as Washington State DOT have turned to *work accomplished* measures.

These three questions will provide a more complete view of what has occurred.

Agencies should be careful not to be burdened by monitoring progress for too many targets. Monitoring includes continuous and regular review of measures as they related to strategic goals. When targets or measures become obsolete, they should be discarded. Measures and targets should be dropped when they are no longer needed or if no change occurs after much attention.

5.10. Make Corrections

Corrections are typically made as budget allocations are developed for the following program or fiscal year. As those allocations are developed, the agency should evaluate how well the strategies employed actually worked. They should also consider any implementation issues that need to be addressed. Finally, they should use any new information to update their models and estimates to reduce future error and risk.

With this step the process circles back to the start. This typically involves returning to the step of defining strategies and actions, since it would involve correcting for previous errors in cost estimates or assumptions of impact. In some cases, it may involve returning to the first step for closer look at the approach taken, the value being pursued, and the strategic context of these values.

6. Conclusion from Phase One

After digesting a lengthy literature review, analyzing a survey of states and provinces, and grappling with many in-depth interviews, the research team was able to come to a number of conclusions.

- Setting targets is but one of the issues with which agency managers should be concerned. Of equal importance are issues related to gaining acceptance of the performance measurement effort, gaining external support and understanding, using measures as a management tool, collecting and using information reliably, and dealing with issues of risk.
- Agencies vary widely in their structure, their focus, and their procedures.
- A corollary finding to the above point is that interstate comparisons are difficult to make. Therefore, efforts to do cross state benchmarking or to define national maintenance performance measures that have meaning when two states are compared is difficult to impossible.
- Few agencies have done much to prioritize the various maintenance functions. Therefore, many have some difficulty in allocating resources in a manner that could be considered strategic. With some exceptions, those with some prioritization scheme have arrived at those priorities in a fairly informal manner.
- The tie between agency goals and maintenance goals and direction is usually informal. Few states have a formal mapping of program goals to agency strategic goals.
- Few agencies have structured their maintenance programs in a way that allows the great number of maintenance actions or activities to be rolled up to the significant few that might have greater meaning to policy makers.
- Targets tend to be established based on historic trends or what would be desirable. Few agencies have a strong tie between targets and available funding levels.
- Related to the above is the relative lack of sound cost information. Few agencies have a good grasp of the cost of the outputs that contribute to the accomplishment of program goals.
- Existing targets tend to be fairly static. They are reviewed at long intervals and often changed only in response to some external pressure or rule change. For example, one agency reported significant changes to their targets for sidewalks and walkways as a result of the Americans with Disabilities Act rules.
- The approach to risk management in most states is very informal, relying on the experience and judgment of experienced personnel. Little analysis is brought to bear on issues related to risk.
- While some agencies use measures and targets for reporting to policy makers, few seem to have been able to engage their policy makers in true discussions of the consequences of policy and budget decisions.
- Few states actually use measures for basic management functions: allocating resources, developing plans or monitoring performance.
- Little sound information is available on the impact of maintenance actions on outcomes, safety improvement, longer-lived pavements, etc. Therefore, maintenance programming tends to be focused on the very short term, usually one year.

6.1. Implications of Conclusions for the Guide

With the above conclusions it is clear that measures could be used more aggressively by many agencies. If they were used more aggressively they could have a greater impact on the direction, effectiveness and efficiency of the maintenance programs. They could also be more useful in gaining support and understanding from the general public and policymakers.

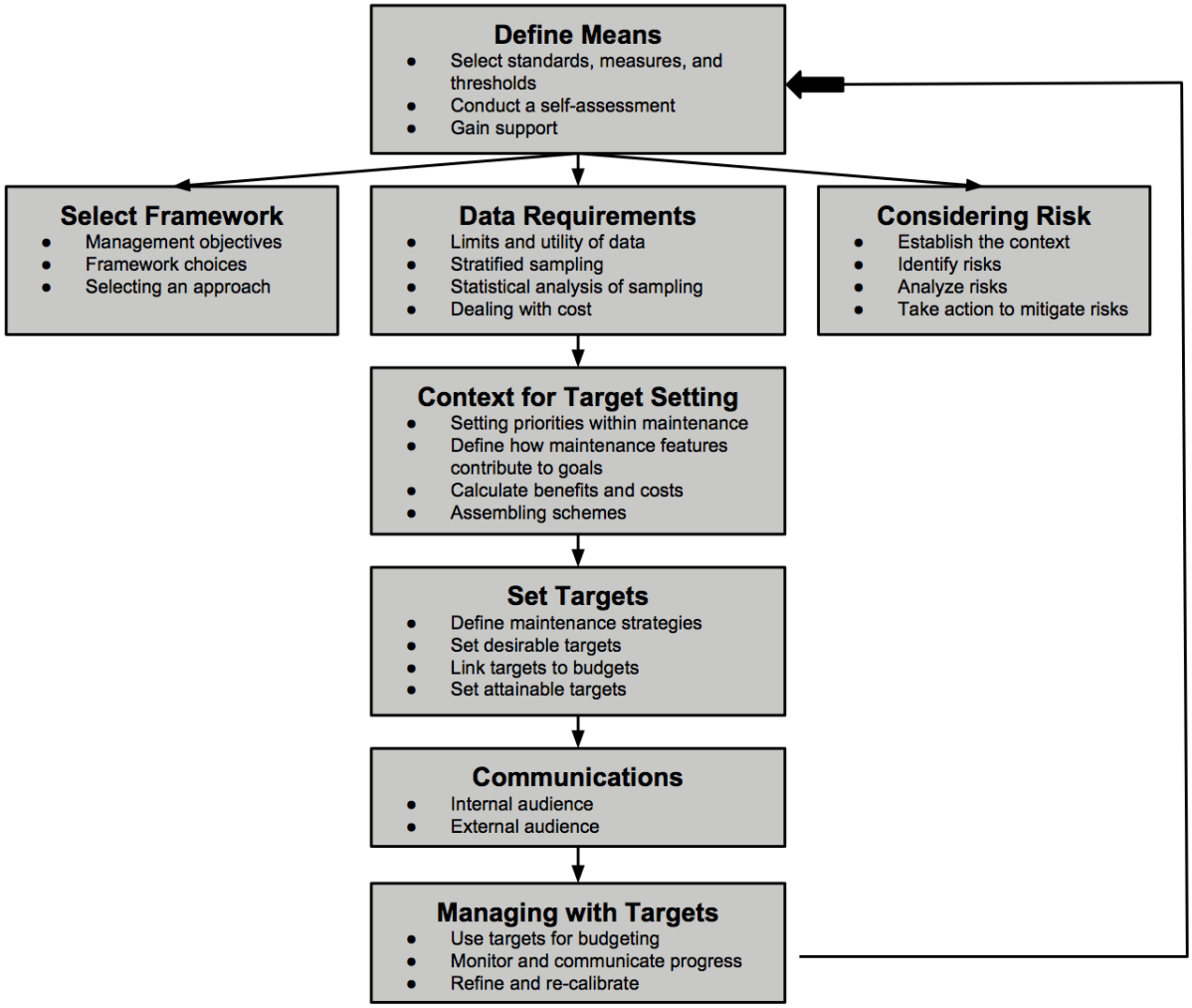


Figure 13. An Overview of the Maintenance Performance Management Process

With these conclusions and implications in mind, the Guide outlines direction for a number of topics, as shown in Figure 13.

- It begins with tools, including a self-assessment to help the manager and concerned staff to think about the agency, its resources, its culture, and its needs, to define the most appropriate approach to performance measurement. It offers a number of different approaches to how measures can be defined and portrayed and links those approaches to various management objectives and constraints within the agency. This allows the use of measures and targets to be tailored to the needs of specific agencies. It offers references and suggestions on how support can be gained for the effort within the agency. All of these things must be considered before any effort is made to implement performance measures or management or to use and existing performance measurement system more rigorously.

- Next it offers tools and approaches to dealing with data. How can data be sampled? How can it be summarized? How can costs be estimated? And many other topics. These are critical to the success of a measurement system and a critical need for many agencies.
- Then it offers ideas and tools for dealing with risk. These are straight-forward, common-sense tools that will be appropriate for this agencies that are just beginning to deal with risk in a systematic manner.
- The Analytical Hierarchy Process (AHP) tools are offered as an approach to systematically defining what is most important from the perspective of attaining the goals of the agency and the maintenance program. AHP allows the agency to determine the extent to which maintenance features contribute the attainment of specific maintenance goals. It provides methods for better understanding the relative cost of each feature in attaining goals. And it provides the input to the linear programming tool that can assist in defining attainable targets.
- The use of a hierarchy of features and categories and of maintenance and agency goals in the AHP illustrates both how maintenance goals can be tied to agency goals and how maintenance activities can be rolled up to a meaningful few for discussions with agency leaders and policy makers. Chapter 8 of this Report includes a very detailed discussion of the AHP method and how it evolved in the process of developing the Guide.
- The Guide dedicates an entire chapter to the topic of communicating targets. It offers a process for defining a communications plan, methods of analyzing the needs of targeted audiences and many examples from states on how targets and progress toward targets has been communicated.
- The Guide also dedicates a chapter to the process of setting targets. Chapter 9 of this report provides detail on how the thinking of the research team evolved on the process of actually setting targets with the support of the AHP tool.
- The Guide dedicates another entire chapter to management. It outlines procedures to managing with a plan. It details how that plan should contain input, output, intermediate outcome and outcome targets to allow the manager to understand the impacts of the actions taken, the efficiency with which those actions have been taken, the effectiveness of actions and basic task completion. All of these levels are necessary for effective management. It also provides process for monitoring and modifying plans as part of an overall management process.

Since all of the topics listed above are critical to the success of establishing targets and using performance measures, the Guide takes the broad view and outlines an overall process for establishing and rigorously using a maintenance performance management program.

7. Data Requirements

The method for LOS target setting makes use of the inventory, condition, and cost data sets that most agencies have assembled for their maintenance quality assurance programs. Some of the self-assessment questions explored the availability of the prerequisite data sets. At minimum the agency must gather the data listed in Table 10 for the assets to be included in the target-setting effort.

Table 10. Data Requirements for LOS Target Setting

Data Set	Description
Maintenance goals	Descriptions of the agencies maintenance goals and the relative importance of those goals. Ideally the maintenance goals relate to the agency's goals.
Inventory and condition assessment	Number of units of the highway features in total on the system and number of units not performing to the standard.
LOS definitions	LOS ranges for the highway features. The number of units of each feature not performing to the standard indicates the appropriate LOS.
Maintenance costs	Cost to perform maintenance activities expressed in units of work consistent with maintenance condition assessment: e.g., feet of edgelining, cost per face or square foot of sign replacement, acres of mowing, etc.
Effectiveness of maintenance activities	The expected outcome of maintenance activities and the relative contributions of maintenance activities for achieving the desired maintenance goals. The importance of edge lines for reducing crashes, mowed roadsides for reducing animal hits, better pavement markings for improving driver satisfaction, etc.

7.1. Understanding the Limits and Utility of Data

Agencies should have in place the office and field procedures for collecting and managing their maintenance cost and condition data. These procedures minimize data errors and inconsistencies. For example, data inconsistencies can be avoided by eliminating redundant data entries. Database design and development of user interfaces for data entry are basic information technology requirements that apply to any application and are therefore beyond the scope of this project.

Maintenance performance assessment programs use inventory data for some categories, such as pavement and signs. The inventory and condition databases that support the agency's asset management systems may also support maintenance performance assessment systems. Agencies commonly use inventory and condition databases for features such as pavements and bridges to assess maintenance performance.

Maintenance target setting requires reliable inventory and condition data of roadway features. Many agencies do not have the inventory and condition data for some features, such as drainage and shoulders. For these, it is necessary to survey roadways to collect the data required to assess maintenance quality. To minimize the workload, agencies try to identify a single key measure for each feature and design a random sampling strategy that will provides enough measurements for a valid analysis.

7.1.1. Using Inventory Databases

Virtually all states have 100 percent inventory and condition data for their pavements and bridges and these data are updated at least every two years. Many states inventory roadway pavement conditions from vehicle detection systems. All states follow federal guidelines for bridge inspection and inventory. Pavements and bridges are considered capital assets and are the major focus of the agencies asset management system. Many states are developing these data sets for culverts and signs.

When inventory and conditions databases are available, agencies tend to want to use many or all of the measures for maintenance performance assessment system. However setting targets for and reporting of all of the condition measures may lead to confusing and diffused messages to executives, the legislature, and the public. To reduce workload and confusion, agencies should identify and use the most pertinent, meaningful features for the audiences.

7.1.2. Quality Assurance of Data

Quality data–accurate, complete, and timely–is important at all steps of the target setting process. Complete data is free of missing elements and is uniformly representative of the population.

7.1.2.1. Managing Data Accuracy

Data accuracy indicates how well the data represents the true value. Accurate data is free of random and systematic measurement errors and biases. Data accuracy concerns are well justified if data is being collected by teams of human inspectors whether through the windshield or on foot along the roadside. Variations in how individual inspectors identify, measure, and record condition readings are virtually certain. Some common types of accuracy errors include:

- **Measurement error.** Most measurement errors may be attributed to human factors. The field data collection requires some physical agility, vigilance for safety precautions, and human interpretation. Quality assurance strategies should be established for training and monitoring the field inspection teams order to reduce and detect measurement bias.
- **Missing observations.** Segments can be excluded from the sample because they are deemed unsafe for measurement by the data collection teams. So there is good reason to believe that excluded segments may have higher deficiency rates than those included in the sample, and therefore that their exclusion results in sample bias.
- **Sampling methodology.** The exclusion of segments reviewed in the previous year's sample means the current year's sample is not completely random.

Most states recognize the likelihoods for inaccurate data collection. To combat the problem they have developed tools for their field inspectors such as pocket manuals with photographs showing various conditions states and pre-printed inspection forms.

An effective proactive strategy for assuring the accuracy of maintenance condition data collected in the field is to retrain the field inspectors annually. Wisconsin DOT requires all inspectors, new and returning, to attend a one-day training session with office and field components. Seasoned inspectors often serve as instructors. The annual training resets everyone's understanding of how condition measures are to be identified, measured, and recorded.

Data accuracy can be tested by using quality assurance (QA) tests. For maintenance condition data, the QA tests compare the field measurements collected by a Quality Assurance team to the measurements collected by the Field Review (FR) team for randomly selected highway segments. The tests look for measurement variations between the two teams. The results point to emphasis

areas for future training and modifications to the measuring techniques and/or deficiency thresholds. Additionally, data quality trends over time may indicate measures that should be deleted or changed because they simply cannot be reliably collected. The QA technique does not work well for hazardous debris, litter, mowing, and other vegetation control and other features if the condition varies between when the FR team and the QA team completed their ratings. A few questions that might be answered by the QA tests are the following.

1. Do the FR and QA teams agree on the observed existence of features?
2. When the teams agree that a feature exists, do the teams agree on the quantity of the feature?
3. When the teams agree that a feature exists, do the teams agree on whether or not that feature is deficient?
4. When the teams agree that a feature is deficient, do the teams agree on the magnitude (severity) of the deficiency?

7.1.3. Managing Data Completeness

Timely data is up-to-date and consistent with business cycles and maintenance cycles. Agencies know the number of centerline miles in the inventory so for features like shoulders, centerlines, and edge lines, the necessary sample size is easy to determine. Problems arise because some features are rare or the distribution of their maintenance condition is skewed (not normal). The number of features and sample bias cannot be dealt with simply by increasing the sample size.

7.2. Stratified Sampling for LOS Assessment

If the feature inventories and condition ratings are available, then sampling is not necessary. However, for many features, inventories are not available. For some features, the agency many have an inventory of the asset but condition data is not available or inadequate. For these cases, agencies must collect the LOS assessment information in the field.

To be successful, agencies must be able to design a stratified sampling plan including the sample size. Rather than surveying all features, an impractical and expensive effort, agencies can use probability and statistics to develop a sampling design. The sampling design is extremely important because it directly affects the precision of the confidence intervals and the ability to interpret whether an LOS target has been met or not. Two important considerations are stratified sampling and sample size.

DOTs are organized into jurisdictional areas, usually called districts or regions. These may be a subset of counties. No matter whether highway maintenance activities are centralized, decentralized, or by contracts, there is likely to be some variation across the jurisdictional areas. The variation could be due to local priorities of decision makers or due to geography, weather, and traffic. A stratified sampling design accounts for differences between jurisdictions so that one area does not disproportionally influence the statewide results.

Stratified sampling is a probability sampling technique wherein the analyst divides the entire state into areas, called *strata*, and then allocates a proportion of the total sample units to be randomly selected from each *stratum*. Each random observation is called a *sample unit***.** The collection of all the individual *sample units* is called the *sample*. The *sample size* is the total number of sample units.

Usually the sample units are 0.1 centerline miles and the number of sample units from each is proportional to the centerline miles in that area relative to the total centerline miles statewide. Table 11 shows a simple example for allocating eight sample units to the two regions

proportionally based on the number of centerline miles. The number of sample units allocated to each region must be rounded up or down to an integer value.

Table 11. Stratified Sampling Design Based on Proportion of Centerline Miles

Region	Centerline Miles	Percentage of total centerline miles	Computed sample size	Allocated Samples
Dane	350.00	62.8	5.02	5
Portage	207.64	37.2	2.98	3
Total	**557.64**	**1**	**8**	**8**

The number of sample units needed depends on the desired precision of the maintenance deficiency estimates. Unfortunately, precision for all estimates is difficult to achieve at an affordable cost. The highway features are not uniformly distributed across all areas, the deficiencies are measured in different ways, and some features occur much more frequently than others. A sampling rate of 2 to 5 percent is common for state DOTS. Due to budget constraints, some states go as low as one percent.

The sampling rate is calculated from Equation 1 where n is the number of sample units and N is the total number of possible sample units.

Equation 1. Sampling Rate

$$f = \frac{n}{N}$$

For the simple example in Table 11, the sampling rate is 0.143 percent. The unit of measure for the numerator and denominator should be the same. The sample unit is a 0.1 mile highway segment. The centerline miles must be converted to segments for computing the sampling rate.

$$f = \frac{n}{N} = \frac{8}{557.64\ (10)} = 0.00143$$

The number of sample units required for a sampling rate of 1 percent is 56, rounded up from 55.76.

$$n = fN = (0.01)557.64(10) = 55.76$$

The purpose of stratified sampling is to reduce the number of samples required. The effectiveness of stratified sampling will vary from feature to feature for the reasons mentioned above. There is a way to assess the effectiveness of the stratified sampling design. The design effect is a measure of how effectively the stratified design reduces variances compared to the simple random sampling designs. A design effect value of one indicates that the stratification had no effect on reducing the variance of the estimated deficiency rate. For most features, the design effect is less than one indicating improved sample efficiency (lower variance for the same sample size). Thus, simply stratifying the sample can be an easy, no-cost way to effectively increasing the sample size, as long as the stratifying variables have known population values.

7.3. Statistical Analysis of Stratified Sampling for Maintenance Performance Assessment

Three dominant geometries of features are apparent and there are patterns in the way certain geometries are measured. The geometry of features as they relate to roadway segments may be continuous linear, discontinuous linear, or point. These geometries are compatible with linear location referencing systems used for transportation asset management. The geometry of a feature

tends to prescribe the dimensions for its measure. Most states use measures that are parametrically related to the feature geometry.

1. A *continuous linear* feature is one that has a linear geometry and is continuous along a highway segment. Edgeline striping is an example of a continuous linear feature. The deficiency of a continuous feature is often expressed as per mile or as a percentage of the segments sampled. For many continuous features percent deficient is an adequate measure because the quantity of the feature is implied by the segment length.
2. A *discontinuous linear* feature is one that has a linear geometry, but does not appear continuously along a highway segment. Examples of discontinuous linear features are guardrails and concrete barriers. The deficiency of a discontinuous linear feature is expressed in terms of the length or linear feet of damage. For discontinuous linear features, it is best to record the quantity (area or length) the feature along with the quantity or percentage that is defective so that both the absolute (magnitude) and relative (percentage of total) deficiency are known. These values are important for calculating average deficiency over a region.
3. The final type of geometry identified is the *point*. The location of these features is associated with a point along the roadway length. These features are not uniformly distributed along a highway segment and follow no pattern or density. An example of a point feature is a sign. A culvert can be considered a point feature since its location can be associated with a specific position along the roadway length. The deficiency of a point feature is often expressed in terms of the total number deficient. For point features, it is best to record the quantity (number) of the feature along with the quantity or percentage that is defective so that both the absolute (magnitude) and relative (percentage of total) deficiency are known.

Because the maintenance condition of features is measured in different ways, and because some features are not present on every sample unit, the method for estimating the regional or statewide LOS performance varies. The estimation technique may also vary by what sort of auxiliary information is available. For example, variables that are related to the features can be used to adjust estimates. This report addresses three methods for estimating some common feature types. The estimator types are listed and described in Table 12 along with examples of possible common features. These methods are not applicable to all feature types. Furthermore, for specific features, there may be much better methods for estimation, even if one of those given below is valid.

The overall process for statistical analysis of the sampled data is shown in Figure 14. The outcome of the analysis is an estimate of the deficiency rate that can be used to estimate the LOS. The sampled condition data for each one of the agency's highway features is to be classified as one of three estimator types, which in turn indicates the appropriate statistical analysis procedure to follow for estimating the deficiency rate and confidence interval for the estimate. The final step is to create plots showing the estimated deficiency rates and confidence intervals along with the range of LOS grades that fall into the confidence interval.

The specific equation for each of the statistical analysis procedures is slightly different, but the overall process is the same. Figure 15 shows that overall process. First, use the sampled condition data for a single feature to estimate the deficiency rate and the confidence interval for that estimate. Next, test whether the confidence interval meets any predefined precision requirements set by the agency. If the precision requirements are not met, then the minimum required sample size is determined; this sample size is advisory for the next data collection effort. Whether or not the agency's predefined precision requirement is met, the confidence interval is plotted on the LOS grading scale for the feature. The analysis procedures include steps for assessing the sufficiency of

the sample size and if sufficient, whether the target has been met; otherwise, if insufficient, the steps for calculating the required sample size are included.

Table 12. Classification of Statistical Analysis Procedures for Evaluating Maintenance Condition Data from Stratified Sampling

Type	Estimator	Descriptions	Example Features
A	Simple binomial proportions	• The feature inventory size is *known*. • The frequency of features in the sample is *known* prior to sampling. • The feature occurs on all sampled segments. • Each sampled segment gets a pass/fail rating for the maintenance condition of the feature.	• Litter • Hazardous debris
B	Domain binomial proportions	• The feature inventory size is *unknown*. • The frequency of features in the sample is *unknown* prior to sampling. • Only segments that have the feature get rated. • Each of those segments gets a pass/fail rating for the maintenance condition of the feature. • Some features are rare and therefore sparsely represented in the sample.	• Paved shoulders • Unpaved shoulders • Centerline markings • Mowing for Vision
C	Ratios	• The feature inventory size is *unknown*. • The frequency or amount of features in the sample is *unknown* prior to sampling. • All sampled segments are rated. The total quantity of the feature and the quantity in deficient condition are measured on each segment. • If the feature is not present on the segment, then the total quantity and quantity in deficient condition are zero. Thus, we can generalize to say all sample segments are rated. • Some features are rare and therefore sparsely represented in the sample.	• Ditches (linear feet) • Fences (linear feet) • Special pavement markings (number of markers) • Protective barriers (linear feet)

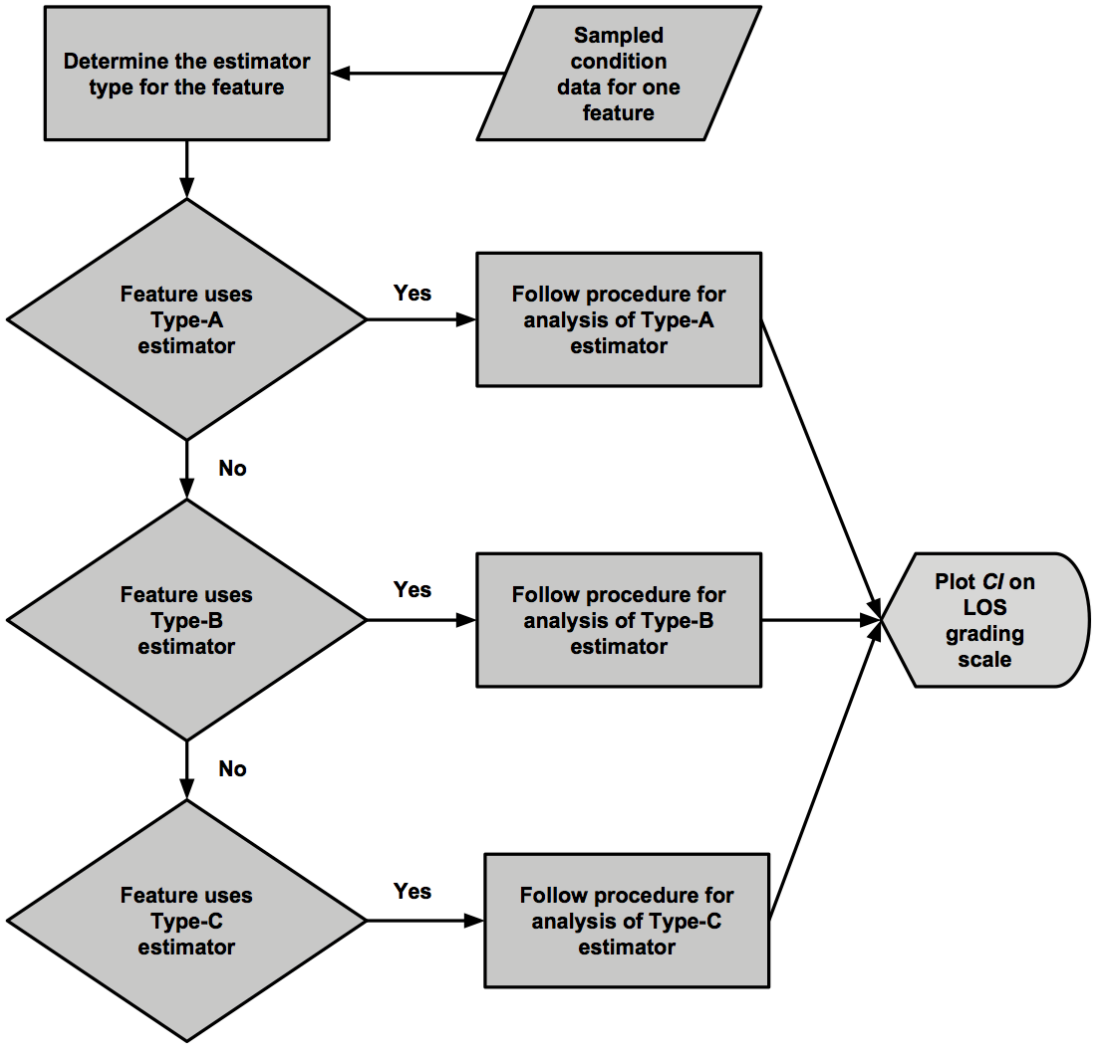


Figure 14. Process for Using the Statistical Analysis Procedures

The following standard notation will be used throughout this section.

Table 13. Notation Used in the Formulas for Statistical Analysis of Sampled Assessments of Maintenance Performance

Symbol	Definition	
i, I	Index of sample segments: $i \in I = \{1, 2, \dots, n\}$	
L	Number of strata	
h, H	Index of strata: $h \in H = \{1,2, \dots, L\}$	
n_h, n	Sample size (in segments) for stratum h, total sample size:	$n = \sum_{h=1}^{L} n_h$
N_h, N	Population size (in segments) for stratum h, total population size:	$N = \sum_{h=1}^{L} N_h$

Symbol	Definition	
f_h, f	Sampling rate for stratum h: $f_h = \frac{n_h}{N_h}$, total sampling rate: $f = n/N$	
$\mathcal{S}_h$	The subset of sample units in stratum h: $\mathcal{S}_h \subset I$	
$\bar{x}_h$ or $\hat{p}_{xh}$	The sample mean or sample proportion of the feature measure X in stratum h:	$\hat{p}_{xh} = \bar{x}_h = \frac{1}{n_h}\sum_{i \in \mathcal{S}_h} x_i$
$\bar{y}_h$ or $\hat{p}_{yh}$	The sample mean or sample proportion of the domain measure Y in stratum h:	$\hat{p}_{yh} = \bar{y}_h = \frac{1}{n_h}\sum_{i \in \mathcal{S}_h} y_i$
w_h	The population weight of stratum h:	$w_h = \frac{N_h}{N}$
s_{xh}^2	The sample variance of the feature measure X in stratum h:	$s_{xh}^2 = \left(\frac{1 - f_h}{n_h - 1}\right) n_h \hat{p}_{xh}(1 - \hat{p}_{xh})$
s_{xh}	The standard deviation of the feature measure X in stratum h:	$s_{xh} = \sqrt{s_{xh}^2}$
s_{yh}^2	The sample variance of the domain measure Y in stratum h:	$s_{yh}^2 = \left(\frac{1 - f_h}{n_h - 1}\right) n_h \hat{p}_{yh}(1 - \hat{p}_{yh})$
s_{yh}	The standard deviation of the domain measure Y in stratum h:	$s_{yh} = \sqrt{s_{yh}^2}$
s_{xyh}	The sample covariance of the feature measure X and domain measure Y in stratum h:	$s_{xyh} = \frac{(1 - f_h)1}{n_h - 1}\sum_{i \in \mathcal{S}_h} (x_i - \bar{x}_h)(y_i - \bar{y}_h)$

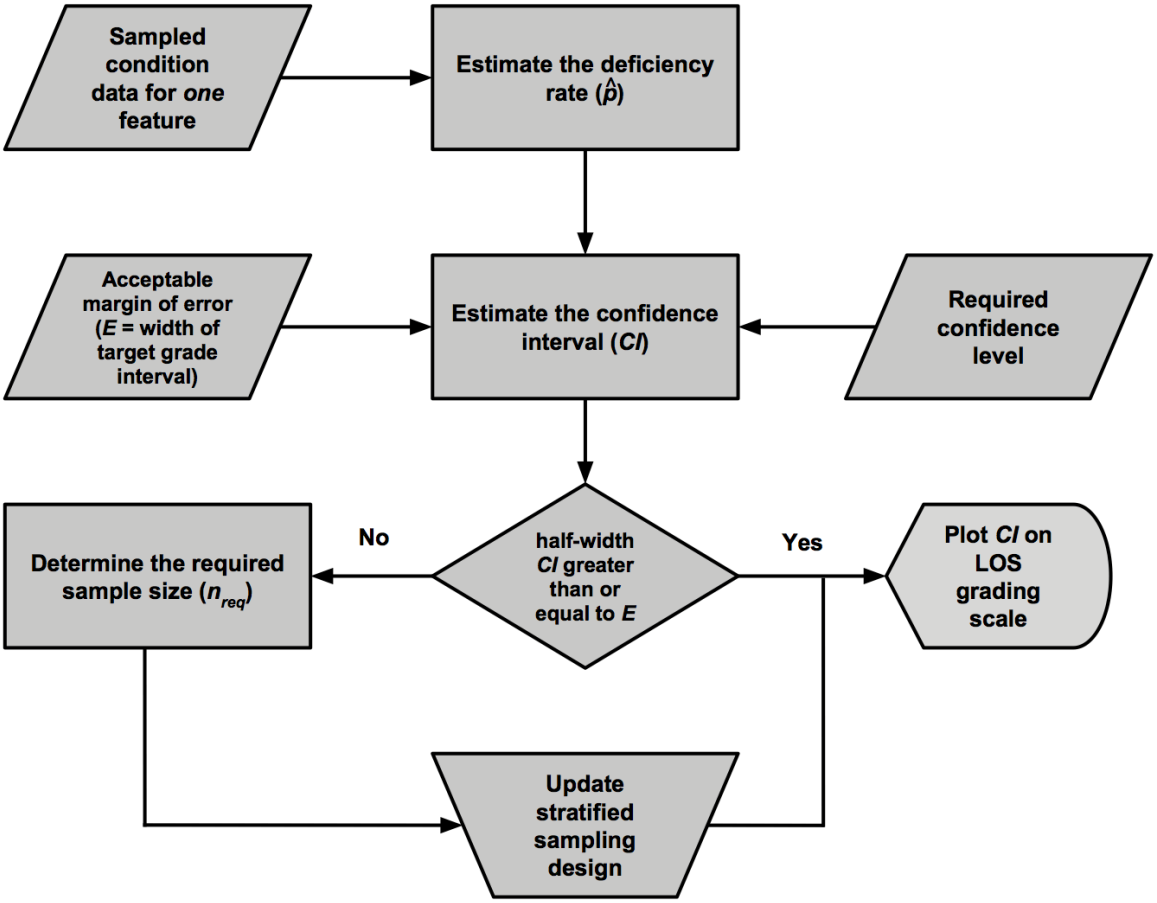


Figure 15. Process for Analyzing Confidence Level, Margin of Error, and Sample Size of Sampled Feature Condition

Examples for calculating the sample size and precision of the confidence interval for each type of estimator illustrate the methodology. A small dataset with two counties (or strata) and eight sample units (sample size = 8) is listed in Table 14. The example data table shows sample measures for one feature using each of three estimator types listed in Table 12. They are hazardous debris, drop-off/buildup (paved), and ditches, for estimator types A, B, and C, respectively. Note that this example is only used to demonstrate calculations. It is too small to meet any of the assumptions required by the different methods.

Table 14. Simple Data Set of Stratified Samples for Assessing Maintenance Performance

i	*h*	County	Hazardous Debris (Pass = 0, Fail = 1)	Paved Shoulder (Yes = 1, No = 0)	Drop-off/ Buildup (Paved shoulder) (Pass = 0, Fail = 1)	Linear Feet of Ditch	
						Total	In deficient Condition
1	1	Dane	0	1	1	1056	528
2	1	Dane	0	1	0	1056	450
3	1	Dane	1	1	0	923	250
4	1	Dane	0	1	1	0	0
5	1	Dane	0	0	0	330	0

i	h	County	Hazardous Debris (Pass = 0, Fail = 1)	Paved Shoulder (Yes = 1, No = 0)	Drop-off/ Buildup (Paved shoulder) (Pass = 0, Fail = 1)	Linear Feet of Ditch	
						Total	In deficient Condition
6	2	Portage	1	1	0	230	0
7	2	Portage	1	1	1	1056	100
8	2	Portage	0	1	0	760	0

Because a random sample is being used to estimate the maintenance performance, the analyst must use statistical tools to interpret results of the random sampling. The following are the basic metrics used for statistical analysis of sampled data for maintenance performance assessment.

1. An estimate of percentage of features that are deficient;
2. The variance and standard error of the estimate;
3. The 95 percent confidence interval for the true percent deficient;
4. The required sample size for desired confidence level and acceptable margin of error; and,
5. The effect of the stratified sampling design relative to simple random sampling.

7.3.1. Analysis of Features using Type-A Estimators

7.3.1.1. Deficiency Rate

The Type-A estimator is useful when the condition of the feature is measured on every sample segment and compared to a pass/fail criterion. For example, the number of pieces of hazardous debris on each segment is counted and if that value is not zero then the segment is deficient. The binary variable of interest, arbitrarily called x, then has a value of one if a segment is deficient and zero otherwise. The Type-A estimated statewide percent deficiency for the feature can then be calculated as the sum of the weighted mean of all strata,

Equation 2. Multi-stratum Percent of Deficient Segments for Type-A Estimators

$$\hat{p}^A = \sum_{h=1}^{L} w_h \hat{p}_{xh}$$

where $\hat{p}^A$ is an estimate of the true statewide percent deficient (referred to as p), w_h, the weight factor, is the percent of total centerline miles in stratum h, and $\hat{p}_{xh}$ is the average rating for all sample units from strata h. The superscript denotes the Type-A estimator. Using the dataset in Table 14, $L = 2$ counties and $N = 557.64 * 10 = 5{,}576.4$ 1/10th mile segments. For Dane County $n_1 = 5$ segments and $N_1 = 350 * 10 = 3{,}500$ segments, and thus $w_1 = \frac{3{,}500}{5{,}576.4} = 0.6276$. Using the observations for Dane County (hazardous debris column), the stratum sample mean is $\hat{p}_{x1} = 1/5$. Similarly, for Portage County, $n_2 = 3, N_2 = 2{,}076.4, w_2 = \frac{2{,}076.4}{5{,}576.4} = 0.3724$, and $\hat{p}_{x2} = 2/3$. Then from Equation 2, $\hat{p}^A = 0.6276 * 0.2 + 0.3724 * 0.6667 = 0.3738$. The result indicates that approximately 37 percent of the state's centerline miles are deficient due to hazardous debris.

Alternatively, $\hat{p}^A$ could be computed as the average of all observations from all strata without weighting. With this alternative form, $\hat{p}^A = \frac{3}{8} = 0.375$, which is slightly higher than the estimate using Equation 2. The alternative form gives weight to each stratum according to the number of samples units from that stratum not the portion of total centerline miles. Equation 2 requires more computations but is preferred because it can compensate for situations when the allocation of

sample units to each county requires some rounding, when sample units are missing, or for some reason, extra sample units are being used.

7.3.1.2. Standard Error

Because $\hat{p}^A$ is itself a random variable it can vary from sample to sample. One measure of this variability is the standard error in Equation 3. The standard error measures the accuracy with which the sample represents the population.

Equation 3. Standard Error for Type-A Estimators

$$StdErr(\hat{p}^A) = \sqrt{\sum_{h=1}^{L} \frac{{w_h}^2 \boldsymbol{s_{xh}^2}}{n_h}}$$

Where s_{xh}^2 is the sample variance defined in Table 13 and requires values for $f_h, \hat{p}_{xh}$, and n_h for each of the strata to be available. Both $\hat{p}_{xh}$ and n_h are given above so only the sampling rates f_h are needed. For Dane County, $f_1 = \frac{5}{3{,}500} = 0.00143$ and for Portage County, $f_2 = \frac{3}{2{,}076.4} = 0.00144$. Thus for the proportions in Table 14, the stratum sample variances are

$$s_{x1}^2 = \left(\frac{1 - 0.00143}{5 - 1}\right)(5)\left(\frac{1}{5}\right)\left(\frac{4}{5}\right) = 0.1997$$

$$s_{x2}^2 = \left(\frac{1 - 0.00144}{3 - 1}\right)(3)\left(\frac{2}{3}\right)\left(\frac{1}{3}\right) = 0.3329$$

The standard error of the estimate for the percentage of centerline miles that are deficient is about 18 percent. The standard error is the standard deviation of the estimator.

$$StdErr(\hat{p}^A) = \sqrt{\frac{(0.6276^2 * 0.1997)}{5} + \frac{(0.3724^2 * 0.3329)}{3}} = 0.1764$$

7.3.1.3. Design Effect

Standard error is then used to calculate design effect. The design effect for Type-A estimators can be computed using Equation 4.

Equation 4. Design effect of Type-A Estimators

$$DEFF^A = \left(\frac{StdErr(\hat{p}^A)}{StdErr_{srs}(\hat{p}^A)}\right)^2 = \frac{\left(StdErr(\hat{p}^A)\right)^2}{\frac{(1-f)\hat{p}^A(1-\hat{p}^A)}{n-1}}$$

At the state level $f = \frac{8}{5576.4}$ and $n = 8$. Furthermore, from Equation 2, $\hat{p}^A = 0.3304$. Thus, at the statewide level, the design effect for estimating the percentage of segment deficient due to hazardous debris is,

$$DEFF^A = \frac{0.1764^2}{\frac{\left(1 - \frac{8}{5{,}576.4}\right) * 0.3738(1 - 0.3738)}{8 - 1}} \approx 0.93$$

This means that the stratified sampling design has less variance than the simple random sampling. The stratified design has 93 percent of the variance, an approximately 7 percent reduction in the variance relative to simple random sampling.

7.3.1.4. Confidence Interval

The confidence intervals offer another measure of an estimator's variance by providing an interval that aims to contain the true backlog rate p with a pre-determined level of confidence. For example, a 95 percent confidence interval should include the true percent deficient for 95 out of 100 samples on average.

The statistical analysis of the Type-A estimator assumes that $\hat{p}^A$ is normally distributed as shown in Figure 16.

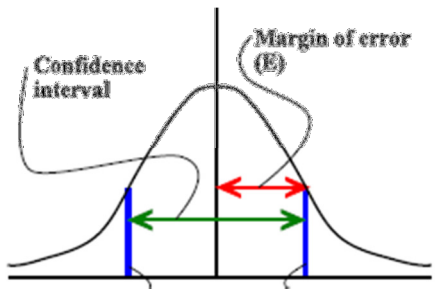


Figure 16. Confidence Interval and Margin of Error for a Normal Distribution

As a binomial proportion, the estimator $\hat{p}^A$ has several confidence interval estimation methods available. One approach is to approximate the distribution of $\hat{p}^A$ with a normal distribution and from that determine the confidence interval limits (also known as a Wald Interval). Using this method, a 95 percent confidence interval is calculated according to Equation 5

Equation 5. Wald Interval Equation for Estimating Confidence Intervals for Type-A Estimators

$$\hat{p}^A \pm z * StdErr(\hat{p}^A)$$

In Equation 5, the z critical value depends on the desired confidence level as listed in

Table 15. $\hat{p}^A$ is the computed estimate of the statewide percent deficient using Equation 2. For a 95 percent confidence interval, and the z critical value is 1.96. Thus, the center of the interval is $\hat{p}^A = 0.3738$ and the half-width is

$$z * StdErr(\hat{p}^A) = 1.96 * 0.1764 \approx 0.3458$$

giving a confidence interval of 0.3738 ± 0.3458, or roughly 3 percent to 72 percent. In some cases, the calculated lower bound of the interval will be negative. This does not make sense and is a result of the normal distribution sometimes being a poor approximation. While recommendations vary, one rule of thumb to determine if a normal approximation is reasonable is to ensure that the sample size and proportion are large enough to satisfy both $n > \frac{5}{p}$ and $n > \frac{5}{(1-p)}$. In the case of the hazardous debris feature in the example, this rule can be checked using the estimate $\hat{p}^A$: $\frac{5}{\hat{p}^A} \approx \frac{5}{0.3738} \approx 14$ and $\frac{5}{(1-0.3738)} \approx 8$. Since in the example $n = 8$, the normal assumptions are not valid.

Such rules exist because the sampling distribution for a binomial proportion converges to the normal distribution more slowly when the population proportion is small or large (close to zero or one, respectively). While confidence interval estimation may be poor in these circumstances, that may not be problematic. A small backlog rate is certainly desirable and is likely to have relatively low variance, even if it is underestimated in the confidence interval.

If accurate intervals for small proportions are desired, different estimation methods are available. For example, a Wilson Score interval with an adjustment to the effective sample size, based on

design effect, may be used for stratified samples. For larger sample sizes, however, this will give approximately the same results as a Wald interval.

Table 15. Z Critical Values for Desired Confidence Levels of Normal Distributions

Confidence level	*z* critical value
0.70	1.04
0.75	1.15
0.80	1.28
0.85	1.44
0.90	1.645
0.92	1.75
0.95	1.96
0.96	2.05
0.98	2.33
0.99	2.58

By making the confidence interval narrower with lower variability and higher sample size, it becomes more precise. Often high levels of confidence are associated with wide confidence intervals, while narrower (more precise) intervals have less confidence. As can be seen in the Wald interval equation (Equation 5), along with the equation for standard error (Equation 3), standard error gets smaller when *n* increases and therefore so does the interval half width. The interval half width is also called the margin of error and is a measure of precision of the estimate.

7.3.1.5. Sample Size

For maintenance performance assessment, the natural question is how large of a sample is necessary in order to obtain a 95 percent confidence interval with a specific margin of error. Equation 6, based on the Wald interval equation, provides an approximation of the required sample size for a specific confidence level, *z*, and margin of error, *E,* assuming that the samples are allocated to the strata with the same proportions,

Equation 6. Required Sample Size for Specified Confidence Level and Margin of Error when Using the Type-A Estimator

$$n_{req}^{A} = \frac{z^2}{E^2} * StdErr^2(\hat{p}^A) * n$$

Where *z* is the critical value for the desired confidence levels from

Table 15. For the hazardous debris feature, if the desired confidence level is 95 percent (z=1.96) and the allowable margin of error is 5 percent (E = 0.05), then the required number of sample segments is

$$n_{req}^{A} = \frac{1.96^2}{0.05^2} * 0.1764^2 * 8 = 383$$

This value should only be used if it is greater than that required by the normality assumptions, as described above. In this case, 383 is much larger than 14, and so should be used as the required

sample size. Another practical consideration is that there should be at least two samples in each stratum, in order for variance calculation to be possible.

7.3.2. Analysis of Features using Type-B Estimators

7.3.2.1. Deficiency Rate

The Type-B estimator is useful when the condition of the feature being assessed is measured on a subset of the sample segments. The difference between the Type-A and Type-B estimators is that for Type-B, the feature occurs on a subset of the sampled segments while for Type-A, the feature occurs on all sampled segments. The condition of the feature is compared to a pass/fail criterion for both Type-A and Type-B estimators. For example, the drop-off/buildup (unpaved) feature is measured only on segments with unpaved shoulders. For segments with unpaved shoulders, if there are at least 20 linear feet with more than 1.5 inches of drop-off or buildup between the unpaved shoulder and the road, the entire segment is deficient. For the Type-B estimator, the binary variable x receives a value of one if the segment is deficient, and zero otherwise. But because the number of segments with the feature is random, domain analysis must be used. In order to conduct the domain analysis, another variable y takes a value of one if a segment is in the domain (e.g., has an unpaved shoulder), and zero otherwise. Using both of these variables, Equation 7 is used to estimate the proportion of highway segments in the feature domain that are deficient.

Equation 7. Type-B Estimator for Proportion of Deficient Segments in Feature Domain

$$\hat{p}^B = \frac{\hat{p}_x}{\hat{p}_y} = \frac{\sum_{h=1}^{L} w_h \, \hat{p}_{xh}}{\sum_{h=1}^{L} w_h \, \hat{p}_{yh}}$$

The denominator ($\hat{p}_y$) is an estimate of the proportion of 0.1 mile segments in the population that contain the feature, while the numerator ($\hat{p}_x$) is an estimate of the proportion of segments that are deficient. Using the data from Table 14, the random variables y and x are the columns "paved shoulder" and "drop-off/buildup (paved)", respectively. For Dane County, $\hat{p}_{x1} = \frac{2}{5}$ and $\hat{p}_{y1} = \frac{4}{5}$. For Portage County $\hat{p}_{x2} = \frac{1}{3}$ and $\hat{p}_{y2} = 1$. From Table 11, the stratum weights, $w_h = \frac{N_h}{N}$, are $w_1 = 0.6276$ and $w_2 = 0.3724$. Then from Equation 7, the estimated proportion of segments in the feature domain that are deficient is 42.9 percent

$$\hat{p}^B = \frac{0.6276\left(\frac{2}{5}\right) + 0.3724\left(\frac{1}{3}\right)}{0.6276\left(\frac{4}{5}\right) + 0.3724(1)} = \frac{0.3752}{0.8745} = 0.4290$$

7.3.2.2. Standard Error

Under some circumstances, variance estimates for ratios, such as $\hat{p}^B$ can be very inaccurate and significantly underestimated. One reason is that the estimator is a non-linear function of two means. Because it is not possible to calculate the variance of such estimators directly, a linear approximation of the function is used instead and the variance of that linear approximation is calculated. The approximation works well when the function is approximately linear, but not when the denominator is small; for example, when a feature is rarely present. For rare features, the ratio may fluctuate wildly with different samples. However, if the denominator is not too close to zero and the sample size is large enough, the linearization method of variance estimation will be reasonable. Equation 8 is for estimating the linearized standard error for Type-B estimators.

Equation 8. Linearized Standard Error for Type-B Estimators

$$StdErr(\hat{p}^B) = \sqrt{\frac{1}{\hat{p}_y^2}\sum_{h=1}^{L} w_h^2\left(\frac{1}{n_h}\right)(s_{xh}^2 + s_{yh}^2(\hat{p}^B)^2 - 2\hat{p}^B s_{xyh})}$$

Equation 8 requires stratum variance and co-variance s_{xh}^2, s_{yh}^2, and s_{xyh} in addition to the number of sample units from each stratum, the stratum weights, and $\hat{p}_y$. Using equations in Table 12 and the same values of f_h as for Type-A estimator, the following values are calculated:

$$s_{x1}^2 = \frac{1-f_1}{n_1-1}n_1\hat{p}_{x1}(1-\hat{p}_{x1}) = \frac{1-0.00143}{5-1}(5)\left(\frac{2}{5}\right)\left(1-\frac{2}{5}\right) = 0.2996$$

$$s_{y1}^2 = \frac{1-f_1}{n_1-1}n_1\hat{p}_{y1}(1-\hat{p}_{y1}) = \frac{1-0.00143}{5-1}(5)\left(\frac{4}{5}\right)\left(1-\frac{4}{5}\right) = 0.1997$$

$$s_{x2}^2 = \frac{1-f_2}{n_2-1}n_2\hat{p}_{x2}(1-\hat{p}_{x2}) = \frac{1-0.00144}{3-1}(3)\left(\frac{1}{3}\right)\left(1-\frac{1}{3}\right) = 0.3329$$

$$s_{y2}^2 = \frac{1-f_2}{n_2-1}n_2\hat{p}_{y2}(1-\hat{p}_{y2}) = \frac{1-0.00144}{3-1}(3)(1)(1-1) = 0$$

$$s_{xy1} = \frac{1-f_1}{n_1-1}\sum_{i\in S_1}(x_i-\bar{x}_1)(y_i-\bar{y}_1) = \frac{1-0.00143}{5-1}\left[2\left(1-\frac{2}{5}\right)\left(1-\frac{4}{5}\right)+2\left(0-\frac{2}{5}\right)\left(1-\frac{4}{5}\right)+\left(0-\frac{2}{5}\right)\left(0-\frac{4}{5}\right)\right] = 0.0999$$

$$s_{xy2} = \frac{1-f_2}{n_2-1}\sum_{i\in S_2}(x_i-\bar{x}_2)(y_i-\bar{y}_2) = \frac{1-0.00144}{3-1}\left[2\left(0-\frac{1}{3}\right)(1-1)+\left(1-\frac{1}{3}\right)(1-1)\right] = 0$$

Substituting into Equation 8 gives

$$StdErr(\hat{p}^B) = \sqrt{\frac{1}{0.8745^2}[0.6278^2\left(\frac{1}{5}\right)\left(0.2996+0.1997(0.4290)^2-2(0.4290)(0.0999)\right) \quad 0.3724^2\left(\frac{1}{3}\right)(0.3329+0-0)]} = 0.2143$$

7.3.2.3. Design Effect

Standard error is then used to calculate the design effect. The design effect for Type-B estimators can be computed as for Type-A estimators using Equation 9. Note that n_d is used instead of n in the denominator. n_d is the number of sample segments in the domain of the feature being assessed.

Equation 9. Design Effect of Stratified Sampling Relative to Simple Random Sampling for Type-B Estimators

$$DEFF^B = \left(\frac{StdErr(\hat{p}^B)}{StdErr_{srs}(\hat{p}^B)}\right)^2 = \frac{\left(StdErr(\hat{p}^B)\right)^2}{\frac{(1-f)\hat{p}^B(1-\hat{p}^B)}{n_d-1}}$$

For the drop-off/buildup (paved) feature, $n_d = 7$ because 7 of the 8 sample segments have a paved shoulder, For $StdErr(\hat{p}^B) = 0.2143$ and from Equation 8, $\hat{p}^B = 0.4290$. $f = \frac{n}{N} = \frac{8}{5{,}576.4}$. Thus the design effect is,

$$DEFF^B = \frac{\left(StdErr(\hat{p}^B)\right)^2}{\frac{(1-f)\hat{p}^B(1-\hat{p}^B)}{n_d-1}} = \frac{0.2143^2}{\frac{\left(1-\frac{8}{5{,}576.4}\right)0.4290(1-0.4290)}{6}} = 1.13$$

The design effect for Type-B estimators captures the effects of both the stratified sampling and domain analysis. The two effects, stratification and domain analysis usually have opposing design effect. The domain analysis usually increases the variance of an estimate. In this example, the net result of the two effects is an increased design effect. The design effect of 1.13 indicates the

stratified and domain analysis design requires about 13 percent more than the random sample design to achieve the confidence level and margin of error of the simple random sampling.

7.3.2.4. Confidence Interval

As for Type-A features, the Wald interval equation (Equation 10) is used to estimate the confidence interval where z is the critical value for the desired confidence level from

Table 15. From the table, $z = 1.96$.

Equation 10. Wald Interval Equation for Estimating Confidence Intervals for Type-A and Type-B Estimators

$$\hat{p}^B \pm z * StdErr(\hat{p}^B)$$

For this example, the center point of the Wald confidence interval is again the estimate itself, $\hat{p}^B = 0.4290$ and the half-width of the confidence interval is

$$z * StdErr(\hat{p}^B) = 1.96 * .2143 \approx 0.426$$

Thus the interval is about 1 percent to 85 percent, meaning that there is a 95 percent chance that the true deficiency rate in the domain is between 1 and 85 percent. As for Type-A estimates, this interval is based on the assumption that the estimate $\hat{p}^B$ is approximately normally distributed. When the rule of thumb, $n > \frac{5}{p}$ and $n > \frac{5}{(1-p)}$, is satisfied, the normality assumption is reasonable. If not, however, alternative methods may be necessary for an accurate confidence interval. And again, the modified Wilson interval described for Type-A estimators may give better results.

7.3.2.5. Sample Size

Equation 11, based on the Wald interval equation, provides an approximation of the required sample size for a specific confidence level. z is the critical value for the confidence level from

Table 15 and E is the acceptable margin of error.

Equation 11. Required Sample Size for Specified Confidence Level and Margin of Error when Using the Type-A or Type-B Estimator

$$n_{req}^B = \left(\frac{z^2}{E^2}\right) StdErr^2(\hat{p}^B) * n$$

Notice that in Equation 11, n is the number of sample segments used to estimate the value of $StdErr^2(\hat{p}^B)$ while n_{req}^B is the total sample segments required, including domain and non-domain segments, to achieve the desired confidence level and acceptable margin of error. The number of sample segments required is,

$$n_{req}^B = \left(\frac{1.96}{0.05}\right)^2 0.2143^2(8) \approx 565$$

This value should again only be used if it is greater than that required by the normality assumptions, as described for Type-A estimates. In this case, 565 is much larger than $\frac{5}{\hat{p}^B} = \frac{5}{0.4290} \approx 12$ and $\frac{5}{1-\hat{p}^B} = \frac{5}{1-0.4290} \approx 9$, and so should be used as the required sample size.

7.3.3. Analysis of Features using Type-C Estimators

7.3.3.1. Deficiency Rate

Roadway features that use the Type-C estimators are assessed according to the portion the inventory that is deficient. The inventory unit for these features is not the highway segment,

although the highway segment serves as the primary sampling unit and determines the sample sizes. Instead, each segment contains a quantity of the feature that varies from segment to segment. For example, the ditches inventory is measured by linear feet of ditch, and each 0.1 mile segment may have 0 to 2,112 feet of ditches, assuming the ditches traverse both sides of a divided highway.

For Type-C estimators,

- y_i = the total quantity of the feature inventory on segment i.
- x_i = the quantity of the feature inventory that is deficient on segment i.
- Note that $0 \leq x_i \leq y_i$; the feature may not occur on segment i, or the total inventory of the feature on segment i may be deficient. In any case the deficient quantity cannot be greater than the total inventory quantity.

The last two columns in the example dataset in Table 14 are the linear feet of ditches and the linear feet of deficient ditches in each segment. These values are y_i and x_i, respectively. Equation 12 are for computing the stratum sample means.

Equation 12. Stratum Level Mean Inventory and Mean Inventory in Deficient Condition

$$\bar{y}_h = \frac{1}{n_h} \sum_{i \in S_h} y_i$$

$$\bar{x}_h = \frac{1}{n_h} \sum_{i \in S_h} x_i$$

Using Equation 12, the mean inventory and inventory deficient for each stratum, Dane County has estimated means of $\bar{y}_1 = 673$ linear feet of ditch for each 0.1 mile segment and $\bar{x}_1 = 245.6$ linear feet of deficient ditches per segment. In Portage County, $\bar{y}_2 = 682$ and $\bar{x}_2 = 33.3$.

$$\bar{y}_1 = \frac{1}{5}(1056 + 1056 + 923 + 0 + 330) = 673$$

$$\bar{x}_1 = \frac{1}{5}(528 + 450 + 250 + 0 + 0 + 0) = 245.6$$

$$\bar{y}_2 = \frac{1}{3}(230 + 1056 + 760) = 682$$

$$\bar{x}_2 = \frac{1}{3}(0 + 100 + 0) = 33.33$$

The sample means are the weighted mean of the strata means as shown in Equation 13. $\bar{y}$ is an estimate the average quantity of inventory per segment in the population and $\bar{x}$ is an estimate for the average quantity of deficient inventory per segment.

Equation 13. Mean Quantity of Feature Inventory and Deficient Inventory per Sample Segment

$$\bar{y} = \sum_{h=1}^{L} w_h \, \bar{y}_h$$

$$\bar{x} = \sum_{h=1}^{L} w_h \, \bar{x}_h$$

In Equation 13 the stratum weights w_h are the portion of total centerline miles represented by each stratum. Using data from Table 11, the stratum weights are again $w_1 = 0.6276$, and $w_2 = 0.3724$.

$$\bar{y} = 0.6276(823\ ) + 0.3724(682) = 676.35 \text{ Linear feet}$$

$$\bar{x} = 0.6276(245.6) + 0.3724(33.33) = 166.56 \text{ Linear feet}$$

Finally, Equation 14 is the estimated overall percent of the feature quantity that is deficient:

Equation 14. Estimated Percentage of Feature Quantity that is Deficient Using Type-C Estimators

$$\hat{r}^C = \frac{\bar{x}}{\bar{y}} = \frac{\sum_{h=1}^{L} w_h \bar{x}_h}{\sum_{h=1}^{L} w_h \bar{y}_h}$$

For the example, from Equation 14, 24.63 percent of all linear feet of ditch are deficient.

$$\hat{r}^C = \frac{166.56}{676.35} = 0.2463$$

7.3.3.2. Standard Error

Because the random variables, x and y are not proportions of the number of sample segments, formulations for computing the sample mean and variance are different from those used for the Type-A and Type-B estimators. Equation 15 is used to compute the standard error for the Type-C estimator and Equation 16 are for computing the strata variances. As for Type-B estimators, this is a linearized estimate of the ratio's variation, and thus requires that the denominator, $\bar{y}$, not be too close to zero.

Equation 15. Standard Error for Type-C Estimators

$$StdErr(\hat{r}^c) = \sqrt{\frac{1}{\bar{y}^2}\sum_{h=1}^{L} w_h^2 \left(\frac{1}{n_h}\right)(s_{xh}^2 + s_{yh}^2(\hat{r}^c)^2 - 2\hat{r}^c s_{xyh})}$$

Equation 16. Strata Sample Variances for Type-C Estimators

$$s_{x1}^2 = \frac{1-f_1}{n_1-1}\sum_{i\in S_h}(x_i - \bar{x}_1)^2$$

$$s_{y1}^2 = \frac{1-f_1}{n_1-1}\sum_{i\in S_h}(y_i - \bar{y}_1)^2$$

$$s_{x2}^2 = \frac{1-f_2}{n_2-1}\sum_{i\in S_h}(x_i - \bar{x}_2)^2$$

$$s_{y2}^2 = \frac{1-f_2}{n_2-1}\sum_{i\in S_h}(y_i - \bar{y}_2)^2$$

$$s_{xy1} = \frac{1-f_1}{n_1-1}\sum_{i\in S_1}(x_i - \bar{x}_1)(y_i - \bar{y}_1)$$

$$s_{xy2} = \frac{1-f_2}{n_2-1}\sum_{i\in S_2}(x_i - \bar{x}_2)(y_i - \bar{y}_2)$$

For the ditch feature example, the strata variances are found by substituting into Equation 16,

$$s_{x1}^2 = \frac{1-0.00143}{5-1}((528-245.6)^2 + (450-245.6)^2 + (250-245.6)^2 + 2(0-245.6)^2) = 60{,}460.3$$

$$s_{y1}^2 = \frac{1-0.00143}{5-1}(2(1056-673)^2 + (923-673)^2 + (0-673)^2 + (330-673)^2) = 23{,}1283.1$$

$$s_{x2}^2 = \frac{1-0.00144}{3-1}(2(0-33.33)^2 + (100-33.33)^2) = 3{,}328.5$$

$$s_{y2}^2 = \frac{1-0.00144}{3-1}\left((230-682)^2 + (1056-682)^2 + (760-682)^2\right) = 174{,}879.0$$

$$s_{xy1} = \frac{1-0.00143}{5-1}((528-245.6)(1056-673)+(450-245.6)(1056-673)+(250-245.6)(923-673) + (0-245.6)(0-673)+(0-245.6)(330-673)) = 109{,}112.2$$

$$s_{xy2} = \frac{1-0.00145}{3-1}\big((0-33.33)(230-682)+(100-33.33)(1056-682)+(0-33.33)(760-682)\big) = 18{,}672.9$$

Substituting these values into Equation 15 gives

$$StdErr(\hat{r}^C) = \left(\frac{1}{676.35^2}\left(\left(\frac{0.6276^2}{5}\right)(60{,}460.3+231{,}383.1(0.2463)^2-2(0.2463)(109{,}112.2)) + \left(\frac{0.3724^2}{3}\right)(3328.5+174{,}879.0(0.2463)^2-2(0.2463)(18{,}672.9))\right)\right)^{1/2} = 0.0637$$

Thus, the standard error for the estimate of deficient ditch length in percent of total linear feet is 6.37.

7.3.3.3. Confidence Interval

The confidence interval may be calculated using the Wald interval equation again (Equation 17).

Equation 17. Confidence Interval for the Type-C Estimator

$$\hat{r}^c \pm z * StdErr(\hat{r}^c)$$

For the ditch example, the 95 percent confidence interval is 12 to 37 percent. This estimate requires that $\hat{r}^c$ be approximately normal. For large sample sizes (a minimum sample of 30 is often used as a rule of thumb), this assumption is usually reasonable.

7.3.3.4. Sample Size

As this example makes clear, larger samples are required to make the confidence intervals meaningful for setting LOS targets. Equation 18 is used to estimate the required sample size for Type-C estimators. E is the desired margin of error.

Equation 18. Required Sample Size for Type-C Estimators

$$n^c_{req} = \left(\frac{z}{E}\right)^2 StdErr^2(\hat{r}^c) * n$$

This gives:

$$n^c_{req} = \left(\frac{1.96}{0.05}\right)^2 0.06365^2 * 8 \approx 50$$

Thus, for this example, to achieve the desired precision of 5 percent when estimating the overall percentage of deficient ditch length with a confidence level of 95 percent, 50 sample units would be required.

7.3.3.5. Design Effect

The design effect for Type-C estimators is the squared fraction of the standard error for the stratified design to the standard error for a simple random sample (srs) design.

Equation 19. Design Effect for the Type-C Estimator

$$DEFF^C = \left(\frac{StdErr(\hat{r}^c)}{StdErr_{srs}(\hat{r}^c)}\right)^2$$

The standard error for the simple random sampling design $StdErr_{srs}(\hat{r}^c)$ requires a new formula. Because strata are being ignored for the random sample, the different weights of the samples must be accounted for in another way. The standard error is found using Equation 22, which itself requires the use of Equation 20 and Equation 21, as follows. First, compute the individual sample errors e_i, using Equation 20, as the difference between the measured and expected lengths of deficient feature on each sampled segment. Then, calculate the new weights v_i for each sample. These values are summarized in Table 16, which can then be used along with $\hat{r}^c$ in Equation 22.

Equation 20. Segment level error for Type-C Estimators

$$e_i = x_i - \hat{r}^c y_i$$

The segment level weights v_i, depend on the stratum weight w_h and the number of samples from the stratum n_h. Equation 21 is used to compute the segment level weights. The sum of the segment weights over all strata is one.

Equation 21. Segment level weight for Type-C Estimators

$$v_i = \frac{w_h}{n_h}$$

For example, $v_1 = \frac{w_1}{n_1} = \frac{0.6276}{5} = 0.126$. The stratum level values w_1 and n_1 are used because segment one is in stratum 1. Similarly, $w_2 = 0.3724$ and $n_2 = 3$ can be used to calculate $v_6 = \frac{0.3724}{3} = 0.124$, since segment 6 is in stratum 2. Table 16 lists the segment weights for the example of the ditch feature.

Table 16. Data Table Example for Computing Design Effect of a Type-C estimator

			Linear feet			
i	h	County	Ditches y_i	Deficient Ditches x_i	Segment Error e_i	Segment Weight v_i
1	1	Dane	1056	528	267.94	0.126
2	1	Dane	1056	450	189.94	0.126
3	1	Dane	923	250	22.70	0.126
4	1	Dane	0	0	0.00	0.126
5	1	Dane	330	0	-81.27	0.126
6	2	Portage	230	0	-56.64	0.124
7	2	Portage	1056	100	-160.06	0.124
8	2	Portage	760	0	-187.16	0.124

Finally, the components are combined in Equation 22, where f is the region-wide sampling rate, n is the total number of sample units, and, $\bar{y}$ the weighted mean feature quantity on all sampled segments.

Equation 22. Standard Error for a Type-C Estimator for the Simple Random Sampling Design

$$StdErr_{srs}(\hat{r}^c) = \sqrt{\frac{n(1-f)}{\bar{y}^2(n-1)} \sum_{i=1}^{n} v_i^2 e_i^2}$$

For the ditch feature example, $f = \frac{8}{5{,}576.4} = 0.00143$, $n = 8$, and, $\bar{y} = 676.35$ feet.

$$StdErr_{srs}(\hat{r}^c) = \left(\frac{8(1-0.00144)}{676.35^2(8-1)}\left(\left(0.126(267.94)\right)^2 + \left(0.126(189.94)\right)^2 + \left(0.126(22.70)\right)^2 + \left(0.126(0)\right)^2 + \left(0.126(-81.27)\right)^2 + \left(0.124(-56.64)\right)^2 + \left(0.124(-160.06)\right)^2 + \left(0.124(-187.16)\right)^2\right)\right)^{\frac{1}{2}}$$

This gives a value of $StdErr_{srs}(\hat{r}^c) = 0.0835$. Then from Equation 19, the design effect for the Type-C estimator is:

$$DEFF^C = \left(\frac{StdErr(\hat{r}^c)}{StdErr_{srs}(\hat{r}^c)}\right)^2 = \left(\frac{0.0637}{0.0835}\right)^2 = 0.827.$$

Since $DEFF^C < 1$, the stratified sampling design has the effect of decreasing variance when compared to the simple random sampling design.

7.4. Plot of Confidence Interval on LOS Scale

The results of the statistical analysis are the estimated deficiency rate ($\hat{p}$) and confidence interval (CI) for each maintenance feature. These results are useful for communicating the current LOS and the uncertainty associated with the estimates derived from the sample.

Figure 17 is simple chart that shows the range of the estimated deficiency rate superimposed on the LOS grading scale. The chart readily shows which maintenance activities are doing well and which are not. The chart also shows for which features the agency can be most confident about the estimated deficiency rates. The confidence interval for drop-off/build-up (unpaved) is wide but fully within the F grade range. The estimated deficiency rate is not precise but the LOS score is clearly F.

The plot was created using the *High-Low-Close* chart template in Microsoft Excel. This type of plot, normally used for stocks, requires three data series in thise order: upperbound of confidence interval, lowerbound of confidence interval, estimated deficiency rate. The data used to create the chart in Figure 17 is shown in Table 17.

The plot is a visual assessment of the confidence interval showing the range of LOS grades that fall into the confidence interval. Interpreting results for protective barriers and regulatory/warning signs may be problematic because the wide confidence intervals span multiple LOS grade ranges. The statistical analysis procedures in the previous section include equations for determining the required sample size for a specified acceptable margin of error. The margin of error should be equal to one half of the width of the LOS grade range.

Table 17. Sample Data for Creating a High-Low-Close Plot of Deficiency Rate and Confidence Interval

	Confidence Interval		Estimated Deficiency Rate
	Upper Bound	Lower Bound	
Drop-off/build-up (paved)	4.60%	2.40%	3.33%
Drop-off/build-up (unpaved)	40.31%	34.57%	37.39%
Edgeline Markings	9.31%	6.27%	7.66%
Hazardous Debris	9.53%	6.47%	7.87%
Protective Barriers	8.26%	1.68%	4.97%

	Confidence Interval		Estimated Deficiency Rate
	Upper Bound	Lower Bound	
Regulatory/Warning Signs (emergency repair)	4.83%	0.64%	2.74%
Centerline Markings	7.81%	5.05%	6.29%

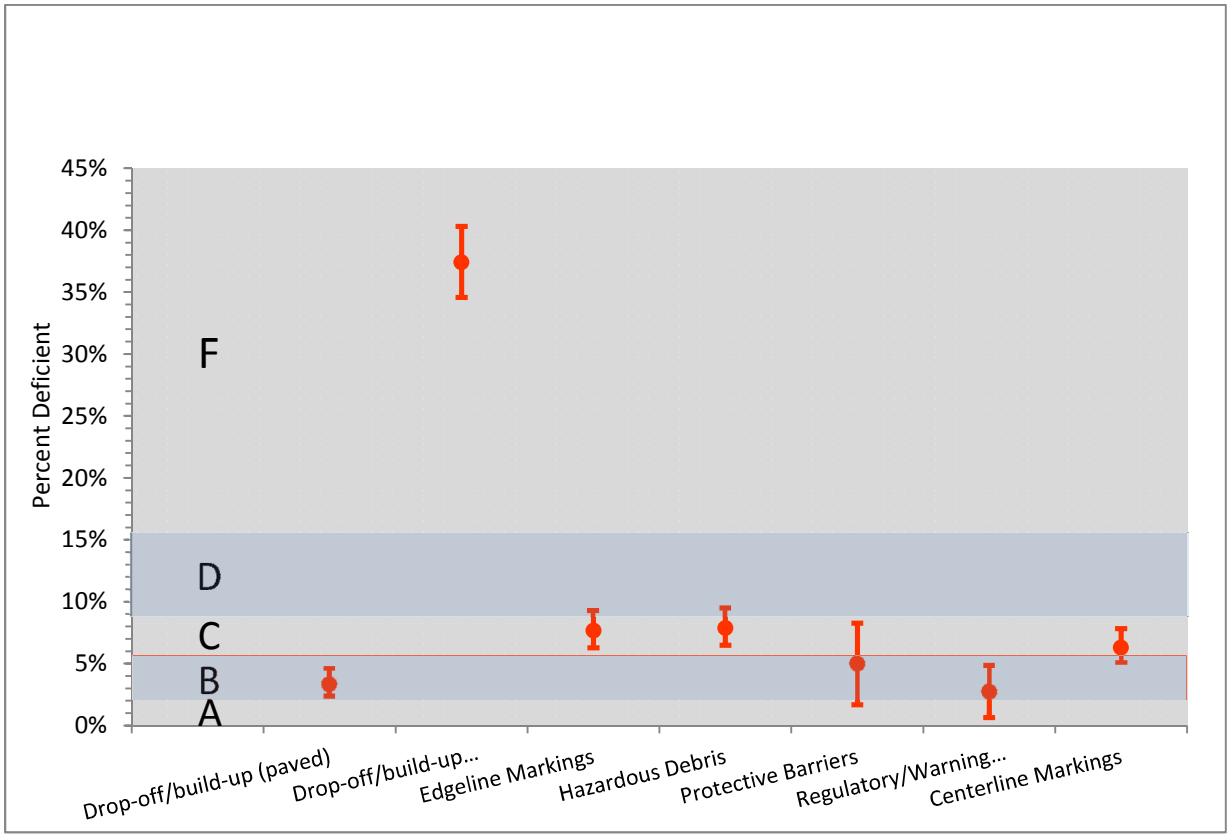


Figure 17. Hi-Lo Plot Showing Estimated Deficiency Rates and Confidence Intervals on the LOS Scale

7.5. Maintenance Cost Data

Agencies vary widely in what they know about the costs of the things they do. A few have fairly sophisticated cost accounting systems. Most know the cost of inputs: labor, machines, and materials. Most agencies interviewed for this project admitted that they had little confidence in their cost information. But cost information is essential to setting meaningful targets, for managing a maintenance program, and making trade-off decisions. Regardless of where a given agency is on this continuum, the following ideas are aimed at moving from the cost of inputs to the cost of strategies.

Strategies are actions taken in pursuit of a maintenance goal: cleaning ditches, mowing, filling cracks, placing pavement markings, replacing signs, patching potholes, etc. The estimates in question are useful at a system planning level. Like all management system data, these estimates become less useful as they are applied to specific locations.

System level cost information can be generated in two basic ways, with variations within each: 1) cost accounting; or 2) estimates made by knowledgeable people using the best data available.

7.5.1. Cost Accounting

Cost accounting is the preferred way of gathering the information on the cost of strategies. Under this approach unique accounts or projects are established for specific strategies—cleaning ditches, mowing etc.—and all labor, machine hours, and materials related to that particular strategy are charged to this specific project. This project, or account, might relate to all such work completed across the state, within a region, a county, or at a specific location. More detail provides better information for management. Information at specific locations can provide ideas that can improve the overall process. Information at the region or county can provide internal benchmarks that might improve the overall efficiency of operation.

More detail also means more accounts to be charged by workers and crews, with more opportunity for confusion and error. The fear of "making highway workers into bookkeepers" is probably the reason that few states have full-blown cost accounting systems. It is also probably the reason that many states have some questions about the accuracy of cost data. Automated data collection and geographic positioning and vehicle locator systems are helping cut and manage the data entry burden, while improving data quality.

Washington State has grouped its maintenance activities into ten groups, or categories. Each group is divided into several subgroups, and each subgroup is comprised of a number of charge codes. Figure 18 is a sample of their structure. The subgroups are generally at the level that we call strategies in this Guide.

Group 1 Roadway Maintenance and Operations

1A1 Pavement Patching, Repair & Crack Sealing

1111	MECHANICAL PATCHING GRADER	TONS
1112	MECHANICAL PATCHING PAVER/BOX	TONS
1122	MANUAL PATCHING	SQ FT
1134	PAVEMENT MILLING/FULL DEPTH PATCHING	TONS
1137	CONCRETE PATCHING WITH SUBGRADE REPAIR	SF
1147	CHIP SEAL PATCHING	100 GALS
1148	PATCHING - DURA PATCHER	GALLON
1149	MAINTENANCE OF STOCKPILE SITES (NON-WINTER RELATED)	NONE
1161	CRACK & JOINT SEALING	LBS
1173	ASPHALT PATCHING WITH SUBGRADE REPAIR	SQ FT
1198	YARD & SHOP CLEANUP (NON-WINTER RELATED)	NONE
1199	OTHER ROADWAY MAINTENANCE AS APPROVED BY SUPERINTENDENT	NONE

Figure 18. Washington State's MAP Structure

The third level codes tie directly to the maintenance accounting system. Therefore, the agency can roll-up cost to the group and subgroup levels. The part that they are working on, and have only partially solved, is agreement on units. Subgroup 1A1 (Figure 19) uses several different units. Without common units, it is not possible to obtain unit costs at the higher levels without further work and estimating.

Washington State's work provides an advantage for maintenance management. Each of the subgroups comes with a short description of what it involves and an overview of the typical workforce and equipment needed to perform the work. Figure 19 is an example:

Group 1 - Roadway Maintenance & Operations

1A1 - Pavement Patching, Repair & Crack Sealing

Activities required to fix pavement deficiencies such as pot holes, alligator cracking, and rutting in order to extend the life of the pavement. Work may include digging out old broken pavement, any unstable base material, and placing and compacting new, free draining, base material and asphalt mix with mechanized equipment or hand tools depending on the size of the patch. For large areas of deficient pavement, an asphalt distributor spraying hot liquid asphalt or emulsified asphalt on the distressed pavement area, and covering it with crushed rock and rolling the rock to compact and seat the stones together, or an overlay patch with hot-mix asphalt may be used. This activity also includes grading and repairing gravel roads, repairing broken curbs, and engineering services for special, roadway-related projects. Crews doing this work may vary from 2 to more than 10 people, depending on the size of the repair and amount of equipment needed to accomplish the work. Equipment may include dump trucks, front end loader, motor grader, paving machine, steel roller, rubber tire roller, chip spreader, oil tank trailer and/or an oil distributor.

This activity also includes crack sealing, which is repairing cracks in asphalt and concrete pavement in order to extend the life of the pavement. Random cracking will appear in pavements due to natural aging and traffic action. Cracks over 1/4 inch wide should be filled to prevent water from entering into and weakening the underlying subgrade. Crack sealing is accomplished by pouring hot liquid asphalt in the cracks. Crews doing this work may vary from 2 or more people, depending on the amount of cracking and amount of equipment needed to accomplish the work. Equipment may include dump trucks, compressor, and asphalt kettle.

Figure 19. Subgroup Definition

This definition of the effort required to do the work helps the agency move toward templates that allow the agency to plan work and budgets.

7.5.2. Expert Judgment

The second approach, used by most states is to estimate the costs of strategies from informed estimates. Often these informed estimates come from assembling a group of experienced maintenance people, using the information that is available, and asking a series of questions.

- What is the appropriate unit to measure this work? Using the example of pothole patching, what is the appropriate unit: a square foot of patching or a lane mile? When considered at a system level, the appropriate unit is probably the lane mile. The square foot could probably also be used, but it would require more calculation to make it useful, since the system is measured in miles.
- Next, what is the appropriate cost unit to be applied to a lane mile of patching? It's probably tons of patch material. How many tons of patch materials are typically used per lane mile of pothole patching?
- How many labor hours and machine hours are needed to place a ton of patch materials?

Then using the estimated factors, a reasonable estimate of the cost of doing a lane mile of patching:

(Tons of patch material x cost) + ((machine hours/ton x tons/mile) x cost) + ((labor hours/ton x tons/mile) x cost) = Cost per lane mile

Colorado DOT uses this method.

A variation of this method is to have the expert panel estimate all of the inputs for a procedure directly. Using the pothole patching example, this would require estimates for the tons of patch material, labor, and machine hours per lane mile. The cost of a lane mile could then be calculated:

(tons of material x cost per ton) + (hours of labor x cost per hour) + (machine hours x hourly rates) = cost per lane mile.

Wisconsin DOT uses this method.

Still another way of making expert estimates is to allocate the costs of inputs to the strategies you want to track. Table 18 illustrates this method. The table links expenditures on work functions to maintenance of features.

Table 18. Allocation of Costs

		Stategies and Percent Distribution								
Work function	**Total Cost**	Pothole patching	Pavement repair	Turf	Centerline marking	Edgeline marking	Shoulder Repair	Crack sealing	Misc. Vegetation	Total
Asphalt overlay	$ 100,000		71%				29%			100%
Longline marking	$ 300,000				62%	38%				100%
Seeding & Plant	$ 50,000			84%					16%	100%
Adopt a hwy	$ 20,000			20%					80%	100%
Maintenance of cracks & joints	$ 120,000							100%		100%
Sweep & wash	$ 11,000	16%	11%		22%	17%	20%	14%		100%
Mowing	$ 95,000			88%					12%	100%
Patch material	$ 150,000	100%								100%
Traffic control	$ 75,000	15%	20%	8%	16%	12%	8%	17%	4%	100%
Total Costs	$ 921,000	$ 163,010	$ 87,210	$ 135,600	$ 200,420	$124,870	$ 37,200	$ 134,290	$ 38,400	$921,000
Units		56	27	123	145	176	49	38	15	
Unit cost		$2,911	$3,230	$1,102	$1,382	$709	$759	$3,534	$2,560	

In practice, the matrix would be much larger. North Carolina, which uses this method, has a 50x25 matrix. The objective is to allocate percentages of costs, as the maintenance accounting system collects them, to the features being used in the maintenance management system. This method has the advantage of connecting to the actual expenditures in the program. In Table 18, $921,000 was spent and $921,000 allocated. It has the disadvantage of removing the estimate one or more degrees from the specific question being asked. Rather than asking directly: What is the cost of a unit of pavement repair? The expert is asked: How much of each of these cost categories should to attributed to pavement repair? This would seem to be a more difficult question to answer.

7.5.3. Final Analysis, Adjustments, and Considerations

With any cost estimation it is necessary to specify the year to which the costs apply. Cost indices should be used to adjust costs as needed. Significant changes in labor, materials, and equipment costs should be considered. For example, the agency construction staff probably has some estimate of future asphalt prices. Again, apply those changes to the cost of materials. These adjustments should yield a reasonable estimate of the direct costs associated with doing a lane mile of pothole patching.

Another consideration deals with indirect, or overhead, costs. If the agency uses some type of overhead charge to distribute costs that are not easily attached to specific activities, apply that

overhead rate to the direct cost to arrive at a full cost. In this way, the cost estimate reflects agency accounting and budget practices.

The final step is to prepare the cost information for the subsequent analysis. The cost estimates must be expressed in terms of cost to the remove deficiencies. For ease of calculation and understanding, the target setting approaches uses cost to eliminate deficiency in one percent of the inventory. To make this calculation, use the cost per unit, as arrived at earlier in this section and the total inventory for each feature. Divide the total inventory for each feature by 100 to yield 1 percent of that inventory. Multiply the 1 percent by the unit cost for that feature:

*(Total inventory for each feature/100)*unit cost for each feature=Cost of treating deficiencies in 1 % of inventory*

These methods work well when maintenance is done using state DOT labor. Increasingly, work is being done by contract. When it is done by contract, one of two broad methods is usually employed. The first is contracting for specific work. Using this example, a contract might be let to do pothole patching on 100 lane-miles of road. In that case the cost per mile can be calculated quite easily.

The other contracting method is to charge the contractor with performing all maintenance on route X from point A to point B, or all routes in county C. The agency will probably supply the standards to be attained, but not the anticipated units of work. The payment will likely be based on lane miles or simply an annual lump sum. If the agency has chosen this method of contracting, determining costs becomes dependent on information provided by the contractor.

7.6. Estimating the Level of Effort

So much of public budgeting tends to focus on incremental change, but to be fully in charge of a program the manager needs to clearly understand how the ongoing resources are being spent. What is the level of effort needed to maintain the status quo? In this section, the concept of cycle times will be discussed as a method of estimating the level of effort required to maintain the current service levels. To do this, several things are required:

1. A sound inventory, either an actual count or a good estimate, for each feature for which targets are being set.
2. Actual units accomplished for the past year, better still for the past three years, at the feature level.
3. Actual or estimated cost per unit of accomplishment at the feature level.
4. In addition to the above, the expert judgment of people familiar with the activities of the program in the past.

Using this information, it will be possible to estimate the cycle times for actions at the feature level—that is the intervals at which each segment of the inventory must be treated to maintain current service levels. For some actions the cycle time might be 0.5—twice a year—for others it might be 20—once every twenty years. The average cycle time is preferable to using last year's information because the annual level of activity will probably vary from year to year.

With an estimated average cycle time, the total inventory can be used to calculate the average number of units to be done each year. The average cost per unit can then be applied to determine the budget requirements to maintain the status quo.

The first step in this process is assembling a group of people who are familiar with the maintenance program. Field managers or foremen are preferred. These are the people with the closest hands-on knowledge of the program and also a sound knowledge of state or regional budgets.

The next step is assembling the information listed above. This inventory information is an absolute requirement. If it is not available, get it and start again.

If actual units of activity are available for one or more years, the experts should be asked to review that information and agree on what is the average level of effort required to maintain the status quo. With this agreed information, a cycle time for each feature can be calculated by dividing the inventory by the annual average level of effort.

If the actual number of accomplished units is not available for past years at the feature level, the experts should be asked to make estimates of the cycle times for each feature. As estimates are made, all participants must recognize that these are estimates for the entire state. Urban areas with high traffic volumes will be very different for many items than rural, low traffic volume areas. Different terrain and weather conditions that may exist in different areas will also tend to make the estimates differ. Indeed, if the differences within the state are pronounced, consideration should be given to doing estimates at a regional level and aggregating them to the state level. Aggregate by weighting the cycle time estimate for each region by the feature inventory for that region and dividing the sum of those weighted numbers by the total inventory.

Regardless of how the estimates are done, remember that these are planning level state or region-wide estimates. They will be wrong when applied to specific locations.

Using inventory and cost information from the Wisconsin Department of Transportation, cycle time estimates were made for each feature, as shown in Table 19.

Table 19. Inventory, Cycle times, Costs, and Budget Requirements for WisDOT

Elements and Features	Statewide Inventory	Cycle Time	Maintenance Unit Cost	Units	Average Budget Required
Shoulders:					
Hazardous Debris	11,774	7	$112.05	Call out, 10/CL	$188,468.10
Cracking (paved shoulder)	21,591	15	$2.00	LF, 440 per mile	$1,266,672.00
Drop off/Build up (paved shoulder)	21,591	15	$7,250	mi	$10,435,650
Potholes/Raveling (paved shoulder)	21,591	15	$5.65	SF, 200/mi	$1,626,522.00
Drop off/Build up (unpaved shoulder)	21,619	4	$330	mi	$1,783,567.50
Cross slope (unpaved shoulder)	21,619	12	$2,000	mi	$3,603,166.67
Erosion (unpaved shoulder)	21,619	7	$116.47	mi, 10/mi	$3,597,092.76
Drainage:					
Ditches (Clean)	18,236.48	20	$8,000	mi	$7,294,592.00
Culverts (Clean)	36,266	10	$226	ea	$819,611.60
Under/edge Drains (Clean)	33,424	4	$16.01	ea	$133,779.56
Flumes (Clean)	11,631	8	$37.35	ea	$54,302.23
Curb and Gutter (Clean)	3,396.91	2	$141	mi	$239,482.16
Storm Sewer (Clean)	48,926	5	$115	ea	$1,125,298.00

Elements and Features	Statewide Inventory	Cycle Time	Maintenance Unit Cost	Units	Average Budget Required
Roadsides:					
Litter	29,625	1	$547.03	LM	$16,205,763.75
Mowing	29,625	1	$83.41	LM	$2,471,021.25
Mowing Vision	39,117	1	$83.41	ea	$3,262,748.97
Woody Vegetation	29,625	6	$1,033.40	LM	$5,102,412.50
Woody Vegetation Vision	39,117	3	$258.35	ea	$3,368,625.65
Fences	14,169,357	30	$6.28	LF	$2,966,118.73
Traffic Control and Safety Devices:					
Centerline Markings	56,799,150	1	$0.15	LF	$8,519,872.50
Edgeline Markings	156,417,624	1	$0.15	LF	$23,462,643.60
Special Pavement Markings	48,910	1	$169.60	ea	$8,295,136.00
Reg/Warning Signs (emergency)	159,004	20	$171.48	ea	$1,363,300.30
Reg/Warning Signs (routine)	159,004	8	$121.90	ea	$2,422,823.45
Other Signs (emergency)	122,970	20	$171.48	ea	$1,054,344.78
Other Signs (routine)	122,970	10	$121.90	ea	$1,499,004.30
Delineators	155,793	8	$52	ea	$1,012,654.50
Protective Barriers (Beam Guard)	3,704,457	15	$34.36	LF	$8,485,676.17
					$121,660,351.02

Abbreviations:
CL = centerline miles
LF = linear feet
mi = miles
SF = square feet
ea = each
LM = lane miles

As with other illustrations in this Report, Table 19 is intended to help the user understand the concepts. While the cycle times are reasonable, they should not be used without further thought and analysis.

In Table 19, it is estimated that Beam Guard will have to be treated once in every 15 years—a 15-year cycle time. The inventory is 3,704,457 linear feet. Therefore, in an average year 1/15th, or about 247,000 feet, of the inventory must be treated. The cost per linear foot is $34.36. With this level of effort, about $8.5 million is needed for Guard Beam maintenance each year. Doing the similar computation for all of the features produces an annual budget requirement of $121, 660,000.

The bottom-line budget requirement should be used as a first-level reality check. Is it close to the amount that is actually available to the state? If the answer is no, the cycle times have probably been underestimated and the experts should be asked to consider them a second time.

Another question that must be asked is: Are the cost estimates fully loaded, or are other demands made on the maintenance budget not shown in the list of features? For example, are all administrative, standby, patrolling, or research costs included? If not, additional items must be added to the table before the bottom line comparison can be made.

With the information contained in Table 19, it is possible to model the total maintenance budget within a linear program. It is also possible to use targets to create a budget. The research team prepared the guidebook to illustrate how the maintenance budget can be linked to the target setting process.

8. Context for Target Setting

Not all maintenance activities are equally important; some are more important than others. As targets are established, this concept of relative importance must be understood and explicitly considered. Without this understanding of importance, everything is equal, and, if everything is equal, decision-making is impossible and targets have little meaning.

This chapter presents methods for establishing the agency-specific elements that are essential for LOS target setting. These include methods for quantifying priorities that reflect both internal and external perspectives within maintenance; a systematic approach for quantifying the relative effectiveness of maintenance activities in achieving maintenance goals; and an approach for determining the marginal cost and efficiency of maintenance expenditures.

The methods in this chapter build upon existing elements already established for the agency's maintenance program. These include the agency and maintenance program goals, the maintenance quality assurance program, and maintenance cost data. The process for using the tools and analytical steps in this chapter is shown in Figure 20. The first step in the process will establish the parameters that set the context for LOS target setting. Those parameters are indicated by the shaded boxes in Figure 20.

1. Set priorities within maintenance.
    a. Define maintenance goals to support the agency's strategic goals.
    b. Array maintenance goals in terms of their relative priority.
    c. Perform comparisons of maintenance goals using the AHP tool.
2. Define how maintenance features contribute to goals.
    a. Define maintenance categories and features.
    b. Relate categories and features to maintenance goals.
    c. Array features in terms of their relative priority within each maintenance goal.
    d. Perform a pair-wise comparison of features within each goal using the AHP method.
3. Calculate costs and marginal costs.
    a. Use unit cost and inventory information to estimate the cost of treating 1 percent of the inventory of each feature to reduce deficiencies.
    b. Create the hierarchy showing the relative contribution of each feature to each maintenance goal and the relative contribution of each goal to the overall maintenance program using the results of the pair-wise comparisons.
    c. Use the results of the pair-wise comparison and the cost analysis to estimate marginal cost reflecting the cost effectiveness of each feature.

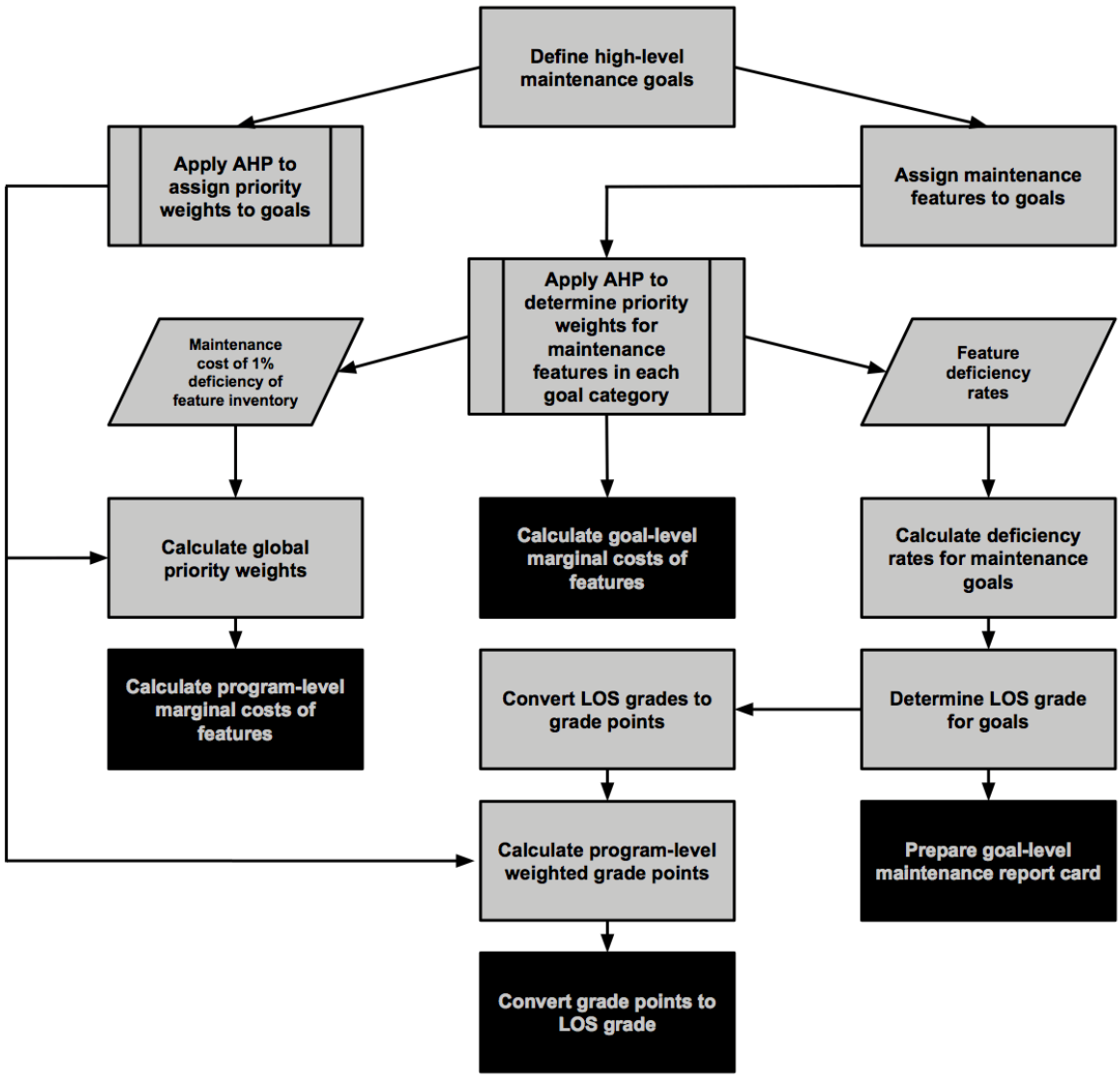


Figure 20. Process for Developing the Context for Target Setting: Priorities, Effectiveness, and Marginal Costs

8.1. Setting Maintenance Priorities

Two major perspectives must be recognized in defining priorities: the external (or customer) and the internal (or technical). Several methods are available to systematically determine priorities.

8.1.1. External Customer Input

External customers include elected policymakers, interest groups, local government officials, and the general public. External groups tend to be more interested in higher order issues, what we have called strategic objectives. A variety of tools can be used to gather their input:

- Customer comment cards can help the agency gauge satisfaction with rest areas.
- Surveys can help understand how the public sees various issues.
- Focus groups can help dig deeper into specific topics with specific groups.

- Organized road trips, where people are driven over a defined course and asked specific questions about the route they have just travelled, can provide very detailed information on preferences and values.

All of these tools and others can give the agency an understanding of what is important to the external groups who pay the bills, use the system, or have an influence over the policy direction of the agency. The tools that will be described later in this chapter can also be used with external groups to gain a more quantitative understanding of the importance that external groups assign to maintenance issues and activities.

8.1.2. Internal Agency Perspective

A transportation agency's highway maintenance section shares goals, strategies, and even resources with other functional areas within the agency. Ideally, upper management will have taken action to support internal alignment so that all staff has some understanding of the agency's strategic direction and of how their efforts contribute to the goals.

Wisconsin's Compass Program is used to illustrate an approach to defining priorities and the relationship between maintenance features and the LOS outcomes for achieving the priorities.

Each of the agency's high-level maintenance goals is clearly defined in a way that is meaningful to maintenance employees and agency professionals:

- **Critical safety**. If not properly functioning, critical safety features would require immediate remedial action, achieved with overtime pay if necessary.
- **Safety**. Highway features and characteristics that protect users against, and provide them with clear sense of freedom from, danger, injury, or damage.
- **Stewardship**. Actions taken to help a highway element obtain its full potential service life.
- **Ride/comfort**. Highway features and characteristics, such as ride quality, proper signing, or lack of obstructions that provide a state of ease and quiet enjoyment for highway users.
- **Aesthetics**. The display of natural or fabricated beauty items, such as landscaping or decorative structures, located along a highway corridor. Aesthetics includes the absence of litter and graffiti that detract from the sightlines of the road.

The goal descriptions are informative on how the agency regards the relative importance of safety features. Some safety features are more immediate and critical that others. Understanding this fact and reflecting it in targets and decisions can be useful and important for program outcomes.

8.2. Assigning Priority Weights to Maintenance Program Goals

Assigning priorities requires comparing one goal to another. The comparison will be based on judgment from the perspective of internal or external stakeholders. This can be done very informally, using a technique such as the Simple Multi-attribute Rating Technique (SMART) or a more analytically rigorous approach such as the Analytical Hierarchy Process (AHP) (Saaty, 2009). The research explored both options. Both methods can be used to establish a set of weights that reflect the relative importance of maintenance goals.

The methods can also be used to quantify the relative contributions of particular maintenance activities in accomplishing those goals. These weights can be used to evaluate maintenance performance toward achieving the goals. When combined with cost data, these weights can be used to optimize the allocation of maintenance resources, or to estimate the performance outcome of a maintenance spending plan.

Sections 8.2.1 and 8.2.2 provide examples of SMART and AHP techniques, respectively. The examples are based on the Compass Program for maintenance quality assurance used by the State of Wisconsin and the Level of Service (LOS) ratings for maintenance features that contribute to achieving the Compass goals.

8.2.1. SMART Technique for Assigning Weights

In SMART, weights of importance may be derived from judgment of relative importance. Assignments of relative importance may be based on the judgment of one person or the collective judgments of many persons. The process begins by having the participants order the goals by importance and assign an arbitrary importance value of 10 to the least important attribute. The participants then judge how much more important each of the remaining goals is in relation to the least important and assign importance values in multiples of ten on an open-ended scale. Finally, the importance values are normalized to produce the weights.

The simplicity of the SMART technique makes it suitable for a general audience. The example, illustrated in Table 20, focuses on weighting the priorities of five maintenance goals used by the Compass Program in the State of Wisconsin. In this example we assume the external participants are a focus group of citizens. Each of the maintenance goals must be defined in a way that is meaningful to the public.

The importance values are assigned by the participants. The weights are computed by normalizing the importance values. Each weight is the importance value divided by the sum of all importance values, 260 for this example. The sum of all of the weights is 1.0.

Table 20. Example Use of the SMART Technique to Establish Weights of Importance

Maintenance Goal	Rank	Importance Value	Weight
Critical Safety	1	100	0.38
Safety / Mobility	2	80	0.31
Stewardship	3	50	0.19
Ride / Comfort	4	20	0.08
Aesthetics	5	10	0.04
	Total	**260**	**1.00**

The SMART technique is easy to apply especially if there are many variables. The technique is also known as the ratio weighting method and is considered better than simple ranking because the weights reflect the relative value for each criterion (goal or objective). The individual criterion may be expressed in terms of its native measurement system (e.g., lane mile, centerline mile, or square feet).

The SMART technique can be used to prioritize and assign relative weights to the maintenance goals. However, the method relies on a single comparison of each element to the least important element. Those single comparisons of then lead to weights of relative important between each element and every other element. For our example in Table 20, the judgments of relative important of critical safety to aesthetics and stewardship to aesthetics lead the implied relative important of critical safety to stewardship. In this case the goal of critical safety is weighted twice as important as the goal of stewardship (0.19 X 2 = 0.31). The implied relative weights may or may not be reasonable.

While the SMART method is useful, it has no way to check the consistency of the implied relative weights. Of course the analyst could ask the participants to verify the implied relative weights and make appropriate adjustments. This would be an iterative process. In the following section, we illustrate the use of paired comparisons to develop the weights of importance. The method considers comparisons between all elements and offers a quantitative test for the logical consistency of the full set of comparisons.

8.2.2. Analytical Hierarchy Process Approach for Assigning Weights

An analytical approach for assigning the priorities is the Analytical Hierarchy Process (AHP)(Saaty, 2009). The method can be used to establish a set of weights that reflect the relative importance of the maintenance goals or features. AHP is a structured technique for organizing and analyzing complex decisions that has been refined over the last 40 years. AHP provides a comprehensive and rational framework for structuring the maintenance target setting decision problem, for representing and quantifying maintenance activities, for relating those activities to overall goals, and for evaluating alternative solutions.

AHP helps decision makers find what best suits their goals and their understanding of the problem, rather than trying to arrive at one correct answer. Users first decompose their decision problem into a hierarchy of easily comprehended maintenance goals and maintenance features, which can each be evaluated separately. Evaluators and decision-makers systematically consider the goals (features), comparing two to each other at a time, with respect to their impact on the overall maintenance program (goals) above them in the hierarchy. In making the comparisons, the decision-makers can use concrete data about the goals (features), but they typically use their judgments about the relative meaning and importance. It is the essence of the AHP that human judgments, and not just the underlying information, can be used in performing the evaluations (Saaty, 2008).

AHP converts these evaluations to numerical values so they can be compared over the entire range of the problem. A numerical weight or priority is derived for each element of the hierarchy, allowing diverse and often incommensurable elements to be compared to one another in a rational and consistent way. This capability distinguishes AHP from other decision-making techniques.

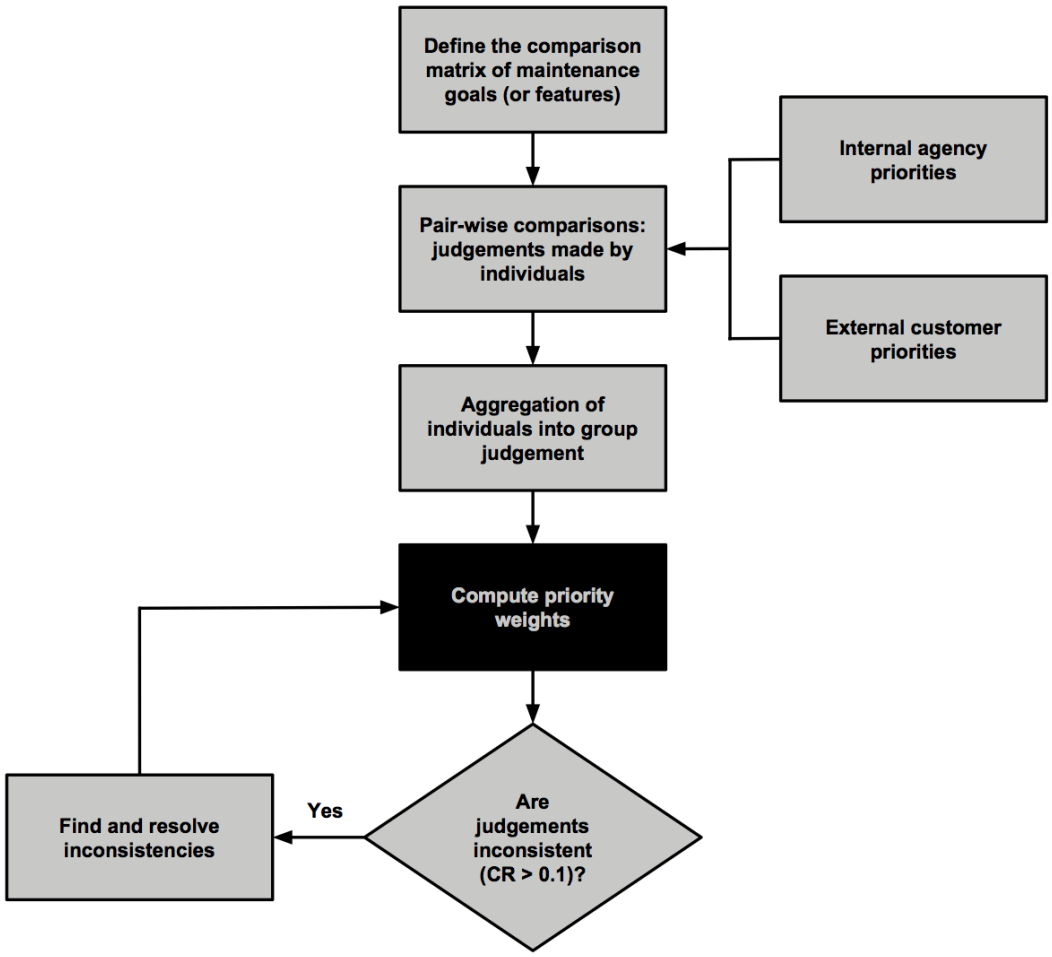


Figure 21. Analytical Hierarchy Process (AHP) for Determining Priority Weights for Maintenance Goals and Features

The methods can also be used to quantify the relative contributions of particular maintenance activities in accomplishing maintenance goals. When combined with cost data, these weights can be used to estimate the marginal cost (a measure of effectiveness) for each maintenance activity, to optimize the allocation of maintenance resources, or to estimate the performance outcome of a maintenance spending plan.

The research team developed detailed examples of the steps in AHP techniques shown in Figure 21. The examples are based on the Compass Program for maintenance quality assurance used by the State of Wisconsin and the Level of Service (LOS) ratings for maintenance features that contribute to achieving the Compass goals.

8.2.3. Pair-wise Comparison Judgments

AHP uses pair-wise comparisons to define relative importance. Each comparison is an expression of an opinion about the dominance (intensity of strength) of one item over another with respect to a single property. The set of all comparisons is organized into a square reciprocal matrix such as the one shown in Table 21. Each element of the matrix represents the dominance of the row item over the column item. The dominance number reflects the answer to two questions: Which of the two elements is dominant with respect to a single criterion and how strongly is that preference or

dominance. For the example in Table 21, the items are the maintenance goals and the single criterion is contribution to the agency's overall mission.

Table 21. Matrix of Relative Importance Ratings for Maintenance Goals of the Wisconsin Compass Program

Maintenance Goal	Critical safety	Safety/mobility	Stewardship	Ride/comfort	Aesthetics
Critical safety	1	3	7	8	9
Safety/mobility	1/3	1	4	7	9
Stewardship	1/7	1/4	1	3	4
Ride/comfort	1/8	1/7	1/3	1	3
Aesthetics	1/9	1/9	1/4	1/3	1

A convenient starting point for creating a comparison matrix is to order the maintenance program goals from most to least important. The first column in Table 21 lists the goals from most through least important. The goals are in the column headings in the same order. For the Wisconsin example, the goal of critical safety is more important than safety/mobility, and safety/mobility is more important than stewardship and so on.

The dominance scale to be used for the pair-wise comparisons is shown in Table 22. The first column shows the scale of numbers that indicate how many times more important the item on the row of the comparison matrix is over the item on the column with respect to the criterion used for the comparison. In this example, the single criterion is contribution to the agency's overall mission.

The dominance scale was used to assign the paired comparisons of importance among the maintenance goals in Table 22. For example, the goal of critical safety is considered to be 7 times more important than the goal of stewardship and 9 times more important than aesthetics. The cost to achieve these goals or willingness to allocate a budget to each goal should not be considered at this time. The matrix values on the diagonal are one because the goal on the row is the same as the goal on the column. Reciprocal values of the importance comparisons are then assigned to the lower triangle such that the reciprocal of the value in cell (i,j) is placed in cell (j,i).

In the AHP method, the dominance of the most important goal must be no more than nine times the least important goal. If the goals differ by more than this range, then the goals should be rearranged into a hierarchy of logical clusters (Saaty, 2009). Within each cluster, the most important goal should be no more than nine times the least important.

Table 22. The Fundamental Scale for Pair-wise Comparisons in the AHP Method

Intensity of Importance	Definition	Explanation
1	Equal importance	Two factors contribute equally to the objective.
2	Weak or slight	
3	Moderate importance	Experience and judgment slightly favor one activity over another.
4	Moderate plus	

Intensity of Importance	Definition	Explanation
5	Strong importance	Experience and judgment strongly favor one activity over another.
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favored very strongly over another; its dominance demonstrated in practice.
8	Very, very strong	
9	Extreme importance	The evidence favoring on activity over another is of highest possible order of affirmation.
Reciprocals of above	If activity *i* have one of the above non-zero numbers assigned to it when compared with activity *j*, then *j* has the reciprocal value when compared with *i*.	A reasonable assumption.

8.2.4. Aggregating Individual Judgments into a Group Judgment

Maintenance targets reflect the values and priorities of the agency so it is important to make a rigorous effort to adequately capture those values and priorities. Some agencies will want to incorporate multiple perspectives, whether internal or external or both.

Combining the individual comparison judgments from multiple participants to produce a single comparison matrix must be done in a special way. The simple arithmetic average will not satisfy the reciprocal relation unless all members of the group have the same individual judgments; and in that case, there is no need to combine the judgments. The way to combine the individual judgments is to multiply them and then take the root equal to the number of individuals. The operation of taking the root *n* of *n* numbers multiplied together, in Equation 23, is known as the geometric mean, $\bar{a}(i,j)_g$.

Equation 23. Geometric Mean

$$\bar{a}(i,j)_g = \sqrt[n]{\prod_{m=1}^{n} a(i,j)_m}$$

8.3. Computing Priority Weights

Given the pair-wise comparisons of priority for item is compared to another. The next step is to calculate overall priorities. The matrix of pair-wise comparisons (known as the *A* matrix for the AHP method) is used to find weights for the importance of each maintenance goal by solving for the normalized principal eigenvector, which measures the relative priorities of the items in the matrix.

There are several methods for calculating the normalized principal eigenvector. Numerical tools, such as Matlab and Mathematica, have built-in functions for solving the generalized eigenvalue problem. There are also downloadable software scripts that can be used in Excel.

In the absence of a computerized numerical solver, two methods below may be used to approximate the vector of priority weights. The research team reviewed the methods and compared results. The guidebook presents on the "preferred" method. Both methods are computationally simple. The preferred method requires slightly more calculations but has an advantage in that the consistency of the pair-wise comparisons can be evaluated as a way to distinguish the expert's comparison judgments from random judgments.

8.3.1. The Preferred Method for Approximating the Priority Weights.

The first, and more accurate of the two methods, uses the geometric mean of each row.

Given the n elements of the ith row of the A matrix, the geometric mean, $\bar{a}_{ig}$, is found by taking the nth root of the product of the n elements. The geometric mean of each row i, where $i = 1, 2, \ldots n$ is:

$$\bar{a}_{ig} = \sqrt[n]{\prod_{j=1}^{n} a_{ij}} = (a_{i1} \cdot a_{i2} \cdot \cdots \cdot a_{in})^{1/n}$$

The following steps can be used to assign weights and check for the consistency of the pair-wise comparisons. First, expand the A matrix with additional rows and columns as shown in Table 23. The subsequent steps are enumerated below.

Table 23. The Preferred Method for Approximating the Priority Weights

Maintenance Goal	Critical Safety	Safety / Mobility	Stewardship	Ride / Comfort	Aesthetics	Geometric mean $(\bar{a}_{ig})$	Priority weight (ω_i)
Critical Safety	1	3	7	8	9	4.324	0.5214
Safety/Mobility	1/3	1	4	7	9	2.426	0.2924
Stewardship	1/7	1/4	1	3	4	1.046	0.1018
Ride/Comfort	1/8	1/7	1/3	1	3	0.404	0.054
Aesthetics	1/9	1/8	1/4	1/3	1	0.226	0.030
Column Sum	**1.712**	**4.504**	**12.583**	**19.333**	**26.000**	**8.293**	**1.000**
$\sum_{j=1}^{n} a_{ij}\,\omega_i$	0.893	1.317	1.281	1.042	0.791	$\lambda_{max} =$	5.325

1. **Compute the geometric mean of each row**. For the example, $n = 5$. For each goal, multiply its five row entries together and take the 5th root of the product. Place this value in the column whose heading is *geometric mean* For example, the 5th root product for Critical Safety would be $(1 \times 3 \times 7 \times 8 \times 9)^{1/5} = 4.324$. Repeat this calculation for the other four maintenance goal.
2. **Estimate the priority weights by normalizing the vector of geometric means.** First, sum the calculated entries in the geometric mean column, i.e., 8.293. Divide each geometric mean by the column sum. For example, the Priority Weight for Critical Safety is 4.324 / 8.293 = 0.521.

The sum of the priority weights must equal 1. The normalized vector of geometric means is a good approximation of the normalized principal eigenvector (the priority weights).

8.3.2. The Shortcut Method for Approximating the Priority Weights

A shortcut for computing the priority weights can be used, but it will produce somewhat different answers. The method assumes the matrix of judgments has acceptable consistency because there is no way to check the consistency of the pair comparisons. The shortcut method is to normalize each column, and then take the average of the corresponding entries in the columns (Saaty & Peniwati, 2008). To apply this method, first, expand the comparison matrix with additional rows and columns as shown in Table 24.

Table 24. The Shortcut Method for Approximating the Priority Weights

Maintenance Goal	Normalized Columns from Table 21					Row Sum	Priority Vector
	Critical Safety	Safety / Mobility	Stewardship	Ride / Comfort	Aesthetics		
Critical Safety	0.5841	0.6661	0.5563	0.4138	0.3462	2.5664	0.5133
Safety/Mobility	0.1945	0.2220	0.3179	0.3621	0.3462	1.4426	0.2885
Stewardship	0.0835	0.0555	0.0795	0.1552	0.1538	0.5275	0.1055
Ride / Comfort	0.0730	0.0317	0.0265	0.0517	0.1154	0.2984	0.0597
Aesthetics	0.0648	0.0246	0.0199	0.0172	0.0385	0.1650	0.0330
Column Sum	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**4.999**	**1.0000**

1. Normalize the comparison ratings in each column. This is done by summing the column then dividing each entry in the column by the sum. Referring to Table 21 for the comparison ratings, the column sum for Critical Safety is $\left(\mathbf{1}+\frac{\mathbf{1}}{\mathbf{3}}+\frac{\mathbf{1}}{\mathbf{7}}+\frac{\mathbf{1}}{\mathbf{8}}+\frac{\mathbf{1}}{\mathbf{9}}\right)=\ \mathbf{1.712}$. The normalized entry for the Safety/Mobility row and the Critical Safety column is, $\frac{\mathbf{1}}{\mathbf{3}} \div \mathbf{1.712} = \mathbf{0.1945}$. Repeat this calculation for the other four maintenance goals.
2. Compute the average of the entries in each row to obtain the priority vector (Saaty & Peniwati, 2008). Table 24 shows the intermediate step of summing the row values before dividing by the number of entries, $\mathbf{n} = \mathbf{5}$.

8.3.3. Comparison of Methods for Computing Priority Weights

Table 25 compares the priority weights derived by using a numerical method to solve for the normalized principal eigenvector (using a tool such as Matlab) with the priority weights computed using the preferred and shortcut methods. For this example, both approximation methods yield priority weights that are close to values found by numerical computation.

Both approximation methods were applied to the same consistent comparison matrix. A measure of the consistency of the comparison is possible from the preferred method but not from the shortcut. The research team reviewed consistency in the following section.

Table 25. Comparison of Priority Weights from Three Computing Methods

	Priority Weights		
		Approximations	
Maintenance Goal	Numerical Method	Preferred Method	Shortcut Method
Critical Safety	0.5277	0.5215	0.5133
Safety / Mobility	0.2866	0.2924	0.2885
Stewardship	0.1005	0.1018	0.1055
Ride / Comfort	0.0541	0.0539	0.0597
Aesthetics	0.0310	0.0304	0.0330

8.3.4. Checking Consistency of Comparison Judgments

The comparison scores represent judgments based on an individual's perspective. By design the matrix requires more comparisons than necessary. The AHP method requires $n(n\text{-}1)/2$ judgments to form a comparison matrix when there are n elements being compared while the minimum set of judgments to construct the matrix is (n-1). Thus $n(n\text{-}1)/2$ - (n-1) comparisons are redundant. The method requires more comparison than necessary because using the minimum number of comparisons may introduce bias whereas redundancy of judgments generally improves the validity of the resulting priority weights. The matrix $A=(a_{ij})$ is consistent if $a_{ij}\, a_{jk} = a_{ik}$, i, j, k= 1, …, n. Real-world pair-wise comparison matrixes are unlikely to be consistent and the possibility for inconsistency increases as the number of elements being compared gets larger.

It is possible to evaluate the quality or trustworthiness of the paired comparisons. AHP provides a way to test the consistency of the comparison matrix. This is done by estimating the maximum eigenvalue, λ_{max}. Eigen theory says that

$$A\omega \;=\; \lambda_{max}\,\omega$$

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{12} & \dots & a_{2n} \\ \vdots & \dots & \vdots \\ a_{1n} & \dots & a_{nn} \end{bmatrix} \begin{Bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{Bmatrix} = \lambda_{max} \begin{Bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{Bmatrix}$$

By adding the equations and rearranging terms:

$$(a_{11} + a_{21} + \cdots + a_{n1})\omega_1 + (a_{12} + a_{22} + \cdots + a_{n2})\omega_2 + \cdots + \;(a_{1n} + a_{2n} + \cdots + a_{nn})\omega_n \\ = \lambda_{max}(\omega_1 + \omega_2 + \cdots + \omega_n)$$

By recognizing that the normalized eigenvector elements $(\omega_1 + \omega_2 + \cdots + \omega_n) = 1,\ \lambda_{max}$.can be estimated as,

$$(a_{11} + a_{21} + \cdots + a_{n1})\omega_1 + (a_{12} + a_{22} + \cdots + a_{n2})\omega_2 + \cdots + \;(a_{1n} + a_{2n} + \cdots + a_{nn})\omega_n = \lambda_{max}$$

Then λ_{max} is used to compute a Consistency Index (CI). The Consistency Index for the A matrix is calculated from Equation 24.

Equation 24. Consistency Index

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

For the example in Equation 24, the steps for computing the consistency index are as follows:

1. Calculate $\sum_{j=1}^{n} a_{ij}\ \omega_i$ for each maintenance goal as the product of its Column Sum and Priority Weight. For example, the product of Column Sum and Priority Weight for Critical Safety is 1.712 x 0.521 = 0.893.
2. Sum the products calculated in Step 1 (= 5.325). This value is known as λ_{max}.
3. Calculate the Consistency Index (CI) using Equation 24, where n = 5.

$$CI = \frac{(5.325 - 5)}{(5 - 1)} = 0.081$$

4. Calculate the Consistency Ratio (CR in Equation 25) for the set of judgments by comparing the CI to the corresponding Random Index (RI). For our example, $RI = 1.12$ and $CR = 0.081/1.12 = 0.072$.

Equation 25. Consistency Ratio

$$CR = \frac{CI}{RI}$$

Ideally, $CR \leq 0.1$. For this example, $CR = 0.072$, which is acceptable.

Table 26. Random Index (RI) for Computing Consistency Ratios

n	3	4	5	6	7	8	9	10	11	12	13	14	15
RI	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.54	1.56	1.57	1.58

CR = 0 means that the comparison judgments are perfectly consistent; that is all $a_{ij}\, a_{jk} = a_{ik}$, for i, j, k= 1, …, n. Saaty and Peniwaty (2008) argue that a CR = 0.1 indicates that the judgments are at the limit of consistency. A CR = 0.9 would mean that the pair wise judgments are nearly random and untrustworthy.

8.3.5. Resolving Inconsistency in Comparison Judgments.

If CR > 0.1, then an offending inconsistent judgment might be identified and resolved by knowing that the matrix A=(a_{ij}) is consistent if $a_{ij}\, a_{jk} = a_{ik}$, i, j, k= 1, …, n. For example, consider the comparison judgments shown in Table 27. For this matrix CR = 0.1132. The challenge is to find the offending inconsistencies and try to resolve them.

Table 27. Example of Inconsistent Pair-wise Comparison Matrix

Maintenance Goal	Critical Safety	Safety/Mobility	Stewardship	Ride/Comfort	Aesthetics
Critical Safety	1	3	7	8	9
Safety/Mobility	1/3	1	4	7	9
Stewardship	1/7	¼	1	5	7
Ride/Comfort	1/8	1/7	1/5	1	3

Maintenance Goal	Critical Safety	Safety/Mobility	Stewardship	Ride/Comfort	Aesthetics
Aesthetics	1/9	1/8	1/7	1/3	1

If i = critical safety, j = stewardship and k= ride, then from the matrix, the assigned values are $a_{ij} = 7$, $a_{jk} = 5$ and $a_{ik} = 8$. The relationship $a_{ij}\, a_{jk} = a_{ik}$ does not hold since the computed value of $a_{ik} = a_{ij}\, a_{jk} = 35$ is very different from the assigned value of 8.

Usually, the inconsistency can be identified by comparing the relative weight across the column in each row. One might notice that critical safety compared to stewardship, ride/comfort and aesthetics are 7, 8, and 9 respectively, meaning that stewardship, ride/comfort and aesthetics are similarly important compared to critical safety. However stewardship compared to ride/comfort and aesthetics are 5 and 7, respectively meaning that stewardship is strongly and very strongly more important than ride/comfort and aesthetics. The relative importance of stewardship, ride/comfort, and aesthetics is inconsistent. We can reduce the inconsistency by changing stewardship compared to ride/comfort from 5 to 3 and the reciprocal from 1/5 to 1/3 and by changing stewardship compared to aesthetics from 7 to 4 and the reciprocal from 1/7 to 1/4. The result is the comparison matrix shown in Table 21 and Table 27. With these simple revisions, the consistency ratio drops from 0.1132 to 0.0725, which is acceptable.

8.4. Relating Maintenance Categories and Features to Goals

The next step in the process is defining how maintenance activities relate to agency goals. This is done by considering the asset elements and features targeted by the maintenance activity. Table 28 shows an example of how each maintenance feature may be assigned to the agency's maintenance goal it primarily supports. The classification relies heavily on professional judgment, an approach that many DOTs can easily apply. The features should be related to one and only one goal. If the feature contributes to more than one goal, the analyst should choose the goal that is most relevant.

Most importantly, the maintenance measures should be reported in a summarized version that is consistent with how they contribute to the maintenance goals. In Table 28, notice the maintenance features are organized by element. The elements relate to the business functional areas of the maintenance program. For example, the traffic control and safety devices are managed by the traffic engineering group. Usually the features for each element share common data sources, condition assessment frameworks, and reporting structures. For LOS target setting, the features should be grouped by the maintenance goal to which they contribute. Stepping back to look at Table 28, we see a many-to-many relationship; elements contribute to multiple goals and goals are accomplished by multiple elements. For LOS target setting, we are interested in the many-to-one relationship between features and goals; many features contribute to one goal. Only one goal column has a check mark for each feature.

Table 28. Assigning Highway Features to Maintenance Outcome Goal as a Means for Establishing Maintenance Priorities (Wisconsin Compass Program)

Category	Feature	Maintenance of this feature contributes primarily to goal of:				
		Critical Safety	Safety / Mobility	Stewardship	Ride / Comfort	Aesthetics
Traffic control & safety devices	Centerline markings	✓				
	Edge line markings	✓				
	Delineators		✓			
	Emergency repair of detour				✓	

Category	Feature	Maintenance of this feature contributes primarily to goal of:				
		Critical Safety	Safety / Mobility	Stewardship	Ride / Comfort	Aesthetics
	markers/ recreation / guide signs					
	Routine replacement of detour markers/ recreation / guide signs				✓	
	Protective Barriers	✓				
	Emergency repair of regulatory / warning signs	✓				
	Routine replacement of regulatory / warning signs		✓			
	Special pavement markings		✓			
Shoulders	Hazardous debris	✓				
	Cracking on paved shoulder			✓		
	Drop-off / build-up on paved shoulder	✓				
	Potholes / raveling on paved shoulder				✓	
	Cross-slop on unpaved shoulder		✓			
	Drop-off / build-up on unpaved shoulder	✓				
	Erosion on unpaved shoulder			✓		
Drainage	Culverts		✓			
	Curb & Gutter			✓		
	Ditches			✓		
	Storm Sewer System		✓			
	Flumes			✓		
	Drains			✓		
Roadside	Fences		✓			
	Litter					✓
	Mowing					✓
	Mowing for Vision		✓			
	Woody Vegetation (*clear zone*)		✓			
	Woody Vegetation Control for Vision		✓			

8.5. Estimating Effectiveness of Maintenance Strategies

The benefits of improving the condition of any element or feature usually cannot be estimated. It is simply not possible to attribute the number of crashes that will be avoided or the number of lives saved by improved pavement markings or better signs, but we still need to understand the relative merit of one set of decisions versus another. AHP allows us to estimate the relative contribution of

one set of actions versus another. Knowing the relative effectiveness is important for tradeoff analysis and for strategic allocation of resources.

As defined in Table 28, each of the maintenance features contributes to the achievement of one of the agency goals. Since the features do not contribute equally, weights may be used to represent their individual contribution. If the maintenance section could eliminate all deficiencies on all features in all of the categories it would be achieving its goals to the maximum possible level. Of course, eliminating all deficiencies is not possible, but eliminating the deficiencies on some features is considered more important than eliminating the deficiency on other features.

8.5.1. Goal-level Priority Weights

The contribution of each feature for achieving the maintenance goals may be assigned using the same pair-wise comparison method as was used to set the weight of importance for the goals (see Section 8.2.3). Again, the analyst orders the maintenance features from most to least important and creates the matrices of comparison judgments. Table 29 through Table 32 show the comparison judgments for the features in each of the goal categories for our Wisconsin example.

An important consideration is the number of features being compared in each category. Comparing more than two features allows for redundancy and therefore greater validity of the judgments. Having too many features opens the possibility of inconsistent judgments. For a set of n elements in a matrix one needs $n(n\text{-}1)/2$ comparisons. Some authors recommend no more than seven elements in order to obtain priorities with admissible consistency (Saaty & Peniwati, 2008).

Table 29. Matrix of Comparison Judgments on Relative importance of Roadway Features for Critical Safety

Roadway Features for Critical Safety	Emergency repair of regulator / warning signs	Hazardous debris	Protective barriers	Centerline markings	Edge line markings	Unpaved shoulder drop-off / buildup	Paved shoulder drop off / build up	Priority Weight
Emergency repair of regulator / warning signs	1	3	5	7	7	8	9	0.42
Hazardous debris	1/3	1	3	5	5	7	9	0.24
Protective barriers	1s/5	1/3	1	3	3	5	7	0.13
Centerline markings	1/7	1/5	1/3	1	3	5	7	0.09
Edge line markings	1/7	1/5	1/3	1/3	1	5	7	0.07
Unpaved shoulder drop-off / buildup	1/8	1/7	1/5	1/5	1/5	1	3	0.03
Paved shoulder drop off / build up	1/9	1/9	1/7	1/7	1/7	1/3	1	0.02

The method requires comparison ratios on a scale that ranges from one to nine. Most humans have difficulty appropriately judging comparisons when the ratios get beyond 9. If the features cannot be compared on the one to nine scale, the analyst should consider redefining the features of the categories. Alternatively, a clustering approach may be used (see Saaty & Peniwati, 2008).

Table 30. Matrix of Comparison Judgments on Relative Importance of Roadway Features for Mobility Safety

Roadway Features for Mobility Safety	Woody vegetation control for vision	Mowing for vision	Special pavement markings	Woody vegetation *(clear zone)*	Culverts	Storm Sewer	Cross slope on unpaved shoulders	Delineators	Routine replacement of regulatory/ warning signs	Fences	Priority Weight
Woody vegetation control for vision	1	3	4	5	5	6	6	7	8	9	0.31
Mowing for vision	1/3	1	3	3	3	5	5	7	7	8	0.20
Special pavement markings	¼	1/3	1	3	3	4	4	5	5	7	0.14
Woody vegetation *(clear zone)*	1/5	1/3	1/3	1	3	3	3	5	5	7	0.10
Culverts	1/5	1/3	1/3	1/3	1	3	3	5	5	7	0.08
Storm sewer	1/6	1/5	1/4	1/3	1/3	1	3	3	5	7	0.06
Cross slope on unpaved shoulders	1/6	1/5	1/4	1/3	1/3	1/3	1	5	5	5	0.05
Delineators	1/7	1/7	1/5	1/5	1/5	1/3	1/5	1	3	5	0.03
Routine replacement of regulatory/ warning signs	1/8	1/7	1/5	1/5	1/5	1/5	1/5	1/3	1	3	0.02
Fences	1/9	1/8	1/7	1/7	1/7	1/7	1/5	1/3	1/3	1	0.01

Table 31. Matrix of Comparison Judgments on Relative Importance of Roadway Features for Infrastructure Stewardship

Roadway Features for Stewardship	Ditches	Curb and Gutter	Flumes	Cracking on paved shoulders	Erosion on unpaved shoulders	Drains	Priority Weight
Ditches	1	3	5	7	7	8	0.46
Curb and gutter	1/3	1	3	5	5	7	0.25
Flumes	1/5	1/3	1	3	3	5	0.13
Cracking on paved shoulders	1/7	1/5	1/3	1	1	5	0.08
Erosion on unpaved shoulders	1/7	1/5	1/3	1	1	5	0.06
Drains	1/8	1/7	1/5	1/5	1/5	1	0.03

Table 32. Matrix of Comparison Judgments on Relative Importance of Roadway Features for Ride Comfort

Roadway Features for Ride / Comfort	Potholes / raveling on paved shoulders	Emergency repair non-regulatory signs	Routine replacement of non-regulatory signs	Priority Weight
Potholes / raveling on paved shoulders	1	5	7	0.73
Emergency repair non-regulatory signs	1/5	1	3	0.19
Routine replacement of non-regulatory signs	1/7	1/3	1	0.08

Table 33. Matrix of Comparison Judgments on Relative Importance of Roadway Features for Aesthetics

Roadway Features for Aesthetics	Mowing	Litter	Priority Weight
Mowing	1	5	0.83
Litter	1/5	1	0.17

After creating the comparison matrices, the analyst will need to determine the priority weights and check the consistency of the comparison judgments. The priority weights can be found by following one of the methods in Section 8.3. The consistency ratio is an important quality check (see Section 8.3.4). For the examples in Table 29 through Table 32, the computed priority weights are shown in the last column.

The priority weights in Table 29 through Table 32 are known as the *goal-level priority weights*. The goal-level priority weights are useful for assessing tradeoffs and effectiveness of maintenance activities when the scope is limited to a single goal category. The vertical sum of the weights in each goal category is 1 and those weights indicate the feature's priority in the goal category.

8.5.2. Global Priority Weights

LOS target setting requires the assessment of tradeoffs in maintenance activities within a single goal category using goal-level priority weights, and across goal categories using global priority weights. The *global priority weights* for the features can be computed by multiplying its goal-level weight by its category weight. The weights are used to evaluate the efficiency of maintenance strategies.

Figure 22 shows the entire hierarchical structure for the example with the goals and priority weights for maintenance contributions to agency high-level goals at the top followed by the weights of importance of maintenance features for achieving the goals. Developing the hierarchy requires care in identifying the goals, maintenance measures, assigning the comparison judgments, and solving for the weights. This effort does not need to be repeated if the agency does not change any of the input components.

The following rules govern the weight values for the entire hierarchy. These rules can be used to check the hierarchy for possible errors.

- The horizontal sum of the weights for the high-level goals is 1
- The vertical sum of the goal-level weights for the features in each goal category is 1.
- The vertical sum of the global weights for the features in each goal category is equal to the priority weight of the category goal.

- The sum of the global weights of all features in all categories is 1.

The effectiveness of each maintenance strategy is directly related to its contribution to improved performance of a single maintenance goal or performance of the overall maintenance program. This effectiveness is directly related to the feature weights in the AHP hierarchy and application of those weights produces the maximum effectiveness contribution. For the example in Figure 22, the maximum relative effectiveness that can be derived from each maintenance feature is the value of its weight. In this way, the contribution of each feature and goal is bounded. Perfect condition of one feature cannot compensate for lack of maintenance on another feature. Similarly, achievement of one maintenance goal cannot compensate for failure to achieve another maintenance goal.

Maintenance Performance

Critical Safety	**0.52**	
	Priority Weight	
Feature	goal	global
Emergency repair of regulator / warning signs	0.42	0.218
Hazardous debris	0.24	0.125
Unpaved shoulder drop-off / buildup	0.03	0.016
Centerline markings	0.09	0.047
Edge line markings	0.07	0.036
Paved shoulder drop off / build up	0.02	0.010
Protective barriers	0.13	0.068

Safety / Mobility	**0.28**	
	Priority Weight	
Feature	goal	global
Routine replacement of regulatory / warning signs	0.02	0.006
Woody vegetation *(clear zone)*	0.10	0.028
Special pavement markings	0.14	0.039
Mowing for vision	0.20	0.056
Culverts	0.08	0.022
Woody vegetation control for vision	0.31	0.087
Cross slope on unpaved shoulders	0.05	0.014
Storm sewer	0.06	0.017
Delineators	0.03	0.008
Fences	0.01	0.003

Stewardship	**0.13**	
	Priority Weight	
Feature	goal	global
Ditches	0.46	0.060
Cracking on paved shoulders	0.08	0.010
Flumes	0.13	0.016
Curb and gutter	0.25	0.032
Drains	0.03	0.004
Erosion on unpaved shoulders	0.06	0.008

Ride / Comfort	**0.05**	
	Priority Weight	
Feature	goal	global
Potholes / raveling on paved shoulders	0.73	0.037
Emergency repair of non-regulatory signs	0.19	0.009
Routine replacement of non-regulatory signs	0.08	0.004

Aesthetics	**0.03**	
	Priority Weight	
Feature	goal	global
Mowing	0.83	0.025
Litter	0.17	0.005

Figure 22. Category and Feature Weights Used to Assess the Contributions of Categories and Features to Overall Maintenance Performance

8.6. Relating Asset Condition to Maintenance Performance

Maintenance performance is assessed by considering the condition of roadway features. The priority weights can be used to combine the feature-level LOS scores to determine composite LOS scores for a goal category or program-wide.

At most agencies, the measurement units for the LOS scale is percentage of inventory. For some agencies, the scale is percentage of inventory in deficient maintenance condition; other agencies use percentage of inventory in non-deficient maintenance. In the Guide and in this Report, we use an LOS scale with thresholds for percentage of inventory in deficient conditions. This scale can be easily transformed to one with thresholds for percentage of inventory in non-deficient condition by subtracting the given threshold values from 100.

8.6.1. LOS Grading Scale

An example LOS grading scale, based on the scale Wisconsin uses, is shown in Table 34. In this example, the LOS scale is more stringent for maintenance of the most important features. The different thresholds levels reflect different expectations for features in the different categories. For example, only 7 percent of highway miles having poor ride quality would be consider excellent. However if motorists must swerve to avoid hazardous debris every 14 miles (7 percent of highway miles), critical safety is a real concern.

Table 34. Example LOS Grading Scale for Percentage of Inventory in Deficient Maintenance Condition

	Percent of Feature Inventory in Deficient Condition				
Maintenance Goal	**A**	**B**	**C**	**D**	**F**
Critical safety	0-2.5%	2.5-5.5%	5.5-9.5%	9.5-15%	>15%
Safety/mobility	0-4.5%	4.5-9.5%	9.5-18.5%	18.5-30%	>30%
Stewardship	0-6.5%	6.5-15.5%	15.5-29.5%	29.5-50%	>50%
Ride/comfort	0-7.5%	7.5-18.5%	18.5-35.5%	35.5-60%	>60%
Aesthetics	0-10.5%	10.5-25.5%	25.5-47.5%	47.5-80%	>80%

8.6.2. Using Priority Weights to Determine LOS scores for Performance on Maintenance Goals

The goal-level priority weights can be used to combine LOS scores to determine composite score for the goal categories. These LOS scores for the goal categories form a report card indicating how well the agency is meeting its maintenance goals. The report card is an assessment of overall maintenance performance and can be used to communicate the agency's progress in meeting its targets.

An example for using the priority weights and to compute a composite LOS score is shown in Table 35 for the features in the critical safety category. The estimated deficiency rate of each feature, $\hat{p}$, is $\hat{p}^A$, $\hat{p}^B$,or $\hat{r}^C$ depending upon the estimator type associated with the deficiency measurement of the feature (see Chapter 7 of this Report). All features in each category follow the same LOS grading scale, so the composite deficiency rate will use that same scale too. The composite goal-level deficiency rate $\hat{P}$ is the weighted sum of the feature deficiency rates $\hat{p}_f$ and is computed as shown in

Equation 26 where ω_f are the goal-level priority weights and the subscript f indicates the summation is over all features in the category.

Equation 26. Composite LOS Deficiency Rate for a Goal Category

$$\hat{P} = \sum_f \hat{p}_f \omega_f$$

For the example in Table 35, goal-level priority weight is multiplied by the LOS deficiency rate for each feature. The sum of the weighted deficiency rates is the composite deficiency rate for the goal category. When considering all features in the critical safety goal, the composite deficiency rates for the critical safety are 4.81 and 5.79 percent in years 2010 and 2011, respectively. By referring to the LOS grading scale for critical safety in Table 34, the equivalent LOS grades are B and C in years 2010 and 2011, respectively.

Table 35. Using Priority Weights to Score Maintenance Performance in Goal Categories

		2010		2011	
Critical Safety Feature	**Goal-level Priority Weight ω_f**	**LOS Deficiency rate $\hat{p}_f$**	**Weighted deficiency rate $\hat{p}_f\omega_f$**	**Deficiency rate $\hat{p}_f$**	**Weighted deficiency rate $\hat{p}_f\omega_f$**
Emergency repair of regulator / warning signs	0.42	1	0.42	3	1.26
Hazardous debris	0.24	8	1.92	7	1.68
Protective barriers	0.13	1	0.13	5	0.65
Centerline markings	0.09	7	0.63	6	0.54
Edgeline markings	0.07	8	0.56	7	0.49
Unpaved shoulder drop-off / buildup	0.03	37	1.11	37	1.11
Paved shoulder drop off / build up	0.02	2	0.04	3	0.06
Column Sum	**1.00**		**4.81**		**5.79**

The method can be used to determine the composite deficiency rates and letter grades for all the goal categories. The deficiency rates and grades can be assembled into a performance report card as shown in Table 36. Comparing deficiency rates from year to year show trends that may not be evident in the letter grades.

Table 36. Using Composite Deficiency Rates to Prepare a Report Card on Performance toward Maintenance Goals

	2010		2011	
Goal Category	**Composite Deficiency Rate (% of inventory) $\hat{P}$**	**LOS Grade**	**Composite Deficiency Rate (% of inventory) $\hat{P}$**	**LOS Grade**
Critical Safety	4.81	B	5.79	C
Safety/Mobility	7.79	B	7.30	B

	2010		2011	
Goal Category	**Composite Deficiency Rate (% of inventory) $\widehat{P}$**	**LOS Grade**	**Composite Deficiency Rate (% of inventory) $\widehat{P}$**	**LOS Grade**
Stewardship	12.43	B	13.14	B
Ride/Comfort	7.36	A	8.26	B
Aesthetics	40.42	C	42.25	C

8.6.3. Using Priority Weights to Determine an LOS score for Program Performance

The next step is to roll up the goal-level grades to create a composite program-level grade. A simple method is to weight the composite deficiency rate for each goal and then sum them. This simple approach will produce a meaningful program-level deficiency rate if all goals in the program follow the same LOS grading scale. However, for this example, having different LOS grading scales for each goal category, the simple weighting method will not work because we would not know how to interpret the resulting composite deficiency rating.

For the case when the goal categories use different LOS grading scales, the composite program level grade is the sum of the weighted category grades (not the weighted deficiency rates). Converting deficiency rates to equivalent grades brings all the category score to the same scale. In Table 36, the composite deficiency rate for each goal was converted to a letter grade. Those letter grades can be converted to numeric grade point values on a common scale such as shown in Table 37. Then the grade point values can be combined to determine the composite program level grade point. The method for combining the grade point values is the weighted sum.

Table 37. Common Numeric Equivalents for Letter Grades

Letter Grade	Numeric Equivalent
A	3 to 4
B	2 to 3
C	1 to 2
D	0 to 1
F	0

The steps for determining the composite program-wide grade are as follows:

1. Refer to Table 37 to assign a numeric value for each category letter grade. There are two ways to assign a numeric value.
    a. Simply assign a numeric value equal to the middle of the grade point range.
    b. Interpolate. The interpolation method will find a value, X, on grade point scale that is equivalent to the effective deficiency rate $\hat{P}$ on the LOS scale. The interpolation method is more precise but requires more effort. The basic formula for the interpolation is Equation 27 where LOS_u and LOS_l are the upper and lower bounds of the LOS range for the assigned letter grade and G_u and G_L are the upper and lower bounds of the grade point range for the assigned letter grade. The interpolation

equation is written for mapping increasing deficiency rates to decreasing grade points.

Equation 27. Interpolation to Find an Equivalent Grade Point

$$\frac{LOS_u - \hat{P}}{LOS_u - LOS_l} = \frac{G_l - X}{G_l - G_u}$$

2. Find the sum of the weighted category grade points by multiply the priority weight and grade point for each category and taking the sum. Table 38 shows the steps and results. The program level grade point is 2.55 using the simple method compared to 2.31 using the interpolation method. If the category weight is high and category deficiency rate is close to the top or bottom of the LOS grade range, the interpolation method is preferred.
3. The program-wide grade point may be transformed to an equivalent letter grade using Table 37. For the example, the program-wide grade is the B for both methods. Both program level grade points 2.55 and 2.31 fall within the B range in Table 37.

Table 38. Using Priority Weights to Measure and Report Program-wide Performance

				Simple Method		Interpolation Method		
Goal Category	**Priority Weight**	**Composite Deficiency Rate $\hat{P}$**	**Category LOS Grade**	**Grade Point**	**Weighted Grade Point**	**Interpolation Formula**	**Grade Point (X)**	**Weighted Grade Point**
Critical Safety	0.52	4.81	B	2.5	1.30	$\frac{5.5-4.81}{5.5-2.5}=\frac{2-X}{2-3}$	2.23	1.16
Safety/Mobility	0.28	7.79	B	2.5	0.70	$\frac{9.5-7.79}{9.5-4.5}=\frac{2-X}{2-3}$	2.34	0.66
Stewardship	0.13	12.43	B	2.5	0.33	$\frac{15.5-12.43}{15.5-6.5}=\frac{2-X}{2-3}$	2.34	0.30
Ride/Comfort	0.05	7.36	A	3.5	0.18	$\frac{7.5-7.36}{7.5-0}=\frac{3-X}{3-4}$	3.02	0.15
Aesthetics	0.03	40.42	C	1.5	0.05	$\frac{47.5-40.42}{47.5-25.5}=\frac{1-X}{1-2}$	1.32	0.04
Program-level Score				**2.55 (B)**			**2.31 (B)**	

8.7. Relating Maintenance Cost to the LOS Rating Scale

Maintenance unit costs must be related to maintenance features. Furthermore, this must be done in a way that enables an estimate of cost to address maintenance needs. Cost and the LOS scale are related through the marginal cost; i.e., the maintenance cost for each increment of condition on the LOS scale. Section 7.5 of this Report illustrates two methods for estimating the cost to address deficiency at the feature level.

Marginal cost is the maintenance cost associated with each increment of condition on the LOS scale. The cost for maintenance to address deficiencies in one percent of the total feature inventory is also the marginal cost for a one percentage point improvement on the LOS scale for the feature.

Since the LOS scale is based on the percentage of inventory in deficient condition, the estimated cost for maintenance should also be expressed in terms of percentage of inventory. To illustrate the concept of marginal cost, Table 39 lists some examples of maintenance costs to address one percent

of the total inventory of various features. These costs are computed from the unit cost for maintenance and the number of units in one percent of the inventory. Expenditure of those amounts would lead to a one percentage point decrease in the features' deficiency rates. The cost for maintenance to address deficiencies in one percent of the total feature inventory is the marginal cost for a one percentage point improvement on the LOS scale for the feature.

Table 39 also lists the marginal costs on the goal and program level LOS scales. The feature unit costs and the priority weights for the features are used to compute the marginal costs. The goal-level marginal cost is the feature unit cost divided by the goal-level priority weight. Similarly, the program-level marginal cost is the feature unit cost divided by the global priority weight.

The marginal costs are expressions of the efficiency of maintenance activities. These are the required expenditures to gain a one percentage point improvement in the LOS rating at the goal and program levels. Removal of hazardous debris has the lowest marginal cost. This means that spending for removal of hazardous debris will yield a greater improvement in the goal or program LOS score than the same spending for any other feature in the critical safety category. The marginal costs are very important for allocating budgets if maximizing maintenance performance is the desire. The marginal costs are also useful for estimating the cost to improve specific LOS scores.

Table 39. Estimated Marginal Costs to Reduce 1 Percent Deficiency on the LOS Rating

Critical Safety Feature	Cost 1% of feature inventory	Priority Weight		Marginal Cost 1% on LOS rating scale		
		Local (goal)	Global (program)	Feature-level	Goal-Level	Program-Level
Reg/Warning Signs (emergency repair)	$272,660	0.42	0.218	$272,660	$649,190	$1,250,734
Hazardous Debris	$131,928	0.24	0.125	$131,928	$549,700	$1,055,424
Protective Barriers (Beam Guard)	$1,272,851	0.13	0.068	$1,272,851	$9,791,162	$18,718,397
Centerline Markings	$85,199	0.09	0.047	$85,199	$946,656	$1,812,745
Edge line Markings	$234,626	0.07	0.036	$234,626	$3,351,800	$6,517,389
Drop off/Build up (unpaved shoulder)	$71,343	0.03	0.016	$71,343	$2,378,100	$4,458,938
Drop off/Build up (paved shoulder)	$1,565,348	0.02	0.010	$1,565,348	$78,267,400	$156,534,800
Total		**1.00**	**0.52**			

With the information generated through the processes outlined in this chapter, the analyst now has the tools to generate and evaluate schemes of actions, to set targets and to develop programs. Chapter 9 of this report further develops this process.

9. Setting Targets

9.1. Objectives

The process for setting attainable targets involves analytical or experiential procedures to determine optimal combinations of actions (schemes) that are affordable and move the maintenance program in the direction of its goals. Accordingly, the research team developed both approaches for setting targets. The group approach that draws upon the experience of program decision-makers to evaluate tradeoffs and the more formal analytical approach that uses a linear programming model to set constraints and optimize tradeoffs.

Both approaches consider the cost of maintenance activities and the effectiveness of those activities. Both approaches are predicated upon the availability of certain information:

- The percentage of inventory that is in deficient condition. The percentage of inventory deficient comes from the agency condition assessment. If condition assessment relies on stratified sampling, the statistical analysis tools in Chapter 8 may be used to estimate the deficiency rates.
- Estimated cost to eliminate deficiencies in one percent of the inventory. The cost of treating one percent of deficiencies comes from the analysis performed in Section 7.6.
- Effectiveness of maintenance activities at the goal- and program-levels. The priority weights derived from comparison judgments of which feature, or action, contribute most greatly to the goals of the program. The tools presented in Chapter 8 may be used to assess cost and effectiveness.

This chapter presents the work done by the team to develop each of the approaches. After developing the approaches separately, the research team came to recognize that the two approaches have complementary strengths. The judgment of agency decision makers drives the process while the analytical tool assists by optimizing performance and resource allocation and by answering "what if" questions. In preparing the Guide, the research team blended the experiential and analytical approaches.

9.2. Group Approach for Determining Attainable Targets

This process can be used to select a set of actions, or schemes that will best move the program in the desired direction, following these steps:

1. **Assemble the program decision makers.** This is a group of people who are most knowledgeable in the program, who have to implement the program, and who have some decision-making authority over it. This group should then work through the Steps 2-10.
2. **Consider the current allocation of funds.** Compare the current distribution of funds with the information contained in your agency's versions of
3. Figure 24. Are dollars being allocated to the program areas that are rated as the highest contributors to the overall goals of the program? If not, some reallocations will be in order.
4. **Agree upon minimum acceptable LOS ratings for highway features.** Compare the conditions or LOS grades of the various features in your agency's inventory to the contribution of the features shown in your agency's versions of
5. Figure 24. Features rated as the highest contributors, or as the most important should be in the best condition. Lower priority features should not be expected to be in equal or better

condition that highest priority items. If importance and condition are not in alignment, some reallocation will be in order.

6. **Allocate resources to the most importance goal first.** Look at the features in the most important goal category in the inventory. In our example, this would be critical safety. Determine the total cost to raise the condition of the features in the most important category to the minimum acceptable LOS rating. Total cost includes both the incremental cost and the cost to maintain the ongoing level of effort .
7. **Allocate resources to the other goals in order of decreasing importance.** After the most important goal has been adequately funded, move on to the other goal categories in their order of importance. For each, determine the total costs to attain the minimum acceptable level of service. Total cost includes both the incremental (or decremental) cost as well as the level of effort cost.
8. **Compare required total cost to the budget.** Sum the total costs of attaining the acceptable LOS in all categories and compare that total to the probable budget. Be sure that all costs are considered on both sides of the comparison. If the required total cost is less than the budget go to Step 9, but first check your math and your assumptions, because that would be an unlikely outcome. If the required total cost is greater than the probable budget, then continue with Step 7.
9. **Use marginal costs to adjust expectations for the least important goal category first.** Review the least priority category, focusing on the features with the greatest marginal cost first. Determine if a reduction in the minimum LOS for the highest cost feature is acceptable. If so recalculate the required resources for the category and total. If the required total cost is still greater than budget move to the next feature with the next highest marginal cost and determine if the minimum LOS can be reduced further. If so, recalculate again. Repeat the process, looking at all features in sequence. If the required total cost is less than the budget then go to Step 9. If the required total cost is still greater than the probable budget, then continue with Step 8.
10. **Use marginal costs to adjust expectations for the other goal categories.** When the LOS expectations and total cost for the least priority goal can be reduced no further, move to the other goal categories in the order of increasing importance. For each, follow the procedure in Step 7 to reduce the minimum acceptable LOS and required total cost. If after each category has been fully examined and the total cost is less than the budget, then go to Step 9. Otherwise, follow the procedure in Step 7 to examine the next more important category. If the procedure is completed for all categories and the required total cost is still greater than the probable budget. Then return to Step 7 and repeat the process of using marginal costs to reduce the minimum expectations.
11. **Create a summary evaluation the LOS targets.** The final set of minimum acceptable LOS scores for the features in each category are the attainable LOS targets constrained by the budget. A summary analysis that reviews the potential impacts due to special circumstances and interests is still required.

These steps reflect sound management decision making. The result provide detailed targets for the condition of each feature and goal as well as a means for estimating the associated inputs and outputs, all of which are necessary for the management of the program.

The group process, supported by sound data, can yield a good result but several iterations will be necessary before resource requirements are made to match available resources.

From the perspective of management decision making, Step 9 is obviously necessary. The research team worked to define the details of what would be involved in the summary analysis and how the analysis should be done. In doing so, the team came to realize that the summary analysis could be best accomplished by following the process outlined for risk assessment. Consequently, the Guide recommends agencies follow a risk assessment procedure to evaluate LOS targets before implementing them.

9.3. Linear Programming Model for Determining Attainable Targets

The research team set out to develop a programming model that accomplishes the work of the group approach. The resulting model is useful, but cannot replace the judgments of experienced decision makers. The strength of the linear program approach is that it can determine optimal targets for features, or actions, that will maximize the performance within a given budget constraint. The programming model provides a decision model that is repeatable. The model can be made to reflect the realities of a specific agency or budget by imposing various constraints. The approach forces that agency to recognize budget constraints and to establish minimum expectations. Inherent in model is a mechanism for recording the decision constraints so they can be tracked from year to year.

The research team developed the model to maximize performance given budget constraints. The model formulation is presented in this section.

9.3.1. Maintenance Performance

The ability to measure performance is important for maintenance assessment and a necessary input for setting attainable targets. Performance is the percentage of inventory that is not deficient (in good condition). If measured in the units used for LOS, then performance is the complement of the target. The analytical hierarchy defines the priority weights for how maintenance features contribute to the agency's performance on maintenance goals and for how performance on maintenance goals contribute program performance. By using the LOS units, maintenance performance can be measured at the feature, goal, and program levels.

For a *single feature*, performance is the percentage of inventory that is not deficient.

$$\text{Performance}_{\text{feature}} = (100 - x_i)$$

Where, x_i is the target percentage of inventory for feature *i* that is allowed to be deficient.

For a *goal category*, performance is the weighted sum of the feature performances.

Equation 28. Performance on a Maintenance Goal

$$\text{Performance}_{\text{goal}} = \sum_{i}^{n} (100 - x_i)\, \omega_i$$

Where, $x_i$is the target percent of inventory for feature *i* that is allowed to be deficient; ω_i is the *goal-level priority weight* for feature *i*; and n is the number of features in the goal category.

At the *program level*, performance is the weighted sum of the goal category performance.

$$\text{Performance}_{\text{program}} = \sum_{j}^{m} \left(\sum_{i}^{n_j} (100 - x_i)\, \omega_i \right) \omega_j$$

Where, $x_i$is the target percent of inventory for feature *i* that is allowed to be deficient; ω_i is the goal-level priority weight for feature *i*; ω_j is the priority weight for goal category *j*; n_j is the number of features in goal category *j;* and m is the number of goal categories contributing in the program.

Using *global priority weights*, the *program level* performance is the weighted sum of all feature performances. The program level performance (Equation 29) can be written in the same form as the goal-level performance (Equation 28).

Equation 29. Performance of the Maintenance Program

$$\text{Performance}_{\text{program}} = \sum_{i}^{n} (100 - x_i)\, \omega_i$$

Where, $x_i$is the target percent of inventory for feature *i* that is allowed to be deficient; $\omega_i$is the *global priority weight* for feature *i*; and n is the total number of features in all goal categories.

Equation 28 and Equation 29 have the same form but the parameters are different depending upon whether the goal or program level performance is desired. The equation for performance forms the basis for a linear programming model. The objective is to maximize performance.

9.3.2. Determining Targets to Maximize Performance

Simple linear programming forms the basis for applying maintenance effectiveness and cost to estimate resource needs and to allocate available budgets.

Generalized from Equation 28 and Equation 29, the objective function is to maximize performance,

$$\max \sum_{i}^{n} (100 - x_i)\, \omega_i$$

Where, x_i is the target percent of inventory for feature *i* that is allowed to be deficient; ω_i is the priority weight for feature *i*; and n is the number of features.

The target deficiency rate is related to the current deficiency rate and desired decrease:

$$x_i = x_i' - \Delta x_i$$

Then, the objective function can be re-written as:

$$\max \sum_{i}^{n} (100 - x_i')\omega_i + \sum_{i}^{n} \Delta x_i\, \omega_i$$

Where, x_i' is the current deficiency rate and Δx_i is the decrease in deficiency rate to achieve the target for feature *i*. Note that Δx_i is positive if the deficiency rate is reduced; Δx_i is negative if deficiency rate is allowed to increase. Since the first term is a constant, it can be ignored in the linear programming model.

The target deficiency rate $x_i$is the current deficiency rate x_i' minus the amount the rate is decreased , Δx_i. If deficiency is allowed to increase, then Δx_i is a negative number.

$$x_i = x_i' - \Delta x_i$$

The following basic constraints must be built into the model. This constraint says the upper bound of deficiencies that can be treated is limited to the deficiencies that exist in the inventory. No more deficiencies can be treated than exist.

$$\Delta x_i \leq x_i'$$

Another constraint recognizes that the agency does not have the ability to completely abandon certain maintenance requirements by limiting the target deficiency rate to be no more than a maximum acceptable rate. This constraint is useful if certain features are to be maintained at the least to a minimum level of service. The minimum LOS is associated with a maximum acceptable deficiency rate, $\max(x_i)$. The following non-negativity constraint requires the target to be no more than a maximum acceptable rate.

$$0 \leq x_i \leq \max(x_i)$$

The model may be constrained to require the target deficiency rate to stay the same or be reduced, but not increase. The following constraint will not allow the deficiency rate to increase by requiring the model to find a target that is less than or equal to the current rate.

$$x_i \leq x_i'$$

9.3.3. Cost Function and Level of Effort for the Status Quo

The maintenance program performance is constrained by available resources. The optimization model is constrained by maintenance costs and those costs cannot exceed the budget. Equation 30 shows the cost function and budget constraint.

Equation 30. Maintenance Program Cost Function

$$\sum_{i}^{n} \left(\frac{100}{t_i} + \Delta x_i \right) c_i \leq \text{Budget}$$

Where, t_i is the maintenance cycle time in years for feature *i*; c_i is the cost to address deficiency in one percent of the inventory for feature *i*, and $\Delta x_i$is the decrease (or increase) in deficiency rate to achieve the target for feature *i*. If deficiency is allowed to increase, then Δx_i is a negative number.

So much of maintenance programming tends to focus on how the ongoing resources are being spent, but to move the program in directions that optimize performance, the manager needs to clearly understand tradeoffs and performance impacts of incremental increases or decreases in resource allocations. The level of effort needed to maintain the status quo is the annual quantity of each feature that receives maintenance. Equation 31 is the cost model for the maintaining the status quo level of effort on feature *i*.

Equation 31. Cost Model for Maintaining the Status Quo

$$\left(\frac{100}{t_i} \right) c_i$$

The total inventory divided by the estimated average cycle time yields the expected portion of the inventory that should be serviced each year to keep the deficiency rate in steady state. The average cost for one percent of the inventory can then be applied to determine the budget requirements to maintain the status quo.

Table 40 is a sample tabulation of the feature level data that is required for the cost function. The table includes inventory and cost information from the Wisconsin Department of Transportation. Beam Guard is estimated to require maintenance service every 15 years; thus in Table 40, beam guard is assigned a cycle time of 15 year. The inventory is 3,704,457 linear feet. Therefore, in an average year 1/15th, or about 247,000 feet, of the inventory must be treated. The cost per linear foot is $34.36. To maintain the status quo, about $8.5 million is needed for beam guard maintenance each year. The maintenance cycle times were estimated using the method described in Section 9.3.4.

Table 40. Inventory, Cycle Times, and Cost for Maintenance of Roadway Features (based on Wisconsin DOT Compass Program)

Feature	Inventory	Units	Unit Cost	Cost (1% inventory) c_i	Deficiency rate (%) x_i'	Cycle Time t_i	% inventory for status quo $\frac{100}{t_i}$	Annual Cost for Status Quo $\frac{100}{t_i} c_i$
Reg/Warning Signs (emergency)	159,004	ea	$171.48	$272,660	3	20	5	$1,363,300
Hazardous Debris	11,774	CL	$1,120	$131,928	3	7	14	$1,884,681
Protective Barriers (Beam Guard)	3,704,457	LF	$34.36	$1,272,851	5	15	7	$8,485,676
Centerline Markings	56,799,150	LF	$0.15	$85,199	6	1	100	$8,519,873
Edgeline Markings	156,417,624	LF	$0.15	$234,626	7	1	100	$23,462,644
Drop off/Build up (unpaved shoulder)	21,619	mi	$330	$71,343	37	4	25	$1,783,568
Drop off/Build up (paved shoulder)	21,591	mi	$7,250	$1,565,348	3	15	7	$10,435,650
Woody Vegetation Vision	39,117	ea	$258.35	$101,059	1	3	33	$3,368,626
Mowing Vision	39,117	ea	$83.41	$32,627	1	1	100	$3,262,749
Special Pavement Markings	48,910	ea	$169.60	$82,951	10	1	100	$8,295,136
Woody Vegetation	29,625	LM	$1,033	$306,145	2	6	17	$5,102,413
Clean Culverts	36,266	ea	$226	$81,961	22	10	10	$819,612
Clean Storm Sewers	48,926	ea	$115	$56,265	17	5	20	$1,125,298
Cross slope (unpaved shoulder)	21,619	mi	$2,000	$432,380	27	12	8	$3,603,167
Delineators	155,793	ea	$52	$81,012	25	8	13	$1,012,655
Reg/Warning Signs (routine)	159,004	ea	$123	$193,826	15	8	13	$2,422,823
Fences	14,169,357	LF	$6.28	$889,836	1	30	3	$2,966,119
Clean Ditches	18,236.48	mi	$8,000	$1,458,918	3	20	5	$7,294,592
Curb and Gutter (Clean)	3,396.91	mi	$141	$4,790	4	2	50	$239,482
Clean Flumes	11,631	ea	$37.35	$4,344	39	8	13	$54,302
Cracking (paved shoulder)	21,591	mi	$880	$190,001	60	15	7	$1,266,672
Erosion (unpaved shoulder)	21,619	mi	$1,164	$251,796	2	7	14	$3,597,093
Under/edge Drains (Clean)	33,424	ea	$16.01	$5,351	33	4	25	$133,780
Potholes/Raveling (paved shoulder)	21,591	mi	$1,130	$243,978	6	15	7	$1,626,522
Other Signs (emergency)	122,970	ea	$171.48	$210,869	4	20	5	$1,054,345
Other Signs (routine)	122,970	ea	$121.90	$149,900	39	10	10	$1,499,004
Mowing	29,625	LM	$83.41	$24,710	38	1	100	$2,471,021
Litter	29,625	LM	$547.03	$162,058	63	1	100	$16,205,764
TOTAL								**$123,356,564**

Abbreviations: CL = centerline miles, LF = linear feet; mi = miles; SF = square feet; ea = each; LM = lane miles

9.3.4. Estimating Maintenance Cycle Time

The concept of maintenance cycle time is useful for estimating the level of effort to maintain the current service levels. Each feature has its own cycle time because some features required maintenance on shorter intervals than others.

For many agencies, the cycle times will need to be estimated. To do this, some information is required:

- The inventory, either an actual count or a good estimate, for each feature for which targets are being set.
- Actual units accomplished for the past year, better still for the past three years, at the feature level.
- The expert judgment of people familiar with the activities of the program in the past.

The cycle time for actions at the feature level is the interval at which each segment of the inventory must be treated to maintain steady state service levels. For some actions the cycle time might be 0.5—twice a year—for others it might be 20—once every twenty years.

The first step in this process is assembling a group of people who are familiar with the maintenance program. Field managers or foreman are preferred. These are the people with the closest hands-on knowledge of the program and also a sound knowledge of state or regional budgets.

If actual units of activity are available for one or more years, the experts should be asked to review that information and agree on what is the average level of effort required to maintain the status quo. With this agreed information, a cycle time for each feature can be calculated by dividing the inventory by the annual average level of effort.

If the actual number of accomplished units is not available for past years at the feature level, the experts should be asked to estimate of cycle time for each feature. As estimates are made, all participants must recognize that these are estimates for the entire state. Urban areas with high traffic volumes will be very different for many items than rural, low traffic volume areas. Different terrain and weather conditions that may exist in different areas will also tend to make the estimated differ. Indeed, if the differences within the state are pronounced, consideration should be given to doing estimates at a regional level and aggregating them to the state level.

The method was used to estimate the cycle times in Table 40. There are two simple checks that may be helpful for evaluating whether the cycle times estimates are reasonable.

1. If deficiency rates are somewhat constant from year to year, then the agency may assume annual expenditures at the feature level are maintaining the status quo. The following equation shows how the cycle time may be estimated using a ratio of inventory or expenditures. If the ratios are mathematically close, say within 10 percent of each other, then the cycle time estimate is reasonable. If not, then the experts should be asked to reconsider the cycle time estimates.

$$\frac{\text{inventory}_i}{\text{average (level of effort)}_i} \approx \frac{100c_i}{\text{average (annual expenditures)}_\text{i}} \approx t_i$$

   The equation shows how the cycle time may be computed by dividing the expected cost, if the entire feature inventory is deficient, by the average annual expenditure for that features.

2. As shown in Table 40, the cycle time is used to estimate the budget requirement for the status quo. The total estimated budget required for all features should be used as a first-

level reality check. If the amount is within 10 percent of the actual expenditures, then on average, the cycle time estimates are reasonable. If not, then the cycle times have probably been underestimated and the experts should be asked to consider them a second time.

9.3.5. Implementing the Linear Programming Model

The model formulation can be easily implemented as an Excel worksheet using the built in solver tool. Figure 23 shows the worksheet with data from Wisconsin. The solver tool allows the equations to be solved iteratively and finds a set of target deficiency rates that maximize total performance. The data cell and columns in Figure 23 are labeled to correspond to the notation used in the model formulation.

If state-level policymakers want to add funding to critical safety, or any other goal, the linear programming model can be used to focus on that single goal to recommend the best use of those dollars. The example in Figure 23 focuses on the features in the critical safety goal category and the statewide inventory for those features. Some of the data for this example comes from the Wisconsin Compass program. Most of the maximum acceptable deficiency rates are slightly higher than the current rates providing the opportunity for reallocate funds and possible savings. Total cost for the status quo is $72.9 million while the recommended allocation of funds totals $69 million; a savings of $3.9 million that could be directed to address deficiencies in the other maintenance goals. By re-allocating funds among the features the effective LOS grade for the critical safety goal can be increased from C to B. The funding allocation would improve the LOS score for emergency signs from B to A, and for centerline markings from C to A. The deficiency rates for hazardous debris, edgeline markings, and paved shoulder would be allowed to increase slightly so that funds can be shifted toward efforts to bring the deficiency rate of unpaved shoulders from 37 percent down to the maximum acceptable rate of 25 percent.

The model could also be applied to a subset of the agency such as a district or region to find the most beneficial use of funds for that area. If so, then the scope of the inventory would be limited to the inventory in the district or region. Similarly, the current deficiency rates would be based on the inventory sample from that district or region. If desired, estimated maintenance costs may also be specialized for the district or region.

LOS Target Setting for CRITICAL SAFETY

Directions: Enter available budget, deficiency rates, and maximum acceptable defiency rates.
Select Solver from the Data Menu. Press Solve in the Solver dialog box.

Budget =	$69,000,000	Estimated Cost =	$68,993,731
Current effective deficiency rate =	5.79	Current LOS Grade =	C
Target deficiency rate reduction =	1.48		
Target effective deficiency rate =	4.31	Target LOS Grade =	B

	PRIORITY, COST &CONSTRAINTS							CURRENT		TARGET		
FEATURE	ω_i Priority weight	Inventory	Inventory unit	Unit Cost	t_i Cycle Time	c_i Cost Deficiency in 1% Inventory	$\max(x_i)$ Maximum Acceptable Deficiency Rate	x'_i Current Deficiency rate	LOS Grade	x_i TARGET Deficiency rate	TARGET LOS Grade	TARGET Budget Allocation
Emergency repair of regulator / warning signs	0.42	159,004	ea	$171.48	20	$272,660	5	3	B	0	A	$2,181,280
Hazardous debris	0.24	117,740	CL	$1,120.50	7	$1,319,277	9	7	C	9	C	$16,208,257
Protective barriers	0.13	3,704,457	LF	$34.36	15	$1,272,851	5	5	B	5	B	$8,485,676
Centerline markings	0.09	56,799,150	LF	$0.15	1	$85,199	6	6	C	1	A	$8,945,866
Edgeline markings	0.07	156,417,624	LF	$0.15	1	$234,626	8	7	C	8	C	$23,228,017
Unpaved shoulder drop-off / buildup	0.03	21,619	mi	$330	4	$71,343	25	37	F	25	F	$2,639,680
Paved shoulder drop off / build up	0.02	21,591	mi	$7,250	15	$1,565,348	5	3	B	5	B	$7,304,955
												$68,993,731

	STATUS QUO		GOAL LEVEL WEIGHTED RATES			
FEATURE	$\left(\frac{100}{t_i}\right)$ Status quo % inventory per year	$\left(\frac{100}{t_i}\right)c_i$ Cost for status quo LOS	$\omega_i x'_i$ Weighted Current Deficiency Rate	Δx_i Deficiency Rate Reduction	$\omega_i \Delta x_i$ Weighted Rate Reduction	$\omega_i x_i$ Weighted Target Deficiency Rate
Emergency repair of regulator / warning signs	5	$1,363,300	1.26	3	1.26	0
Hazardous debris	14	$18,846,810	1.68	-2	-0.48	2.16
Protective barriers	7	$8,485,676	0.65	0	0.00	0.65
Centerline markings	100	$8,519,873	0.54	5	0.45	0.09
Edgeline markings	100	$23,462,644	0.49	-1	-0.07	0.56
Unpaved shoulder drop-off / buildup	25	$1,783,568	1.11	12	0.36	0.75
Paved shoulder drop off / build up	7	$10,435,650	0.06	-2	-0.04	0.1
		$72,897,520	5.79		1.48	4.31

Figure 23. A Simple Spreadsheet Optimization Tool for Setting LOS Targets and Allocating Maintenance Funds

The research team worked to understand how the basic optimization model can be used to address specific questions. Based on the teams experiences, maintenance managers are concerned with issues in a particular jurisdictional areas or program wide performance. Managers are often called upon to address issues or answer questions related to specific maintenance activities or goals. With this understanding the research team identified the two scoping parameters of goal and inventory. The goal scope could be program-wide or a single goal. The inventory scope could be single region or state-wide. Table 41 shows how the scoping parameters define the appropriate features, priority weights, and deficiency rates to be used in the analysis for setting LOS targets.

Table 41. Scope of Goal and Inventory for Setting LOS Targets

Goal Scope	Inventory Scope	Features to include in the analysis	Priority Weights to be used	How to estimate the deficiency rates
Analysis of a single goal	Single region (single stratum)	Goal features	Goal-level	Based on stratum subset of the sample
	State-wide (all strata)	Goal features	Goal-level	Based on full sample
Program-wide (analysis of all goals)	Single region (stratum)	All features	Global weights	Based on stratum subset of the sample
	State-wide (all strata)	All features	Global weights	Based on full sample

9.4. Setting the Attainable LOS Targets

Either of the methods in Section 9.5 can be used to determine attainable targets that fall within budget and performance constraints. In practice, the target setting process is most effective using a blended approach; one that draws upon expert knowledge to set minimum acceptable performance levels and uses the spreadsheet tool for an iterative optimization of tradeoffs.

Whatever the method used the set of targets should be carefully reviewed before they become policy. There may be special situations or circumstances for which the targets are not acceptable. The risk assessment strategy in Chapter 4 of the Guide may be used to identify these special circumstances. For example, if the tentative program calls for a reduction in the effort associated with maintaining fences, the agency should consider risks when implementing the reduction policy. Certainly, sections of fencing that separate a freeway from a subdivision with children would prioritize over fencing along a cornfield.

The risk assessment should consider each feature in turn to identify risky circumstances. The specific areas of concern should be focused upon and additional tradeoffs or prioritization may be necessary.

Once the attainable targets are considered reasonable and supportable, then they should be communicated. The individual feature targets, the mitigation strategy for risky circumstances, and the effective targets for goal categories should be published and communicated to all, inside the agency and outside, who have a need to know or an interest in the program.

An implementation plan should accompany the communication of targets. The implementation plan would include the expected units of output necessary to reach the targets. Both the level-of-effort and the changes in effort must be expressed in the implementation plan. The plan should also include the estimated inputs and estimated budget allocation for achieving the targets.

The plan may include details on how resources are to be spent. For example, if shoulder patching cost estimates are based on a breakdown of labor, materials, and equipment costs, the plan should communicate the estimated tons of patch material, machine hours, and labor hours needed to achieve the specific number of output units to reach the target.

The plan should communicate the goals and output and input targets for each district or region based on the condition of the inventory and the total inventory of that district or region. These provide three points of reference for monitoring:

- Were the targeted conditions achieved?
- Were the planned units of output accomplished?
- Were the planned inputs consumed as planned?

9.5. Setting Desirable Targets

The desirable targets, though non-attainable, are useful for gap analysis. The difference between what is being accomplished and what should be accomplished is the performance gap. The desirable targets can be contrasted to attainable targets to define a gap between what is being accomplished and what should be accomplished.

The research team prepared an example of how the linear programming tool can be used to assist in computing the performance and monetary gap between the attainable and desirable target. The team prepared the example for inclusion in the Guide.

Typically, desirable standards are set based on accepted professional standards, although they may be politically motivated. Motivation notwithstanding, the basis upon which the desirable targets were established must be clear, and reasonable estimates must be made for both the incremental costs of attaining those targets and the incremental benefits that might be found. The concept of diminishing marginal returns may in fact set practical boundaries on what can be achieved.

Several methods can be used to set desirable targets. In many cases, the approach taken will be a combination of one or more of these:

Following established professional standards of practice. Many professional organizations define standards of practice that can be used as a basis for defining standards. Similarly, many manufacturers provide guidance as to how their products should be used and how often they should be replaced. Using a standard endorsed by AASHTO or 3M tends to lend some credibility to the standard. However, even if such organizations suggest some standard, the practitioner would do well to fully understand the basis upon which such recommendations are made.

Consensus. Typically, under this method, agency staff who are expert in a given area use such information as may be available to them—journal articles, experience, conference proceedings, etc.—and come to agreement on what a desirable standard would be. This approach has the advantage of getting the support of program staff and of using the expertise that is available within the agency. It may have the defect of being perceived as

the self-serving statements of people involved in the program. If it is used, it may be desirable to emphasize the resources that were called upon to arrive at the answers, which will make this approach seem similar to the first.

Benchmarking. Most transportation professionals belong to some kind of professional network. They tend to know which agency does a good job at X and which does a good job at Y. Through those network contacts, standards can be drawn from those outstanding performers. If this approach can be bolstered by testimonials from customers of those agencies, or if the customers in a given state agree that a neighboring state does an outstanding job in the area being benchmarked, the standard will likely have credibility.

Customer-responsive. Some states have very aggressive approaches to gathering the views of those who use the facilities. If road trips or focus groups tell the agency that customers value some features very highly, they might logically define a high desirable standard for that feature. This approach probably cannot be used alone since the general public may not fully comprehend the needs in some areas that do not immediately impact their driving experience. Measures related to the long life or structural integrity of the system may be overlooked until the problems become chronic.

Policy direction. Since policy makers often like to make decisions, this approach is often used. An agency head, policy board, or other policymaker dictates that the standard will be set at A. While this probably happens often, it is perhaps the least desirable approach since it begs the question of why A? Whenever possible, policy direction should be combined with and informed by one or more of the other methods.

It is important to demonstrate how the targets were developed, if your agency chooses to use desirable targets. Which of the methods was used? How do the desirable targets compare with targets of other agencies? The answers to these questions can add credibility to the targets and increase their value as points of comparison as policymakers consider the budget for maintenance and as the public attempts to understand the condition of the highway network.

9.5.1. Estimating the Benefit of Desired Targets

A key part of adding credibility to desirable targets is defining the benefit that might be derived from attaining those targets. The spreadsheet tool can calculate the relative benefits of desirable targets using the LOS scale. However, the effective deficiency rate will probably have little meaning to people outside of the agency. Therefore, a more direct, if less quantitative, approach may be desirable.

For example, if your desirable target for pavement is significantly higher than can be attained, why would it be desirable to move to a higher maintenance level? Can you estimate improved life for pavements with greater maintenance, and therefore a life-cycle cost savings? Even if the estimate is broad-brush and qualitative, it could provide a compelling argument for those who make decisions. Similarly, if desirable pavement marking targets are higher than can be attained, can the benefit be illustrated by customer preference surveys; or can the benefit of improved markings be shown by the number of leaving-the-road or bad weather crashes to which worn markings might have been a contributing factor?

Any credible information that illustrates the benefits of striving for higher targets can be useful to make the case for desirable targets. The key word is *credible*. The promise that crashes will be reduced by a certain percent if better pavement markings are placed can

lead to loss of credibility. If the reason for placing better markings is to improve safety and/or customer satisfaction, then discussing those items is appropriate.

9.5.2. Estimating Incremental Cost of Desirable Targets

The linear programming tool can be used to estimate the cost to achieve the desirable targets. Being able to estimate the cost is important if the agency wants to persuade policymakers and the public that greater investment is appropriate.

Figure 24 shows an example of how the spreadsheet tool can be used for estimating the cost of desirable targets. To estimate the cost of desirable targets, the deficiency rate of the desired targets should be entered as the maximum acceptable deficiency rate for the feature (in the max(x_i) column). This tells the solver to constrain the solution. A budget amount should be set to what is available, not what is desirable. Running the solver function will not find a solution but will estimate the cost for the desirable targets. The difference between the estimated target costs in Figure 23 and

Figure 24 is the incremental cost of the desirable targets. The incremental cost to achieve the desirable targets for all of the critical safety features is about $20.2 million ($89,161,387 - $69,993,731). If the agency is making the case for funding to improve edgeline markings and the desirable target is 2 percent deficiency, then the estimated funding gap to achieve the desired target is about $1.4 million $24,635,776 - $23,228,017). The estimated funding gap is the additional amount needed to achieve the desired target, it gap amount is above and beyond the cost to maintain the current level of effort.

LOS Target Setting for CRITICAL SAFETY

Directions: Enter available budget, deficiency rates, and maximum acceptable defiency rates.
Select Solver from the Data Menu. Press Solve in the Solver dialog box.

Budget =	$69,000,000	Estimated Cost =	$89,161,387
Current effective deficiency rate =	5.79	Current LOS Grade =	C
Target deficiency rate reduction =	3.79		
Target effective deficiency rate =	2.00	Target LOS Grade =	A

	PRIORITY, COST &CONSTRAINTS							CURRENT		TARGET		
FEATURE	ω_i Priority weight	Inventory	Inventory unit	Unit Cost	t_i Cycle Time	c_i Cost Deficiency in 1% Inventory	$\max(x_i)$ Maximum Acceptable Deficiency Rate	x_i' Current Deficiency rate	LOS Grade	x_i TARGET Deficiency rate	TARGET LOS Grade	TARGET Budget Allocation
Emergency repair of regulator / warning signs	0.42	159,004	ea	$171.48	20	$272,660	2	3	B	2	A	$1,635,960
Hazardous debris	0.24	117,740	CL	$1,120.50	7	$1,319,277	2	7	C	2	A	$25,443,194
Protective barriers	0.13	3,704,457	LF	$34.36	15	$1,272,851	2	5	B	2	A	$12,304,230
Centerline markings	0.09	56,799,150	LF	$0.15	1	$85,199	2	6	C	2	A	$8,860,667
Edgeline markings	0.07	156,417,624	LF	$0.15	1	$234,626	2	7	C	2	A	$24,635,776
Unpaved shoulder drop-off / buildup	0.03	21,619	mi	$330	4	$71,343	2	37	F	2	A	$4,280,562
Paved shoulder drop off / build up	0.02	21,591	mi	$7,250	15	$1,565,348	2	3	B	2	A	$12,000,998
												$89,161,387

	STATUS QUO		GOAL LEVEL WEIGHTED RATES			
FEATURE	$\left(\frac{100}{t_i}\right)$ Status quo % inventory per year	$\left(\frac{100}{t_i}\right)c_i$ Cost for status quo LOS	$\omega_i x_i'$ Weighted Current Deficiency Rate	Δx_i Deficiency Rate Reduction	$\omega_i \Delta x_i$ Weighted Rate Reduction	$\omega_i x_i$ Weighted Target Deficiency Rate
Emergency repair of regulator / warning signs	5	$1,363,300	1.26	1	0.42	0.84
Hazardous debris	14	$18,846,810	1.68	5	1.20	0.48
Protective barriers	7	$8,485,676	0.65	3	0.39	0.26
Centerline markings	100	$8,519,873	0.54	4	0.36	0.18
Edgeline markings	100	$23,462,644	0.49	5	0.35	0.14
Unpaved shoulder drop-off / buildup	25	$1,783,568	1.11	35	1.05	0.06
Paved shoulder drop off / build up	7	$10,435,650	0.06	1	0.02	0.04
		$72,897,520	5.79		3.79	2.00

Figure 24. Using the Spreadsheet Tool to Estimate Cost of Desirable LOS Targets

10. Implications for Future Research

The Guide and the research that brought it into being bring a number of new or enhanced concepts to maintenance performance management. Use of these concepts will improve the practice of maintenance management and maintenance performance measurement, but more work can be done. A number of research areas might be pursued based on the findings of this study:

- The linear programming tool recommended in the Guide could be refined, made much more user friendly, and developed into a tool that would more easily allow policy ideas and budget allocations to be tested. One could imagine links to condition files and inventories to produce output, defining areas of need, and predicting future conditions. The tool could also be expanded to provide targets for input as well as outputs. A more elegantly defined tool might also make the process more attractive to more agencies.
- More work could be done in refining an overall maintenance performance management process. Such a process would outline how managers should use various types of measures, how they should use them to communicate with policy makers, and how they should use them to improve the maintenance program.
- More work could also be done on improving the links between maintenance accounting systems and maintenance performance management. The goal of such an improved linkage would be to allow cost and accomplishment information to be collected in a manner that complements and feeds into the performance management system.
- Finally, much more work could be done to improve our understanding of how maintenance actions contribute to the larger strategic goals of an agency. Most people "understand" that better pavement maintenance will contribute to longer pavement life, but now one really understands the trade-offs between more maintenance and longer life. How much more is useful? How much more is marginally wasteful. Similar questions could be asked about safety, bridge management and many other maintenance activities. Better information is needed.

References

Adams, T.M. and J. Smith (2006). Synthesis of Measures for Highway Maintenance Quality Assurance. Paper No. MMC2006-051. In Proceedings, 11th AASHTO-TRB Maintenance Management Conference. Transportation Research Circular E-C098. Washington, DC.

Adams, Teresa M., and Smith, Janille (2005). Maintenance Quality Assurance-Synthesis of Measures. Midwest Regional University Transportation Center. Madison, WI.

Adams, Teresa, Scott Janowiak, Will Sierzchula, and Jason Bittner (2009). MQA-2 –Synthesis of Measures. Project 08-15 Report. Midwest Regional Transportation Center. University of Wisconsin-Madison.

American Association of State Highway and Transportation Officials (AASHTO) (2006). Asset Management Data Collection Guide: Task Force 45 Report. Washington, DC.

Applied Pavement Technology, Inc. (2011). Best Practices in Performance Measurement for Highway Maintenance and Preservation. NCHRP U.S. Domestic Scan Program–Scan 10-03. (NCHRP Project 20-68A). Transportation Research Board, Washington, DC.

Bell, Daniel (2005). Maintaining Standards: A Public Report on Trunk Road Maintenance in 2004/5. Scottish Executive Enterprise, Transport and Life Long Learning Department. (http://www.scotland.gov.uk/Publications/2005/09/2392802/28054 as of Jan. 27, 2012).

Bridges, Michael (2011). ERP Support for TAM at Louisiana DOTD. In, TAM Guide Webinar 4: Tools and Techniques for Implementing the TAMP. Slides 52-61. (http://tam.transportation.org/Documents/Webinars/TAM%20Guide%20Webinar%204%20draft%203.ppt as of Feb. 29, 2012).

Brown, L. D., Cai, T. T., & DasGupta, A. (2001). Interval Estimation for a Binomial Proportion. Statistical Science, 16(2), 101-133.

Cambridge Systematics and Applied Research Associates (2011). Technical Guidance for Deploying National Level Performance Measurements. NCHRP 20-24(37) G, Transportation Research Board, Washington, DC.

Cambridge Systematics, I. (2011). Uses of Risk Management and Data Management to Support Target-setting for Performance-based Resource Allocation by Transportation Agencies.

Cambridge Systematics, Inc. (1999). Multimodal Transportation: Development of a Performance-Based Planning Process. NCHRP Web Doc 26, Project 8-32(2)A): Contractors Final Report. Transportation Research Board, Washington, DC. (http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w26-a.pdf as of Jan. 28, 2012).

Cambridge Systematics, Inc. and High Street Consulting (2010). Transportation Performance Management: Insight from Practitioners. NCHRP Report No. 660. Transportation Research Board, Washington, DC.

Cambridge Systematics, Inc., Applied Research Associates, Arora & Assoc., KLS Engineering, PB Consult, and Louis Lambert (2009). An Asset-Management Framework for the Interstate Highway System. NCHRP Report No. 632. Transportation Research Board, Washington, DC

Cambridge Systematics, Inc., Boston Strategies International, Inc., Gordon Proctor and Associates, and Michael J. Markow (2010). Target-Setting Method and Data Management to Support Performance-Based Resource Allocation by Transportation Agencies. NCHRP Report No. 666. Transportation Research Board, Washington, DC.

Cambridge Systematics, Inc., Parsons Brinckerhoff Quade and Douglas, Inc., Roy Jorgensen Associates, Inc., and Paul D. Thompson (2002). Transportation Asset Management Guide: Final Report. NCHRP Project 20-24(11). Transportation Research Board, Washington, DC.

Dye Management, Paul Thompson, and Quality Engineering Solutions (2010). Development of Levels of Service for the Interstate Highway System. NCHRP Report No. 677. Transportation Research Board, Washington, DC.

Eigenvalue Excel vba software. Soft32Download.com. from http://eigenvalue-excel-vba.win.soft32download.com/

Engert, P. E., & Lansdowne, Z. F. (1999). Risk Matrix User's Guide. Bedford, MA: The MITRE Corporation.

Essex County Council (2008). Essex Highway Maintenance Strategy: Maintenance Policy and Standards. (http://www.essex.gov.uk/Travel-Highways/Report-a-highway-problem/Documents/Essex_Highway_Maintenance_Strategy_April_08.pdf as of Jan. 27, 2012).

Falkirk Council (2010). Carriageways Lifecycle Plan (Version 1.0) 2010/11 Falkirk. (http://www.falkirk.gov.uk/services/development/roads_services/asset_management/carriage_plan.pdf as of Jan. 27, 2012).

FHWA (Federal Highway Administration) (2004). Transportation Performance Measures in Australia, Canada, Japan, and New Zealand. Report No. FHWA-PL-05-001. Washington, DC.

Garvin, Michael, Keith Molenaar, Desiderio Navarro, and Gordon Proctor (2011). Key Performance Indicators in Public-Private Partnerships : A State of the Practice Report. FHWA-PL-10-029 FHWA / Office of International Programs; American Trade Initiatives, Alexandria, VA.

Godfrey, P. (1996). Control of Risk: A Guide to the Systematic Management of Risk from Construction. Sir William Halcrow and Partners, Ltd: 71, London, UK.

Gordon, Mark (2011). Levels of Service Overview. In, TAM Guide Webinar 3: The Transportation Asset Management Plan (TAMP). Slides 10-13. (http://tam.transportation.org/Documents/Webinars/TAM%20Guide%20Webinar%203%20draft%202.ppt as of Feb. 29, 2012).

Gordon, Mark, George JasonSmith, Paul D. Thompson, Hyun-A Park, Frances Harrison, and Brett Elston (2011). AASHTO Transportation Asset Management Guide – A Focus on Implementation. American Association of State Highway and Transportation Officials, Washington, DC.

Haas, R, G. Felio, Z. Lounis, and L. Cowe Falls (2009). Measurable Performance Indicators for Roads: Canadian and International Practice. NRCC-52692. National Research Council of Canada. (http://www.nrc-cnrc.gc.ca/obj/irc/doc/pubs/nrcc52692.pdf as of Jan. 27, 2012).

Hoffman, Gary L, Amar Bhajandas and Jagannath Mallela (2010). Issues and Practices in Performance-Based Maintenance and Operations Management. Prepared for American Association of State Highway and Transportation Officials (AASHTO) (prepared as part of NCHRP Project 20-24(61)). Transportation Research Board, Washington, DC.

Hyman, William A. (2004). Guide for Customer-Driven Benchmarking of Maintenance Activities. NCHRP Report 511 (Project 14-13). Transportation Research Board, Washington, DC.

Hyman, William A. (2009). Performance-Based Contracting for Maintenance. NCHRP Synthesis Report 389. Transportation Research Board, Washington, DC.

Institute of Risk Management (2010). A structured approach to Enterprise Risk Management (ERM) and the requirements of the ISO 31000. London, UK.

Korn, E. L., & Graubard, B. I. (1998). Varience Estimation for Superpopulation Parameters. Statistica Sinica, 8, 1131-1151.

Lemlin, Marc (1998). Development of Tools for Performance Measurement, World Road Association. (http://rru.worldbank.org/documents/toolkits/highways/pdf/95.pdf as of Jan. 27, 2012).

Maze T.H., Chris Albrecht, Dennis Kroeger, and Jon Wiegand (2007) Performance Measures for Snow and Ice Control Operations. NCHRP Web-Only Doc. 136, pages 135–153. Contractor's Final Report for NCHRP Project 6-17, Transportation Research Board. (http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w136.pdf as of Jan. 28, 2012).

Morin, Pat (2011). Washington State DOT's Legislative Context for Asset Management. In, TAM Guide Webinar 3: The Transportation Asset Management Plan (TAMP). Slides 21-30. (http://tam.transportation.org/Documents/Webinars/TAM%20Guide%20Webinar%203%20draft%202.ppt as of Feb. 29, 2012).

National Performance Review (NPR) (1997). Serving the American Public: Best Practices in Performance Measures. Benchmark Study Report. Washington, DC.

National Research Council (1995). Measuring and Improving Infrastructure Performance. Committee on Measuring and Improving Infrastructure Performance. Washington, DC.

OECD (Organization for Economic Cooperation and Development) (2001). Performance Indicators for the Road Sector: Summary of Field Test. OECD Publications Service. (www.internationaltransportforum.org/Pub/pdf/01PerformIndicE.pdf as of Jan. 27, 2012).

PB Consult, Inc. Price Warehouse Coopers, LLC, Cambridge Systematics, and Nustats, Inc. (2004). A Review of DOT Compliance with GASB 34 Requirements. NCHRP Report 522. Transportation Research Board, Washington, DC.

Proctor, G. (2012). Evaluating Threats, Capitalizing on Opportunities. Risk-Based Transportation Asset Management. US Department of Transportation, Washington, DC.

Project Management Institute (PMI) (2013). *A Guide to the Project Management Body of* Knowledge. PMBOK Guide, 5th ed., 2013, American National Standards Institute (ANSI) /Project Management Institute (PMI) 99-001-2008.

PWC (2008). A Practical Guide to Risk Assessment. connectedthinking, PricewaterhouseCoopers. (http://www.pwc.com/en_us/us/issues/enterprise-risk-management/assets/risk_assessment_guide.pdf).

Richrath, Scott (2011). Maintenance Level of Service (MLOS) at Colorado DOT. In, TAM Guide Webinar 3: The Transportation Asset Management Plan (TAMP). Slides 14-20. (http://tam.transportation.org/Documents/Webinars/TAM%20Guide%20Webinar%203%20draft%202.ppt as of Feb. 29, 2012).

Road Liaison Group (2011). Well Maintained Highways: Code of Practice for Highway Maintenance Management. UK Department of Transport, London, UK.

Saaty, T. L. (2009). Theory and Applications of the Analytic Network Process : decision making with benefits, opportunities, costs, and risks. RWS Publications, Pittsburgh, PA.

Saaty, T. L. and K. Peniwati (2008). Group Decision Making: Drawing out Reconciling Differences. RWS Publications, Pittsburgh, PA.

Shetland Islands Council (2010). Shetland Islands Council: Road Asset Management Plan. (http://www.shetland.gov.uk/roads/documents/DraftRAMP.pdf as of Jan. 27, 2012).

State of Missouri (2012). January 2012 Tracker Report. Accessed March 4, 2012 at http://www.modot.org/about/general_info/Tracker.htm.

State of Wisconsin. (2010). Final Compass Report. Wisconsin Department of Transportation, Madison, WI.

Tarnoff, Philip J., Stanley E. Young, Joshua Crunkleton, and Nezamuddin Nezamuddin (2008). Guide to Benchmarking Operations Performance Measures. NCHRP Project 20-07 /Task 202. Transportation Research Board, Washington, DC.

Thompson, Paul D (2011). Risk Management, Life Cycle Management, and Condition Forecasting. In, TAM Guide Webinar 4: Tools and Techniques for Implementing the TAMP. Slides 19-23. (http://tam.transportation.org/Documents/Webinars/TAM%20Guide%20Webinar%204%20draft%203.ppt as of Feb. 29, 2012).

Transport Scotland (2007). Road Asset Management Plan for Scottish Trunk Roads: April 2007-March 2008. The Scottish Government, Edinburgh. (http://www.transportscotland.gov.uk/files/Road_Asset_Management_Plan-April07-March09.pdf as of Jan. 28, 2012).

Transportation Association of Canada (TAC) (2006). Performance Measures for Road Networks: A Survey of Canadian Use. (www.tac-atc.ca/english/resourcecentre/.../perf-measures-0306.pdf as of Jan. 27, 2012).

USAF (1998). Operational Risk Management (ORM) Guidelines and Tools. Air Force Pamphlet 91-215. U.S. Air Force, Washington, DC.

Varma, S. (2012). Examining Risk-based Approaches to Transportation Asset Management (F. H. A. Office of Asset Management, Trans.) Risk-Based Asset Management: US Department of Transportation.

Vector Study. (2008). The Deming Cycle. (http://www.vectorstudy.com/management_theories/deming_cycle.htm , Accessed April 6, 2012).

Wireman Terry (1999). Performance Indicators for Maintenance. Engineer's Digest. (http://vestapartners.com/worxcms_files/PerformanceIndicatorsForMaintenance.pdf as of Jan. 27, 2012).

Zimmerman, Katharyn A., and Marshal Stivers (2007). A Guide to Maintenance Conditions Assessment Systems: Final Report. NCHRP Project No. 20-07, Task 206. Transportation Research Board, Washington, DC.

NCHRP 14-25:

Guide for Selecting Level-of-Service Targets for Maintaining and Operating Highway Assets

Webinar Slides

University of Wisconsin – Madison

2013

Setting Targets for Maintenance Performance Measurement

Teresa M. Adams
University of Wisconsin-Madison

National Center for Freight & Infrastructure Research & Education

Gaps in Current LOS Target Setting

- Using condition and cost to set attainable maintenance targets
- Precision of estimates from stratified sample
- Relating maintenance activities to maintenance goals
- Strategy to prioritize maintenance activities
- Cost of maintenance outputs
- Estimating cost to achieve desirable targets
- Budget allocation to maximize LOS outcome
- Considering risk in managing a maintenance program

Contents of the Guide

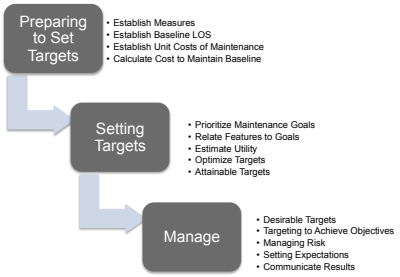
Definitions

- **Performance measure.** An indicator, preferably quantitative, of service provided by the transportation system to users; the service may be gauged in several ways (e.g., quality of ride, efficiency and safety of traffic movements, services at rest areas, quality of system condition, etc.).
- **Performance target**. The threshold value of a performance measure that an agency will strive to achieve to satisfy a policy objective.

Guide: Eight Step Process

1. Preparing to Set Targets
   a. Establish Goals, Measures and Costs
   b. Understand the Baseline
2. Setting Targets
   a. Priority and Utility Weights
   b. Linear Programming Optimization Model
3. Managing with Targets
   a. Communication
   b. Risk Management

Process for Setting Maintenance LOS Targets



Preparing to Set Targets

Highway Feature. In the terminology used in the Guide, a highway feature is the key to relating maintenance activities, to condition, costs, and goals.

- The agency measures the condition of each feature in terms of the percent of the inventory deficient on an LOS scale. The agency uses a set of maintenance activities designed to address the deficiencies in each feature. Maintenance of the highway features contributes to achieving goals. Maintenance costs are allocated to address deficiency in the features.

Prerequisite Data and Information for LOS Target Setting

Data Set	Description
Maintenance goals	Descriptions of the agencies maintenance goals and understanding of the relative importance of those goals. Ideally the maintenance goals relate to the agency's goals.
Inventory and condition assessment	Agencies should have in place the office and field procedures for collecting and managing their maintenance condition data. Target setting requires knowing the number of units of the highway features in total on the system and number of units not functioning as intended.
LOS definitions	LOS ranges for the highway features. The number of units of each feature not performing to the standard indicates the appropriate LOS.
Maintenance costs	Cost to perform maintenance activities expressed in units of work consistent with maintenance condition assessment: e.g., feet of edgelines, cost per face or square foot of sign replacement, acres of mowing, etc.
Utility of maintenance activities	An understanding of the relative contributions of maintenance activities for achieving the desired maintenance goals. For example, the weighted contribution of pavement markings for safety compared to other features that contribute to safety.

Establish Maintenance Performance Measures and LOS Scales

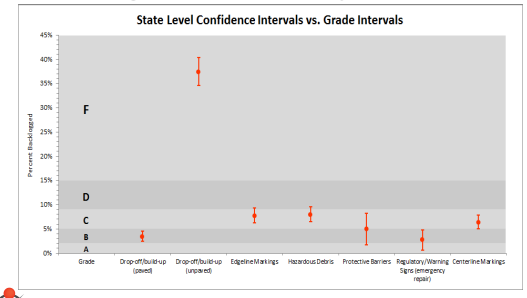
- LOS based on Pass-Fail Assessment
- Direct LOS Rating of Each Feature

LOS Rating Scale. LOS scales based on pass-fail assessment are the most commonly used frameworks for expressing LOS grades. The target-setting procedure in this Guide is based on the pass-fail with LOS assessment. The procedure requires that agencies determine their baseline LOS and the maintenance cost required to sustain the baseline. The Guide includes detailed procedures for analyzing sample data to estimate the percent of inventory that is deficient.

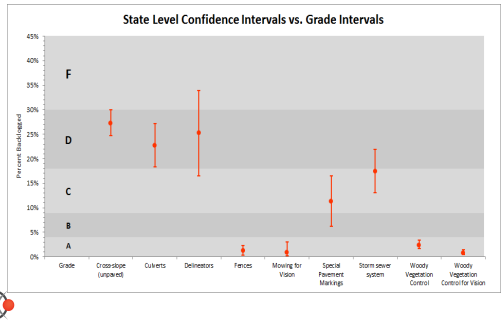
Example LOS Grading Scales for Highway Features (Wisconsin DOT)

	Percent of Feature Inventory in Deficient Condition				
Maintenance Feature	**A**	**B**	**C**	**D**	**F**
Rutting	≤2.5	>2.5 and ≤5.5	>5.5 and ≤9.5	>9.5 and ≤15	>15
Edgeline Markings	≤4.5	>4.5 and ≤9.5	>9.5 and ≤18.5	>18.5 and ≤30	>30
Shoulder Erosion	≤6.5	>6.5 and ≤15.5	>15.5 and ≤29.5	>29.5 and ≤50	>50
Longitudinal Joint Distress	≤7.5	>7.5 and ≤18.5	>18.5 and ≤35.5	>35.5 and ≤60	>60
Litter	≤10.5	>10.5 and ≤25.5	>25.5 and ≤47.5	>47.5 and ≤80	>80

Establish Baseline Performance e.g., Critical Safety Goal



Establish Baseline Performance e.g. Safety/Mobility Goal



Sampling Techniques

Type	Estimator	Characteristics	Example Features
A	Simple binomial proportions	• The feature inventory size is *known*. • The frequency of features in the sample is *known* prior to sampling. • The feature occurs on all sampled segments. • Each sampled segment gets a pass/fail rating for the maintenance condition of the feature.	• Litter • Hazardous debris
B	Domain binomial proportions	• The feature inventory size is *unknown*. • The frequency of features in the sample is *unknown* prior to sampling. • Only segments that have the feature get rated. • Each of those segments gets a pass/fail rating for the maintenance condition of the feature. • Some features are rare and therefore sparsely represented in the sample.	• Paved shoulders • Unpaved shoulders • Centerline markings • Mowing for Vision
C	Ratios	• The feature inventory size is unknown. • The frequency or amount of features in the sample is unknown prior to sampling. • All sampled segments are rated. The total quantity of the feature and the quantity in deficient condition are measured on each segment. • If the feature is not present on the segment, then the total quantity and quantity in deficient condition are zero. Thus, we can generalize to say all sample segments are rated. • Some features are rare and therefore sparsely represented in the sample.	• Ditches (linear feet) • Fences (linear feet) • Special pavement markings (number of markers) • Protective barriers (linear feet)

Sample Data

	Confidence Interval		Estimated Deficiency Rate
	Upper Bound	Lower Bound	
Drop-off/build-up (paved)	4.60%	2.40%	3.33%
Drop-off/build-up (unpaved)	40.31%	34.57%	37.39%
Edgeline Markings	9.31%	6.27%	7.66%
Hazardous Debris	9.53%	6.47%	7.87%
Protective Barriers	8.26%	1.68%	4.97%
Regulatory/Warning Signs (emergency repair)	4.83%	0.64%	2.74%
Centerline Markings	7.81%	5.05%	6.29%

Estimating Unit Cost

- Cost accounting
- Expert judgment
  - Based on placed materials
  - Based on activity
  - Based on available cost data
- Price Tags and Cost Allocation

Cost to Maintain Baseline

Cost Model for Maintaining the Baseline LOS

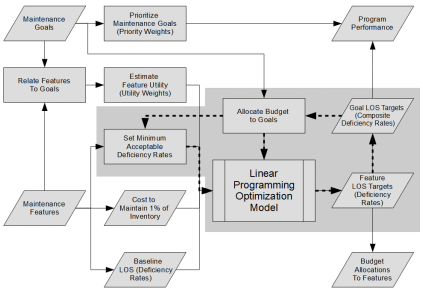
$(100/t_i)c_i$

- where t_i is the cycle time for feature i and c_i is the cost to treat one percent of the inventory of feature i.

Inventory, Cycle Time & Cost (This table or previous slide)

Feature	Inventory	Units	Unit Cost	Cost (1% inventory) c_i	Deficiency rate (%) x_i	Cycle Time t_i	% inventory for status quo $\frac{100}{t_i}$	Annual Cost for Status Quo $\frac{100}{t_i}c_i$
Reg/Warning Signs (emergency)	159,004	ea	$171.48	$272,660	3	20	5	$1,363,300
Hazardous Debris	11,774	CL	$1,120	$131,928	3	7	14	$1,884,681
Protective Barriers (Beam Guard)	3,704,457	LF	$34.36	$1,272,851	5	15	7	$8,485,676
Centerline Markings	56,799,150	LF	$0.15	$85,199	6	1	100	$8,519,873
Edgeline Markings	156,417,624	LF	$0.15	$234,626	7	1	100	$23,462,644
Drop off/Build up (unpaved shoulder)	21,619	mi	$330	$71,343	37	4	25	$1,783,568
Drop off/Build up (paved shoulder)	21,591	mi	$7,250	$1,565,348	3	15	7	$10,435,650
Woody Vegetation Vision	39,117	ea	$258.35	$101,059	1	3	33	$3,368,626
Mowing Vision	39,117	ea	$83.41	$32,627	1	1	100	$3,262,749
Special Pavement Markings	48,910	ea	$169.60	$82,951	10	1	100	$8,295,136
Woody Vegetation	29,625	LM	$1,033	$306,145	2	6	17	$5,102,413

Setting Targets



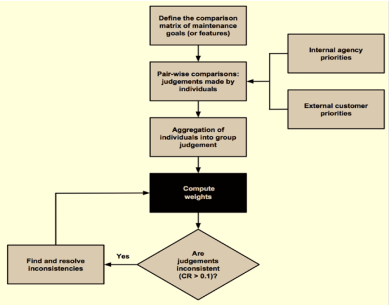
Prioritize Maintenance Goals

- Maintenance goals should align with the agency's strategic goals. Priority of maintenance goals should reflect what's important to internal and external stakeholders.

Example use of the SMART technique to establish weights of importance

Maintenance Goal	Rank	Importance	Weight
Critical Safety	1	100	0.38
Safety / Mobility	2	80	0.31
Stewardship	3	50	0.19
Ride / Comfort	4	20	0.08
Aesthetics	5	10	0.04
	Total	260	1.00

Analytical Hierarchy Process (AHP) for Determining Weights for Maintenance Goals and Features



Relating Maintenance Features to Goals

Element Group	Feature	Maintenance of this feature contributes primarily to goal of:				
		Critical Safety	Safety / Mobility	Stewardship	Ride / Comfort	Aesthetics
Traffic control & safety devices	Centerline markings	✓				
	Edge line markings	✓				
	Delineators		✓			
	Emergency repair of detour markers/ recreation / guide signs				✓	
	Routine replacement of detour markers/ recreation / guide signs				✓	
	Protective Barriers	✓				
	Emergency repair of regulatory / warning signs	✓				
	Routine replacement of regulatory / warning signs		✓			
	Special pavement markings		✓			
Shoulders	Hazardous debris	✓				
	Cracking on paved shoulder			✓		

Estimating Utility of Maintenance
for accomplishing goals

Roadway Features for Critical Safety	Emergency repair of regulator / warning signs	Hazardous debris	Protective barriers	Centerline markings	Edge line markings	Unpaved shoulder drop-off / buildup	Paved shoulder drop off / build up	Priority Weight
Emergency repair of regulator / warning signs	1	3	5	7	7	8	9	0.42
Hazardous debris	1/3	1	3	5	5	7	9	0.24
Protective barriers	1/5	1/3	1	3	3	5	7	0.13
Centerline markings	1/7	1/5	1/3	1	3	5	7	0.09
Edge line markings	1/7	1/5	1/3	1/3	1	5	7	0.07
Unpaved shoulder drop-off / buildup	1/8	1/7	1/5	1/5	1/5	1	3	0.03
Paved shoulder drop off / build up	1/9	1/9	1/7	1/7	1/7	1/3	1	0.02

Maintenance Performance

Critical Safety 0.52

Feature	Priority Weight: pwx	Priority Weight: glob-al
Emergency repair of regulator / warning signs	0.42	0.21 8
Hazardous debris	0.24	0.12 2
Unpaved shoulder drop-off / buildup	0.03	0.01 6
Centerline markings	0.09	0.04 7
Edge line markings	0.07	0.03 6
Paved shoulder drop-off / build up	0.02	0.01 0
Protective barriers	0.13	0.06 8

Safety / Mobility 0.28

Feature	Priority Weight: pwx	Priority Weight: glob-al
Routine replacement of regulator / warning signs	0.02	0.00 6
Woody vegetation (clear zone)	0.10	0.02 8
Special pavement markings	0.14	0.03 9
Mowing for vision	0.20	0.05 6
Culverts	0.08	0.02 2
Woody vegetation control for vision	0.13	0.00 7
Cross slope on unpaved shoulders	0.02	0.01 4
Storm sewer	0.01	0.01 4
Delineators	0.03	0.00 8
Fences	0.01	0.00 3

Stewardship 0.13

Feature	Priority Weight: glob-al
Ditches 0.46	0.06 0
Cracking on paved shoulders 0.08	0.01 0
Flumes 0.13	0.01 6
Curb and gutter 0.25	0.03 2
Drains 0.03	0.00 4
Erosion on unpaved shoulders 0.06	0.00 8

Ride / Comfort 0.05

Feature	Priority Weight: pwx	Priority Weight: glob-al
Potholes / raveling on paved shoulders	0.7 3	0.03 7
Emergency repair of regulator / warning signs	0.1 9	0.00 9
Routine replacement of regulator / warning signs	0.0 8	0.00 4

Aesthetics 0.03

Feature	Priority Weight: glob-al
Mowing 0.83	0.02 5
Litter 0.17	0.00 5

Priority Weights

Cost of Incremental Improvement in Feature and Goal Performance

Critical Safety Feature	Utility Weight	Marginal Cost (1 % deficiency)	
		Feature	Goal
Reg/Warning Signs (emergency repair)	0.42	$272,660	$649,190
Hazardous Debris	0.24	$131,928	$549,700
Protective Barriers (Beam Guard)	0.13	$1,272,851	$9,791,162
Centerline Markings	0.09	$85,199	$946,656
Edge line Markings	0.07	$234,626	$3,351,800
Drop off/Build up (unpaved shoulder)	0.03	$71,343	$2,378,100
Drop off/Build up (paved shoulder)	0.02	$1,565,348	$78,267,400

Linear Programming Model for Setting LOS Targets that Maximize Performance

The objective function to maximize performance is:

$$\max \sum_{i=1}^{n_1} (100 - x_i)\,\omega_i$$

Which leads to the objective function to maximize the increase in performance:

$$\max \sum_{i=1}^{n_1} \Delta x_i\,\omega_i$$

For analysis of a *goal category*: ω_i is the *goal-level* utility weight for each feature i in the goal category and n is the number of features in the *goal category*. For analysis of the *maintenance program*: ω_i is the *global* priority weight all features i in the program and n is the total number of goals in the *program*.

This function is subject to the following budget constraint on the cost function:

$$\sum_{i=1}^{n_1} \left(\frac{100}{t_i} + \Delta x_i\right) c_i \le \text{Budget}$$

Where, t_i is the maintenance cycle time in years for feature i; c_i is the cost to address deficiency in one percent of the inventory for feature i, and Δx_i is the decrease in deficiency rate to achieve the target for feature i.

A non-negativity restriction also exists at the feature level to limit the amount of growth in deficiency to the natural rate:

$$\left(\frac{100}{t_i} + \Delta x_i\right) c_i \ge 0$$

The target deficiency rate x_i is related to the current deficiency rate x_i' and decrease in rate:

$$x_i = x_i' - \Delta x_i$$

The constraints on the change in deficiency are,

$$\Delta x_i \le x_i'$$

$$x_i \le \max(x_i)$$

The constraint $\Delta x_i \le x_i'$ means the deficiency cannot be reduced by more than its current value. The constraint $x_i \le \max(x_i)$ requires the target rate to be no more than the maximum acceptable rate.

An optional constraint may be added to require a deficiency rate to stay the same or be reduced, but not increase. The following constraint will not allow the deficiency rate to increase by requiring the target to be less than the current rate.

$$x_i \le x_i'$$

Scope of Goal and Inventory for Setting LOS Targets

Scope of Analysis	Inventory Scope	Hierarchy Level	Weights	How to estimate the deficiency rates
Allocate feature-level spending to maximize performance	Single region (single stratum)	Features in a goal	Feature utility weights	Based on stratum subset of the sample
	State-wide (all strata)	Features in a goal	Feature utility weights	Based on full sample
Roll-up goal-level performance to evaluate program performance	Single region (stratum)	Goals	Goal priority weights	Based on stratum subset of the sample
	State-wide (all strata)	Goals	Goal priority weights	Based on full sample

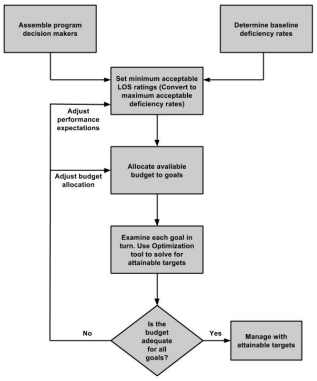
Budget Constraints and Attainable LOS Targets

1. Determine the baseline LOS
2. Assemble the program decision makers
3. Agree upon minimum acceptable LOS ratings for highway features
4. Determine the available budget and allocate the budget to achieving maintenance goals.
5. Examining each goal in turn, determine the highest affordable LOS targets that will satisfy minimum expectations and maximize the performance for the goal

Budget Constraints and Attainable LOS Targets (cont'd)

6. Determine if the budget is adequate to meet all goals
7. Adjust the budget allocation and/or adjust performance expectations
   a. Agree upon a new allocation of the available budget among the goals
   b. Use marginal costs to adjust expectations
8. Use the attainable targets to manage the maintenance program

Setting Targets

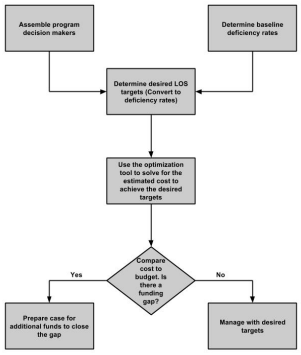


Exploring Cost and Desired LOS Targets

Desirable targets. The desirable targets, though non-attainable, are useful for gap analysis. The difference between what is being accomplished and what should be accomplished is the performance gap.

Methods to Set Desirable Targets

- Following established professional standards of practice
- Consensus
- Benchmarking
- Customer-responsive
- Policy direction



Process for Estimating Cost to Achieve Desired LOS Targets

LOS Target Setting for CRITICAL SAFETY

Budget =		Estimated Cost =	$65,174,500
Baseline composite deficiency rate =	5.79	Baseline LOS Grade =	C
Target composite rate reduction =	3.79		
Target composite deficiency rate =	2.00	Target LOS Grade =	A

FEATURE	FEATURE INPUTS					MAINTENANCE OF STEADY STATE		
	ω_i Utility Weight	Inventory	Inventory Unit	Unit Cost	t_i Cycle Time	$\left(\frac{100}{T_i}\right)$ Percent New Deficient Inventory Per Year	c_i Cost to Mitigate Deficiency in 1% of Inventory	$\left(\frac{100}{T_i}\right)c_i$ Cost to Maintain Baseline
Emergency repair of regulator / warning signs	0.42	159,004	ea	$171.48	20	5	$272,660	$1,363,300
Hazardous debris	0.24	117,740	CL	$1,120.50	7	14	$1,319,277	$18,846,810
Protective barriers	0.13	3,704,457	LF	$34.36	15	7	$1,272,851	$8,485,676
Centerline markings	0.09	56,799,150	LF	$0.15	4	25	$85,199	$2,129,968
Edgeline markings	0.07	156,417,624	LF	$0.15	4	25	$234,626	$5,865,661
Unpaved shoulder drop-off / buildup	0.03	21,619	mi	$330.00	4	25	$71,343	$1,783,568
Paved shoulder drop off / build up	0.02	21,591	mi	$7,250.00	15	7	$1,565,348	$10,435,650
							Total	$48,910,633

FEATURE	TARGET CALCULATIONS					WEIGHTED DEFICIENCY RATES		
	$\max(x_i)$ Maximum Acceptable Deficiency Rate	x_i^* Baseline Deficiency Rate	x_i Target Deficiency Rate	Δx_i Target Rate Reduction	Target Budget Allocation	$\omega_i x_i^*$ Weighted Baseline Deficiency Rate	$\omega_i x_i$ Weighted Target Deficiency Rate	$\omega_i \Delta x_i$ Weighted Rate Reduction
Emergency repair of regulator / warning signs	2	3	2	1	$1,635,960	1.26	0.84	0.42
Hazardous debris	2	7	2	5	$25,443,194	1.68	0.48	1.20
Protective barriers	2	5	2	3	$12,304,210	0.65	0.26	0.39
Centerline markings	2	6	2	4	$2,470,763	0.54	0.18	0.36
Edgeline markings	2	7	2	5	$7,038,793	0.49	0.14	0.35
Unpaved shoulder drop-off / buildup	2	37	2	35	$4,280,562	1.11	0.06	1.05
Paved shoulder drop off / build up	2	3	2	1	$12,000,998	0.06	0.04	0.02
				Total	$65,174,500	5.79	2.00	3.79

Allocating Feature Spending

- Maximize goal-level LOS
- Marginal Costs
- Optimization Model to help

Critical Safety Feature	Utility Weight	Marginal Cost (1 % deficiency)	
		Feature	Goal
Reg/Warning Signs (emergency repair)	0.42	$272,660	$649,190
Hazardous Debris	0.24	$131,928	$549,700
Protective Barriers (Beam Guard)	0.13	$1,272,851	$9,791,162
Centerline Markings	0.09	$85,199	$946,656
Edge line Markings	0.07	$234,626	$3,351,800
Drop off/Build up (unpaved shoulder)	0.03	$71,343	$2,378,100
Drop off/Build up (paved shoulder)	0.02	$1,565,348	$78,267,400

Linear Programming

LOS Target Setting for CRITICAL SAFETY

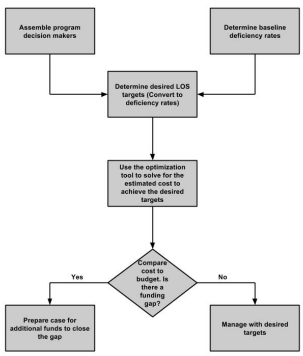
Budget =	$40,000,000	Estimated Cost =	$43,217,550
Baseline composite deficiency rate =	5.79	Baseline LOS Grade =	C
Target composite rate reduction =	-1.07		
Target composite deficiency rate =	6.86	Target LOS Grade =	C

FEATURE	FEATURE INPUTS					MAINTENANCE OF STEADY STATE		
	ω_i Utility Weight	Inventory	Inventory Unit	Unit Cost	t_i Cycle Time	$\left(\frac{100}{T_i}\right)$ Percent New Deficient Inventory Per Year	c_i Cost to Mitigate Deficiency in 1% of Inventory	$\left(\frac{100}{T_i}\right)c_i$ Cost to Maintain Baseline
Emergency repair of regulator / warning signs	0.42	159,004	ea	$171.48	20	5	$272,660	$1,363,300
Hazardous debris	0.24	117,740	CL	$1,120.50	7	14	$1,319,277	$18,846,810
Protective barriers	0.13	3,704,457	LF	$34.36	15	7	$1,272,851	$8,485,676
Centerline markings	0.09	56,799,150	LF	$0.15	4	25	$85,199	$2,129,968
Edgeline markings	0.07	156,417,624	LF	$0.15	4	25	$234,626	$5,865,661
Unpaved shoulder drop-off / buildup	0.03	21,619	mi	$330	4	25	$71,343	$1,783,568
Paved shoulder drop off / build up	0.02	21,591	mi	$7,250	15	7	$1,565,348	$10,435,650
							Total	$48,910,633

FEATURE	TARGET CALCULATIONS					WEIGHTED DEFICIENCY RATES		
	$\max(x_i)$ Maximum Acceptable Deficiency Rate	x_i^* Baseline Deficiency Rate	x_i Target Deficiency Rate	Δx_i Target Rate Reduction	Target Budget Allocation	$\omega_i x_i^*$ Weighted Baseline Deficiency Rate	$\omega_i x_i$ Weighted Target Deficiency Rate	$\omega_i \Delta x_i$ Weighted Rate Reduction
Emergency repair of regulator / warning signs	5	3	5	-2	$817,980	1.26	2.10	-0.84
Hazardous debris	9	7	9	-2	$16,208,257	1.68	2.16	-0.48
Protective barriers	5	5	5	0	$8,485,676	0.65	0.65	0.00
Centerline markings	6	6	6	0	$2,129,968	0.54	0.54	0.00
Edgeline markings	8	7	8	-1	$5,631,034	0.49	0.56	-0.07
Unpaved shoulder drop-off / buildup	25	37	25	12	$2,639,680	1.11	0.75	0.36
Paved shoulder drop off / build up	5	3	5	-2	$7,304,955	0.06	0.10	-0.04
				Total	$43,217,550	5.79	6.86	-1.07

Managing with Targets

1. Exploring Cost of Desired LOS Targets
2. Using LOS Targets to Achieve Management Objectives
3. Assess & mitigate risk
4. Set region and district expectations
5. Monitoring and Communicating



Desired Targets

LOS Target Setting for CRITICAL SAFETY

Budget =		Estimated Cost =	$65,174,500
Baseline composite deficiency rate =	5.79	Baseline LOS Grade =	C
Target composite rate reduction =	3.79		
Target composite deficiency rate =	2.00	Target LOS Grade =	A

FEATURE	FEATURE INPUTS					MAINTENANCE OF STEADY STATE		
	ω_i Utility Weight	Inventory	Inventory Unit	Unit Cost	t_i Cycle Time	$\left(\frac{100}{t_i}\right)$ Percent New Deficient Inventory Per Year	c_i Cost to Mitigate Deficiency in 1% of Inventory	$\left(\frac{100}{t_i}\right)c_i$ Cost to Maintain Baseline
Emergency repair of regulator / warning signs	0.42	159,004	ea	$171.48	20	5	$272,660	$1,363,300
Hazardous debris	0.24	117,749	CL	$1,120.50	7	14	$1,319,277	$18,846,810
Protective barriers	0.13	3,704,457	LF	$34.36	15	7	$1,272,851	$8,485,676
Centerline markings	0.09	56,799,150	LF	$0.15	4	25	$85,199	$2,129,968
Edgeline markings	0.07	156,417,624	LF	$0.15	4	25	$234,626	$5,865,661
Unpaved shoulder drop-off / buildup	0.03	21,619	mi	$330.00	4	25	$71,343	$1,783,568
Paved shoulder drop off / build up	0.02	21,591	mi	$7,250.00	15	7	$1,565,348	$10,435,650
							Total	$48,910,633

FEATURE	TARGET CALCULATIONS					WEIGHTED DEFICIENCY RATES		
	$\max(x_i)$ Maximum Acceptable Deficiency Rate	x_i' Baseline Deficiency Rate	x_i Target Deficiency Rate	Δx_i Target Rate Reduction	Target Budget Allocation	$\omega_i x_i'$ Weighted Baseline Deficiency Rate	$\omega_i x_i$ Weighted Target Deficiency Rate	$\omega_i \Delta x_i$ Weighted Rate Reduction
Emergency repair of regulator / warning signs	2	3	2	1	$1,635,960	1.26	0.84	0.42
Hazardous debris	2	7	2	5	$25,443,194	1.68	0.48	1.20
Protective barriers	2	5	2	3	$12,304,230	0.65	0.26	0.39
Centerline markings	2	6	2	4	$2,470,763	0.54	0.18	0.36
Edgeline markings	2	7	2	5	$7,038,793	0.49	0.14	0.35
Unpaved shoulder drop-off / buildup	2	37	2	35	$4,280,562	1.11	0.06	1.05
Paved shoulder drop off / build up	2	3	2	1	$12,000,999	0.06	0.04	0.02
				Total	$65,174,500	5.79	2.00	3.79

Using LOS Targets to Achieve Management Objectives

- Stretch
- Empowerment
- Cross-agency consistency
- Accountability, transparency, and gap analysis
- Continuous improvement

Management Objectives for Setting LOS Targets

- Stretch
- Empowerment
- Cross-agency consistency
- Accountability and transparency
- Continuous improvement

Managing Risk

- Risk is the positive or negative effects of uncertainty or variability upon the maintenance program objectives.
- Risk management comprises the cultures, processes and structures that are directed toward the effective management of potential opportunities and threats.

Five Step Process

1. Establish context: Define the focus.
2. Risk identification: Record risks
3. Risk analysis: Probability and consequence
4. Risk evaluation: How much risk is acceptable?
5. Risk treatment: Treat, tolerate, terminate, transfer, take advantage

Manage Risk in Setting and Achieving LOS Targets

Risk Register can be useful for:

- Managing risks that may impact the agency's ability to achieve its maintenance goals;
- Listing of specific areas of concern and their ranking in terms of likelihood and seriousness;
- Providing a documented framework for monitoring and reporting the status of risks;
- Documenting predefined risk mitigation, control and response actions to be pursued;
- Ensuring that risk management issues are being appropriately communicated to key stakeholders; and,
- Guiding efforts to seek involvement of the key stakeholders.

Risk Register

Risk Type (Program or Activity)	Risk Event	Context	Impact Severity (S)	Event Likelihood (L)	Rating (S)x(L)	Lead Responsibility	Risk Control and Response Actions
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bridge inspection failure	Bridge failure	Operational	Catastrophic	Rare	5	Chief Engineer	Recruit / train Bridge Inspectors
Truck-related activities	Insufficient CDL-qualified personnel	Compliance	Critical	Remote	8	Maintenance Division Engineer	Recruit / train CDL-qualified personnel
Snow and ice removal	Failure to provide statutory service	Operational	Critical	Occasional	12	Garage Foreman	Provide adequate manpower, equipment, materials
Inadequate traffic plan	Construction zone crashes	Operational	Critical	Occasional	12	Traffic Engineer	Increased training
Program	Insufficient funding	Information & Data	Critical	Occasional	12	Chief Engineer	Request fund transfer or supplemental appropriation
Shoulder drop-off / build up	Highway collision or fatality	Operational	Serious	Occasional	9	District Engineer	Prepare trend data on condition and expenditures; communicate risks; prepare to respond
Edge-line striping	Run off road crash	Operational	Serious	Occasional	9	Traffic Division Engineer	Provide adequate edge lines
Program	Worker injury or fatality	Operational / Compliance	Serious	Occasional	9	Maintenance Division Engineer	Safety training; enforce safety procedures
Program	LOS targets not met	Compliance / Information & Data	Serious	Occasional	9	Maintenance Division Engineer	Re-calibrate LOS ranges,
Program	Staff shortage	Operational	Serious	Occasional	9	Chief Engineer	Conduct training, succession planning, recruit / train employees
Missing delineators	Run off road crash	Operational	Marginal	Probable	8	Garage Foreman	Inspect, replace delineators

Risk Factors

Risk Category	Risk Events related to Risk Factors
Technical	Technology choice In-the-field implementation choice Equipment Material choice and quality
Acts of God	Normal natural calamities Abnormal natural calamities
Economic	Estimation errors Change in material costs
Policy	Change in statutes Inability to comply with statutes
Organizational	Material vendor failure Contractor failure Labor capacity and lack of adequate owner supervision (inspection)
Information and data	Inadequate sampling Inconsistent condition measurement Inconsistent / inaccurate Inspector Daily Reports (IDRs)

Assessing Tolerance to Risk

Level	Quantitative			Qualitative	
	Event Likelihood Category	Description	Nominal Annual Probability	Event Likelihood Category	Description
1	Rare	Return period is greater than 50 years (average of 50 years or more between events)	< 2%	Rare	I would be very surprised to see this happen but cannot entirely rule out the possibility
2	Remote	Return period is greater than 20-50 years	2% - 5%	Unlikely	I would be mildly surprised if this occurred, but cannot entirely rule out the possibility
3	Occasional	Return period is greater than 20 years	5% to 20%	Possible	I think this could maybe occur at some point, but not necessarily in the immediate future
4	Probable	Return period is approximately 1 to 5 years	20% to 100%	Likely	I think this could occur sometime in the coming year or so
5	Frequent	Return period is less than 1 year (average of 1 or more events per year)	100%	Almost Certain	I would not be at all surprised if this happened with the next few months

Severity of the Potential Impacts of Risky Events

		Potential Consequences on:				
Severity of Impact		Public		Preservation	Corridor / Region / Department	
Level	Category	Safety	Mobility	Asset / Environment	Financial Cost	Reputation Impact
1	Negligible	No safety hazard	Minimal delay	Minimal or cosmetic damage	<$100K	None
2	Marginal	Minimal safety hazard	Minor delay	Minor damage can be repaired on routine schedule	$100-$500K	None
3	Serious	Potential minor injuries	Major delay	Moderate damage requiring emergency repair	$500 to $1M	Minor
4	Critical	Potential major injuries	Detour, moderate duration	Extensive damage requiring significant emergency repair	$1M to $10M	Moderate
5	Catastrophic	Potential fatalities and major injuries	Detour, significant duration	Destroyed or large scale damage requiring closure for repair	>$10M	Severe

Risk Heat Map

Impact	Probability				
	Rare	Unlikely	Possible	Likely	Almost certain
Negligible	1	2	3	4	5
Marginal	2	4	6	8	10
Serious	3	6	9	12	15
Critical	4	8	12	16	20
Catastrophic	5	10	15	20	25

Take Action to Mitigate Risk

The Five Ts:

1. Can the risk be treated?
2. Can the risk be tolerated?
3. Termination may be another option
4. Transferring the risk is another option
5. Taking advantage of a risk simply recognizes that risk also present opportunities

Using LOS Targets to Set Expectations for Regions and Districts

An implementation plan would include:

1. The set of baseline deficiency rates and attainable target
2. Estimated budget allocations for achieving the targets
3. Estimated input goals on how resources are to be spent
4. The expected units of output necessary to reach the targets
5. Finally, the plan may include expectations for measurable outcomes

Monitoring and Communicating Progress

- Were the targeted conditions achieved?
- Were the planned units of output accomplished?
- Were the planned inputs consumed as planned?

Communicating

(using composite deficiency rates to prepare a goal-level performance report)

	2010		2011	
Goal Category	Composite Deficiency Rate (Adjusted % of inventory)	LOS Grade	Composite Deficiency Rate (Adjusted % of inventory)	LOS Grade
Critical Safety	4.81	B	5.79	C
Safety/Mobility	7.79	B	7.30	B
Stewardship	12.43	B	13.14	B
Ride/Comfort	7.36	A	8.26	B
Aesthetics	40.42	C	42.25	C

Goal	Priority Weight	Composite Deficiency Rate	Goal LOS Grade	Interpolation Formula	Grade Point X	Weighted Grade Point
Critical Safety	0.52	4.81	B		2.23	1.16
Safety /Mobility	0.28	7.79	B		2.34	0.66
Stewardship	0.13	12.43	B		2.34	0.30
Ride /Comfort	0.05	7.36	A		3.02	0.15
Aesthetics	0.03	40.42	C		1.32	0.04
Program-level LOS Grade					2.31 (B)	

Communicating (Using priority weights to measure and report program-wide performance)

NATIONAL CENTER FOR FREIGHT & INFRASTRUCTURE RESEARCH & EDUCATION

CFIRE

University of Wisconsin-Madison
Department of Civil and Environmental Engineering
1410 Engineering Drive, Room 270
Madison, WI 53706
Phone: 608-263-3175
Fax: 608-263-2512
cfire.wistrans.org