USDOT Region V Regional University Transportation Center Final Report

<u>NEXTRANS Project No 004OY01</u>

Optimal condition sampling of infrastructure networks

By

Rabi G. Mishalani, Principal Investigator
Associate Professor of Civil and Environmental Engineering and Geodetic Sciences
Ohio State University
mishalani@osu.edu

and

Prem Goel, Co-Principal Investigator
Professor of Statistics
Ohio State University
goel.1@osu.edu

Report Submission Date: October 15, 2009

DISCLAIMER

Partial funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

USDOT Region V Regional University Transportation Center Final Report

TECHNICAL SUMMARY

NEXTRANS Project No 004OY01 Final Report, October 2009

Optimal condition sampling of infrastructure networks

Introduction

Transportation infrastructure systems consist of spatially extensive and long-lived sets of interconnected facilities. Over the past two decades, several new non-destructive inspection technologies have been developed and applied in collecting raw condition data and processing them to produce useful condition input to infrastructure inspection, maintenance, and rehabilitation (IM&R) decision-making aimed at minimizing total expected life-cycle cost.

In response to the developments in inspection technologies, decision-making methods evolved whereby the optimum combination of inspection decisions on the one hand and maintenance and rehabilitation decisions on the other are determined based on an economic evaluation that captures the long-term costs and benefits. Recently, sample size has been included in IM&R decision-making as a decision variable when considering a single facility. While, the question of dealing with a network of facilities in making maintenance and rehabilitation decisions has been addressed in the literature, this treatment does not consider condition sampling whereby each facility could require a different set of sample sizes over time. Doing so is valuable given the network nature of facilities that most infrastructure agencies are responsible for, the increasing number of inspection technology choices with possible varying degrees of accuracy and cost, and budget constraints agencies have to work within.

During this reporting period (year 1 of a 2-year project) a methodology was developed to address the extension of the single facility level problem to the network level whereby the uncertainty due to condition sampling is captured and its related decision variables included in the IM&R decision-making process. More specifically, four activities were completed and a fourth was begun. The completed activities

- outlined the issues that could be considered in addressing condition sampling at the networks level,
- formulated the problem taking into account the issues deemed most critical above and devised a solution methodology, and

- developed and wrote computer code to solve the formulation following the devised solution methodology above.

The fourth activity consists of validating the code.

Findings

The preliminary code validation results are indicating that the code is arriving at the correct solutions based on a limited set of scenarios. More specifically, two approaches are being followed to validate the code. The first approach is based on comparing results arrived at by running the developed code that solves the network problem to results that are known to be correct. Identical results were arrived at for three different cost scenarios. The second approach is based on running the developed code that solves the network problem under different scenarios and examining whether the results change from scenario to scenario in a theoretically expected manner. When applying binding budget constraints, the results did change as theoretically expected whereby the total expected cost increased, IM&R cost decreased, and user cost increased (since terminal cost remained effectively unchanged).

Recommendations

While the validation results are encouraging, further testing for a wider set of scenarios is necessary before applying the methodology to more realistic cases and deriving insights regarding sampling at the network level based on that. In preparing for applications to more realistic cases and therefore solving larger problems, it is anticipated that computational issues relating to solving the LP for such larger problems will need to be addressed.

Contacts

For more information:

Rabi G. Mishalani
Principal Investigator
Civil and Environmental Engineering and Geodetic Sciences
Ohio State University
mishalani@osu.edu

Prem Goel
Co-Principal Investigator
Statistics
Ohio State University
goel.1@osu.edu

NEXTRANS Center
Purdue University - Discovery Park
2700 Kent B-100
West Lafayette, IN 47906

nextrans@purdue.edu
(765) 496-9729
(765) 807-3123 Fax

www.purdue.edu/dp/nextrans

NEXTRANS Project No 004OY01

Optimal condition sampling of infrastructure networks

By

Rabi G. Mishalani, Principal Investigator
Associate Professor of Civil and Environmental Engineering and Geodetic Sciences
The Ohio State University
mishalani@osu.edu

and

Prem Goel, Co-Principal Investigator
Professor of Statistics
The Ohio State University
goel.1@osu.edu

Report Submission Date: October 15, 2009

DISCLAIMER

Partial funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT07-G-005 of the U.S. Department of Transportation, Research and Innovative Technology Administration (RITA), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ACKNOWLEDGMENTS

The authors acknowledge the helpful discussions with Prof. Samer Madanat of University of California, Berkeley, Prof. Haris Koutsopoulos of Royal Institute of Technology, Sweden, and Dr. Liying Gong of Capital One relating to the approach and methodological aspects addressed in this report. Professors Madanat and Koutsopoulos are members of the project advisory committee. The authors also thank Graduate Research Assistants Mr. Cheng Chen and Mr. Dunke Zhou. In addition, they acknowledge the financial support offered by the USDOT/UTC program and The Ohio State University. The views and opinions contained in this report are those of the authors and do not represent the views, opinions, or policies of any other individuals, agency or group.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1. INTRODUCTION, PROBLEM, AND APPROACH

1.1. Introduction

Infrastructure systems consist of spatially extensive and long-lived sets of interconnected facilities, which are usually constructed through public, private, or joint endeavors for public or commercial use. The life-cycle of infrastructure systems is a complex and dynamic process, which consists of physical, managerial decision-making, and economic elements interacting with one another. Examples of infrastructure types include roadways, railways, runways, pipelines, and water distribution and wastewater collection systems. Even though the issues addressed in this study are valid for any facility type network, the application of interest relates to roadways.

Over the past two decades, several new non-destructive inspection technologies have been developed and applied in collecting raw condition data and processing them (McGhee, 2004) to produce useful condition input to infrastructure inspection, maintenance, and rehabilitation (IM&R) decision-making aimed at minimizing total expected life-cycle cost over a specified planning horizon. In response to the developments in inspection technologies, decision-making methods have also evolved in that the optimum combination of inspection decisions on the one hand and maintenance and rehabilitation decisions on the other are determined based on an economic evaluation that captures the long-term costs and benefits.

The most recent of such advances addressed the question of optimizing condition sampling for a specific facility. However, agency budgetary constraints and network effects can only be addressed in the decision-making process when all the facilities are considered simultaneously as an interconnected network. While, the question of dealing with a network of facilities in making maintenance and rehabilitation decisions has been addressed in the literature, this treatment does not consider condition sampling whereby each facility could require a different set of sample sizes over time. Doing so is valuable given the network nature of facilities that most infrastructure agencies are responsible for, the increasing number of inspection technology choices with possible varying degrees of accuracy and cost, and budget constraints agencies have to work within. During this reporting period (year 1 of a 2-year project) a methodology was developed to address the extension of the single facility level problem to the network level whereby the uncertainty due to condition sampling is captured and its related decision variables included in the IM&R decision-making process.

This report first presents the various possible considerations of a network treatment and specifies the aspects being considered. Following that, the formulation of the problem, its solution methodology, code development, and code validation methodology are presented. The preliminary finding regarding code validation follow, after which the report closes with conclusions.

1.2. Problem

Inspection deals with the gathering of data on the extent of facility damage expressed by variables such as pavement cracking and rutting (in the case of roadways). The data may be collected by visual inspection, through manual measurements, or by sensors. An average of collected damage measurements (by a single inspection technology) over a field is an estimate of the current condition of that field, and, in turn, is one primary input to forecasting future condition of a facility. A field is defined as a section of infrastructure where condition behaves homogeneously over space and time (Mishalani and Koutsopoulos, 1995, 2002) and where typically one inspection technology is employed.

The developments in non-destructive inspection technologies allow agencies to estimate facilities' conditions using large amount of data. The quality of measurements, the sample size, and the nature of correlation among condition variables at different locations determine the accuracy of condition estimates. Of course, more accurate estimates potentially lead to more effective maintenance and rehabilitation decisions an optimal cost. Consequently, the expected combined user costs and maintenance and rehabilitation costs are reduced over the planning horizon. However, more accurate information requires more resources such as increased inspection frequency, advanced inspection sensor technologies, larger sample sizes, as well as statistical methods that appropriately combine all this information. Therefore, the optimum combination of inspection decisions and maintenance and rehabilitation decisions should be determined based on an economic evaluation that captures the long-term costs and benefits.

In the inspection data collection and utilization process, two sources of uncertainty are present: measurement and spatial sampling. A distress variable is distributed along a specific homogenous section (referred to as field above) with a given constant mean value, constant variance, and spatial correlation function (Mishalani and Koutsopoulos, 1995, 2002). Therefore, the values of condition at n sample locations vary around a mean value reflecting inherent spatial correlations. The corresponding measured condition values are noisy versions of the true condition at the sample locations and reflect an additional degree of uncertainty due to measurement technology limitations and environmental effects (Humplick, 1992). The sample mean of measured condition is calculated from these observations. Therefore, the difference between this condition estimate and the true condition mean reflects both measurement and spatial sampling uncertainty, where the latter is due to the spatial variation in the condition variables and positive correlation among them. Together with forecasting uncertainty, which is introduced when predicting future condition (Carnahan et al., 1987; Madanat, 1993, Madanat and Ben-Akiva, 1994), the combination of these three sources of uncertainty could affect the resulting M&R decisions through their influence on the quality of condition estimation and forecasting.

The question of dealing with an entire network of facilities in making maintenance and rehabilitation decisions was addressed by Smilowitz and Madanat (2000) (henceforth referred to as the S&M study), however, this treatment does not consider condition sampling whereby each facility could require a different set of sample sizes over time. Incorporating Sampling uncertainty in IM&R decision-making and including the sample size as a decision variable for a single homogeneous facility (or field) was addressed by Gong (2006) and Mishalani and Gong (2008, 2009) (henceforth referred to as the G&M study). This study addresses the extension of the single facility level problem to the network level whereby the uncertainty due to condition sampling will be captured and its related decision variables included in the IM&R decision-making process.

1.3. Approach

In this reporting period, the research effort builds upon the S&M study on optimizing IM&R policies for a network of facilities without any condition sampling considerations and the G&M study on optimizing condition sampling for a single facility. Each study has addressed separately one of the two main aspects of this research projects, namely condition sampling and network treatment of decision-making. Thus, building upon these two approaches by incorporating both elements in a consistent manner within a single framework is an effective approach to follow.

The activities during year 1 of this 2-year project have been divided into the following categories:

Activity 1: Outline the issues that could be considered in addressing the condition sampling at the networks level.

Activity 2: Formulate the problem taking into account the issues deemed most critical under Activity 1 above and devise a solution methodology.

Activity 3: Develop and write computer code to solve the formulation following the devised solution methodology under Activity 2 above.

Activity 4: Validate the code developed under Activity 3 above.

CHAPTER 2. METHODOLOGY

This chapter is organized by the four activities outlined above.

2.1. Network sampling issues

When considering a network of facilities, various dimensions can be considered including: length heterogeneity of the facilities, condition similarity of facilities, costs independent and dependent on network structure, and budget constraints. Each dimension is discussed in what follows.

The S&M study assumed a single facility length. However, given the definition of homogeneous facilities referred to in section 1.2, the lengths of facilities within a network will vary from facility to facility. In the presence of spatial correlation among condition observations within a homogenous facility, which is known to exist as is also referred to in section 1.2, the effects of sampling on the uncertainty of condition assessment will depend on facility length. Therefore, in this study the developed formulation considers facilities of varying lengths.

Addressing decision-making under uncertainty is at the core of the S&M and M&G studies and this one as well. In doing so, however, the number of possible future condition related states to consider increases exponentially over time, which results in computational difficulties for realistically long planning horizons. As is done in the previous two studies, similar states are grouped together to reduce the computational cost. The state of the system at a given point in time is characterized by the so-called augmented state, which includes all the observed condition states and actions taken up to that point. To readily address this grouping in a straightforward manner, the sufficient statistic known as the information vector is used to represent the state instead. The information vector is the probability mass function of the true condition state given the augmented state, which can be calculated directly for each augmented state using Bayes' Law. As a result, each information vector corresponds to a given augmented state. An important advantage of resorting to information vectors instead of augmented states is that the similarity between various information vectors can be readily measured and similar vectors can be grouped. This grouping could be achieved by assuming that the elements of information vectors take some predetermined set of discrete values (e.g., 0.0, 0.1, 0.2, … 0.9, and 1.0). Each information vector is mapped into one of the finite possibilities and, therefore, similar vectors will be mapped into the same vector among the set of predetermined vectors, thus effectively grouped.

There are numerous costs at play in IM&R decision-making, some of which depend on the specific network structure and some do not. One of the components of user cost is the vehicle operating cost as a function of the condition of the facility. Another is the delay due to IM&R activities. Both are considered in this study. However, at this stage, delays resulting from network interactions in the form of spillovers upstream and onto other paths are not modeled. IM&R include the fixed and variable costs associated with applying the activities, some of which are a function of the condition state. Again in the case of these costs, at this stage, the elements that depend on the structure of the network including the interactions of labor, equipment, and material transportation costs are not modeled.

Finally, as already discussed, budget constraints can only be considered at the network level, a critical motivation for addressing the optimal IM&R decision-making problem at the network level in light of the common budget constraints public agencies face. In this study, the budget constraints are restricted to apply to maintenance and rehabilitation costs, one of the largest costs an agency incurs in this context. Nevertheless, other constraints do apply, but are not considered at this stage. Such constraints include labor, equipment, and material availability, which could depend on the structure of the network as well.

2.2. Problem Formulation and solution methodology

As already pointed out, based on the S&M study regarding the network level infrastructure management problem without sampling and the G&M study regarding facility level infrastructure management with sampling, this study focuses on network level IM&R decision making with sampling. The following items have been incorporated in the developed integrated formulation for addressing the sampling problem at the network level.

<u>Condition states</u>: As in the S&M and G&M studies, assessed facility condition is assumed to fall into one of a finite number of discrete condition states.

<u>Information vector</u>: Due to measurement error and inherent variability in condition within a facility, it is impossible to know the true condition state with certainty. However, as already discussed in section 2.1, given the current condition assessment through measurement and sampling and historical information including previous measurements and IM&R actions, the posterior probability mass function of the true condition state of each facility (the information vector, I) can be determined. As already discussed, the information vector is a sufficient statistic that corresponds to a given assessed conditino through the calculation discussed above.

<u>Facility age</u>: The age of a facility is defined as the number of years since it was last rehabilitated. It is important to keep track of age in addition to time because the transition probabilities (discussed below) depend on age.

<u>Transition probabilities</u>: Transition probabilities specify how facility condition evolves during the next time period, given its current condition state, age, and the maintenance and rehabilitation action applied. More specifically, a transition probability represents the likelihood of a facility to transition to a certain state in one period given the state it is currently in. Therefore, transition probabilities can be organized into a matrix representing all combinations of transitions from state to state. Two facilities in the same condition state to which the same routine maintenance action is applied, will have different transition probabilities if their age is different. The facility with lower age has a smaller probability of

deteriorating to a poorer state during the next period. Therefore, it is important to allow for non-stationary transition probabilities conditional on the age of a facility.

Sampling uncertainty: As developed in detail in the G&M study, the variance of the assessed condition is a critical element of the formulation and is determined as a function of the measurement technology, sample size, and the characteristics of the facility (in terms of the inherent variability in condition and the spatial correlation).

Costs: In addition to the user and IM&R costs discussed in section 2.1, the hypothetical terminal cost incurred at the end of the time horizon represents the cost of bringing the facility back to the best condition state for the purpose of equalizing the service life from that point onward. The terminal cost is denoted by $TC(l, I)$ where l denotes the length of the facility in question. The expected value of all other costs combined is denoted by $EC(l, I, a, r, n)$ where a denotes the maintenance and rehabilitation action and r the inspection technology.

Decision variables: $W_t(l, y, I, a, r, n)$ denotes the number of facilities whose information vector is I, whose length is l, of age y (i.e., they hasn't been rehabilitated for y years), on which maintenance and rehabilitation action a will be performed, and inspection technology r will be used with n samples. In addition, $W_T(l, I)$ is the number of facilities of length l in information vector I at the end of the planning horizon. As in the case of the S&M study, the nature of the decision variable when solving the network problem is that of a randomized policy. That is, facilities associated with the same information vector could receive different actions (optimally determined). Such a result is possible due to the presence of budget and condition constraint (discussed in more detail below). It is important to note that which of the facilities associated with the same information vector receives which policy when such solutions are arrived at is not part of the solution outcome.

Objective function: For each feasible realization of the decision variables, the expected discounted cost includes the user, inspection, maintenance and rehabilitation, and terminal costs. Therefore, the objective is given by

$$\min \sum_t \alpha^t \sum_l \sum_y \sum_I \sum_a \sum_r \sum_n EC(l, I, a, r, n) \times W_t(l, y, I, a, r, n) + \alpha^T \sum_I \sum_l TC(l, I) \times W_T(l, I)$$

Note that the expected discounted cost is a linear function of the decision variables.

Constraints:

- Non-negativity: These constraints guarantee that each decision variable is non-negative.

- Conservation: These constraints ensure the conservation of facilities over time. That is, the information vectors must transition from one period to the next in a manner consistent with the condition state transition probabilities.
- Initial state: The initial distribution of facilities as defined by a set of information vectors is assumed known.
- Condition states: These constraints require that the proportion or number of facilities in the condition states considered to be poor are bounded by a maximum value each year.
- Budget: These constraints require that the IM&R cost is bounded by a maximum and possibly a minimum value each year.

At this point, the problem is reduced to finding the values of the decision variables that satisfy all the constraints and achieve the minimal objective function value. Since the objective function and all the constraints are linear with respect to the decision variables, the problem can be solved using linear programming (LP). However, the size of the optimization problem grows fast with added dimensionality, and standard LP routines in MATLAB (the currently adopted programming software as is discussed in more detail below) may not be adequate for solving such large problems. Eventually, LP routines that can handle realistically large problems should be resorted to.

It is worthwhile to note here that while the network level formulation can be solved using LP due to the nature of the decision variables and formulation, at the facility level, the decision variables take the form of the actual inspection technology, sample size, and maintenance and rehabilitation action indexes and, therefore, LP does not apply. In the case of the G&M study (and some other facility level studies preceding it), dynamic programming was used to solve the problem.

2.3. Code development

MATLAB 2007a has thus far been used for programming and for solving the LP. The developed code includes five modules, each of which is described below.

<u>Module 1</u>: Prepares the inputs into a readable form. These inputs relate to the condition states, discretization degree of the elements of the information vector, discount factor, maintenance and rehabilitation activities, inspection technologies, spatial correlation, sample sizes, terminal costs, user costs, inspection costs, maintenance and rehabilitation costs, transition probabilities by age, condition constraints by year, and budget constraints by year.

<u>Module 2</u>: Calculates intermediate elements. These include the transition probabilities for the information vectors based on the condition state transition probabilities, measurement accuracy and sample size; and total expected cost for each year, age, information vector, inspection technology, sample size, and maintenance and rehabilitation action.

<u>Module 3</u>: Sets up the objective function. The LP coefficients are specified based on the total expected cost from Module 2 and terminal cost and discount factor from Module 1.

<u>Module 4</u>: Sets up the constraints. The non-negativity, conservation, condition, and budget constraints are specified.

<u>Module 5</u>: Solves the LP. The LP solver in MATLAB is applied to solve the LP setup based on Modules 3 and 4.

While each module was extensively debugged as it was developed, it is also critical to debug and validate the integration of the five modules. The latter is the subject of the next section.

2.4. Code validation

Two approaches are being followed to validate the code that integrates the five modules to arrive at a solution to the network problem. This is an ongoing activity and, therefore, what is presented below reflects the initial stages of this process.

The first approach is based on comparing results arrived at by running the developed code that solves the network problem to results that are known to be correct. The second approach is based on running the developed code that solves the network problem under different scenarios and examining whether the results change from scenario to scenario in a theoretically expected manner.

2.4.1. Single facility versus a network of identical facilities

A network problem is specified in such a manner that renders its solution to be identical to that of a single facility problem. Doing so is desirable because the code that solves the single facility problem developed in the G&M study was thoroughly tested and, therefore, the solutions it arrives at are known to be correct with a high degree of confidence. The manner in which the network problem is set up to ensure equality between the network and facility level solutions consists of the following. First, specify the network to be comprised of multiple facilities all of which are identical in all dimensions to the facility of the single facility problem. Second, do not impose any budget or condition constraints. If the code developed to solve the network problem in this study arrives at the same solution as that of the single facility problem (when it is theoretically expected to), then the confidence in this developed code is established. Of course, this equality must be achieved under a wide set of scenarios. At this stage only a limited set of three scenarios have been specified.

The variable specifications that are common to all three scenarios are the following:

- Three conditions states: 1, 2, and 3; 1 representing the best state and 3 the worst.

- Information vector discretization degree: 0.1. That is, each element in the information vector can take one of the following values: 0.0, 0.1, 0.2, …, 0.9, or 1.0.

- Initial condition state: 1 with probability 1.0.
- Spatial correlation of condition: 0.
- Two inspection technologies: low (1) and high (2) measurement error technologies.
- Four sapling levels: 4 represents a 100% sampling, 3 represents 60%, 2 represents 20%, and 1 represents 0% (and thus no inspection).
- Two maintenance and rehabilitation activities: routine (1), and rehabilitation (2).
- Inspection cost (from the G&M study):

  0.10025 + 0.00023368 × (sample size) for the low measurement error technology (1), and

  0.0134 + 0.000085 x (sample size) for the high measurement error technology (2).
- Rehabilitation cost (from the G&M study): 64.869 $\$/m^2$.
- User cost (from the G&M study): 36.937, 70.835, and 100.000 $\$/m^2$ for conditions states 1, 2, and 3, respectively.
- Terminal Cost (from the G&M study): 0.155, 5.195, and 14.058 $\$/m^2$ for conditions states 1, 2, and 3, respectively.
- Transition probabilities: The transition probabilities under routine maintenance and rehabilitation are set as shown in Tables 1 and 2, respectively.

Table 1: Transition probabilities under routine maintenance

From\To	1	2	3
Age = 1			
1	0.92	0.08	0.00
2	0.00	1.00	0.00
3	0.00	0.00	1.00
Age = 2			
1	1.00	0.88	0.12
2	2.00	0.00	1.00
3	3.00	0.00	0.00
Age = 3			
1	0.75	0.25	0.00
2	0.00	0.94	0.06
3	0.00	0.00	1.00

Table 2: Transition probabilities under rehabilitations

From\To	1	2	3
1	0.92	0.08	0.00
2	0.92	0.08	0.00
3	0.92	0.08	0.00

The scenarios are defined by varying the routine maintenance cost as shown in Table 3. For the Base Scenario, the routine maintenance costs are adopted from the G&M study. For Alternative Scenarios 1 and 2, these costs are increased.

Table 3: Routine maintenance cost ($\$/m^2$) scenarios

	Condition State		
	1	2	3
Base Scenario	1.7872	4.4472	11.4472
Alternative Scenario 1	10.340	13.000	20.000
Alternative Scenario 2	20.000	40.000	60.000

2.4.2. Network with and without budget constraints

In this exercise, the network problem is solved under different scenarios and the changes in the results are verified against theoretically expected characteristics. Only two scenarios are specified at this stage. Further variations are being explored.

The first scenario consists of Alternative Scenario 2 from the validation exercise presented in section 2.4.1. The second scenario consists of the same specification in addition to the application of IM&R budget constraints set at 24.000 $\$/m^2$ for each year. The application of budget constraints, if binding, is theoretically expected to result in an increase in the total minimum expected cost, a reduction in IM&R cost and the application of rehabilitation given its higher cost, and an increase in the combination of user cost and terminal cost.

CHAPTER 3. FINDINGS

The findings during year-1 of this 2-year project relate to the preliminary validation of the developed code. The presentation of these results is organized by the two validation exercises.

3.1. Single facility versus a network of identical facilities

Recall, three routine maintenance cost scenarios are considered in this validation exercise. The value of the objective function at optimality is found to be the same when solving the network and single facility problems as expected (for the specific definition of the network where it consists of a set of facilities identical to the facility of the single facility problem and where no budget and condition constraints are present). The optimal values of the objective function are shown in Table 4 for each of the three cost scenarios considered.

Table 4: Value of objective function ($\$/m^2$) at optimality
for different cost scenarios, network and single facility cases

	Base Scenario	Alternative Scenario 1	Alternative Scenario 2
Optimal Total Cost	139.1036	164.6598	195.63

In addition to examining the value of the objective function at optimality, it is important to examine the actual optimal decision policies and compare them across the two cases of a single facility and a network of identical facilities. As already discussed, in the absence of budget and condition constraints, the optimal policies for each of these two cases must be identical. Note that due to the absence of budget and condition constraint, each information vector (which, as already discussed, is associated with a specific assessed condition state from the sample of measurements) will have only one optimal set of IM&R actions corresponding to it. That is, the randomized policy in the network case will consist of that optimal set with probability one. That set must be the same as the optimal set of IM&R actions determined for the same information vector in the case of the single facility.

Table 5 shows the optimal IM&R policies for the Base Scenario. The results are identical for both the network and the single facility cases and, therefore, they are summarized in one table. The first column indicates the year in question while the second indicates the age of the facility (time since the last rehabilitation). As already discussed, it is important to keep track of age separately from the year in question because the transition probabilities are age-dependent. The information vector set of columns indicates the probability mass function by condition state (recall, state 1 represents the best condition and state 3 the worst). The likelihood is the probability associated with the corresponding information vector. In the network case, it represents the average percentage of facilities reflecting that particular information vector. The last three columns indicate the optimal policy regarding measurement technology, sample size, and maintenance and rehabilitation action for the corresponding information vector.

Table 5: Optimal IM&R policies as a function of information vectors
for the Base scenario, network and single facility cases

	Age	Information Vector			Likelihood	Sampling Tech.	Sample Size	M&R Action
		1	2	3				
Year 1	1	1.0	0.0	0.0	1	2	2	1
Year 2	2	1.0	0.0	0.0	0.801	2	2	1
	2	0.8	0.2	0.0	0.177	2	4	1
	2	0.2	0.8	0.0	0.022	–	1	1
Year 3	3	1.0	0.0	0.0	0.743	–	1	1
	3	0.7	0.3	0.0	0.155	–	1	1
	3	0.2	0.8	0.0	0.072	–	1	1
	3	0.1	0.9	0.0	0.025	–	1	1
	3	0.0	1.0	0.0	0.006	–	1	1

For example, in year 2, the likelihood of the facility in the single facility case or each of the facilities in the network case reflecting information vector [0.8 0.2 0.0] is 0.177, and the optimal actions for facilities whose age is 2 and are found to reflect this information vector are to apply routine maintenance (index = 1) this coming year and to inspect condition at the beginning of the following year by employing the higher measurement error inspection technology (index = 2) and a sample size of 100% (index = 4) of the length of the facility. Note that since in year 1 facilities are in state 1 with probability 1.0, the likelihoods in year 2 represent the transition probabilities to each of the respective information vectors given the condition state of 1 in year 1.

Again, the results regarding the information vectors, the likelihoods, and corresponding optimal policies are found to be identical when applying the dynamic programming code to solve the single facility problem and the newly developed linear programming code to solve the network problem. Identical results are arrived at for the two other cost scenarios. These results provide an initial indication that the code seems reliable under the limited set of scenarios it is tested under. Further examination under a wider set of scenarios is underway.

3.2. Network with and without budget constraints

The budget constraint applied to Alternative Scenario 2 is found to be binding as indicated by an increase in the value of the objective function at optimality with respect to the case of no budget constraint as indicated in Table 6. Table 6 also shows the breakdown of this total cost into IM&R, user, and terminal costs. Notice that as expected when a binding budget constraint is applied, IM&R cost is reduced (to meet the constraint) at the expense of user cost (terminal cost remains effectively the same in this case), resulting in a net loss reflected in the higher total cost at optimality.

Table 6: Total, user, IM&R and terminal costs ($\$/m^2$) at optimality
with and without the budget constraints for Alternative Scenario 2

	With Budget Constraints	No Budget Constraint
Optimal Total Cost	195.69	195.63
Optimal IM&R Cost	65.721	66.298
Optimal User Cost	128.35	127.80
Terminal Cost	1.6189	1.5328

Once again, it is important to also examine the optimal policies with and without the binding budget constraints. Table 7 presents the results. The structure of the table is similar to that of Table 5 except that two likelihood values are presented, one for the case with the binding budget constraints in effect and the other without any budget constraint. The only difference between the two cases – highlighted in bold in Table 7 – is that in year 3 when facilities are associated with information vector [0.2 0.8 0.0] and are of age 3 (a situation that materializes with a 5% likelihood), two different optimal policies are arrived at. In the case of no budget constraints, rehabilitation (index = 2) is the optimal maintenance and rehabilitation action while in the case of binding budget constraints, a randomized optimal policy is in effect whereby the optimal policy is to apply rehabilitation to 60% (0.03/0.05) of these facilities and routine maintenance (index = 1) to 40% (0.02/0.05). The optimal inspection action remains unaffected whereby no inspection is prescribed in both cases.

The shift to rehabilitating 60% of the facilities associated with the information vector case discussed above in the presence of binding budget constraints from the 100% in the absence of such constraints is consistent with the expectation that under binding budget constraints less costly maintenance and rehabilitation actions will be prescribed. Clearly, rehabilitation is more costly than routine maintenance thus the shift from rehabilitation to routine maintenance for 40% of the facilities associated with that information vector. While this exercise is limited and additional such validations are necessarily, the results indicate that at least under the tested scenarios the code is producing results that have characteristics consistent with theoretical expectations.

Table 7: Optimal IM&R policies with and without budget constraints for Alternative Scenario 2

	Age	Info. Vector			Likelihood w/ budget constraint	Likelihood w/o budget constraint	Sampling Tech.	Sample Size	M&R Action
		1	2	3					
Year 1	1	1.0	0.0	0.0	1	1	2	2	1
Year 2	2	1.0	0.0	0.0	0.801	0.801	2	2	1
	2	0.8	0.2	0.0	0.177	0.177	2	4	1
	2	0.2	0.8	0.0	0.022	0.022	2	2	2
Year 3	1	1.0	0.0	0.0	0.017	0.017	–	1	1
	1	0.8	0.2	0.0	0.004	0.004	–	1	1
	1	0.2	0.8	0.0	0.001	0.001	–	1	1
	3	1.0	0	0.0	0.743	0.743	–	1	1
	3	0.7	0.3	0.0	0.155	0.155	–	1	1
	3	0.2	0.8	0.0	**0.020**	**0.000**	–	1	**1**
	3	0.2	0.8	0.0	**0.030**	**0.050**	–	1	**2**
	3	0.1	0.9	0.0	0.025	0.025	–	1	2
	3	0.0	1.0	0.0	0.006	0.006	–	1	2

CHAPTER 4. CONCLUSIONS

Transportation infrastructure systems consist of spatially extensive and long-lived sets of interconnected facilities. Over the past two decades, several new non-destructive inspection technologies have been developed and applied in collecting raw condition data and processing them to produce useful condition input to infrastructure inspection, maintenance, and rehabilitation (IM&R) decision-making aimed at minimizing total expected life-cycle cost.

In response to the developments in inspection technologies, decision-making methods evolved whereby the optimum combination of inspection decisions on the one hand and maintenance and rehabilitation decisions on the other are determined based on an economic evaluation that captures the long-term costs and benefits. Recently, sample size has been included in IM&R decision-making as a decision variable when considering a single facility. While, the question of dealing with a network of facilities in making maintenance and rehabilitation decisions has been addressed in the literature, this treatment does not consider condition sampling whereby each facility could require a different set of sample sizes over time. Doing so is valuable given the network nature of facilities that most infrastructure agencies are responsible for, the increasing number of inspection technology choices with possible varying degrees of accuracy and cost, and budget constraints agencies have to work within.

During this reporting period (year 1 of a 2-year project) a methodology was developed to address the extension of the single facility level problem to the network level whereby the uncertainty due to condition sampling is captured and its related decision variables included in the IM&R decision-making process. More specifically, four activities were completed and a fourth was begun. The completed activities

- outlined the issues that could be considered in addressing condition sampling at the networks level,
- formulated the problem taking into account the issues deemed most critical above and devised a solution methodology, and
- developed and wrote computer code to solve the formulation following the devised solution methodology above.

The fourth activity consists of validating the code. The preliminary results are indicating that the code is arriving at the correct solutions based on a limited set of scenarios. More specifically, two approaches are being followed to validate the code. The first approach is based on comparing results arrived at by running the developed code that solves the network problem to results that are known to be correct. Identical results were arrived at for three different cost scenarios. The second approach is based on running the developed code that solves the network problem under different scenarios and examining whether the results change from scenario to scenario in a theoretically expected manner. When applying binding budget constraints, the results did change as theoretically expected whereby the total expected cost increased, IM&R cost decreased, and user cost increased (since terminal cost remained effectively unchanged).

While the validation results are encouraging, further testing for a wider set of scenarios is necessary before applying the methodology to more realistic cases and deriving insights regarding sampling at the network level based on that. In preparing for applications to more realistic cases and therefore solving larger problems, it is anticipated that computational issues relating to solving the LP for such larger problems will need to be addressed.

REFERENCES

Carnahan, J.V., Davis, E. J., Shahin, M.Y., Keane, P.L., and Wu, M. I. (1987). Optimal Maintenance Decisions for Pavement Management. *Journal of Transportation Engineering*, Vol. 113, No. 5, pp. 554-572.

Gong, L. (2006). *Optimal Spatial Sampling of Infrastructure Condition: A Life-cycle-based Approach under Uncertainty*, Ph.D. dissertation, The Ohio State University, Columbus, OH.

Humplick, F. (1992). Highway Pavement Distress Evaluation: Modeling Measurement Error. *Transportation Research Part B*, Vol. 26B, pp. 135-154.

Madanat, S.M. (1993). Optimal Infrastructure Management Decisions under Uncertainty. *Transportation Research Part C*, Vol. 1, No.1, pp. 77-88.

Madanat, S.M. and Ben-Akiva, M. (1994). Optimal Inspection and Repair Policies for Infrastructure Facilities. *Transportation Science*, Vol. 28, No. 1, pp. 55-62.

McGhee, K.H. (2004). Automated Pavement Distress Collection Techniques: A Synthesis of Highway Practice. *NCHRP SYNTHESIS 334*, Transportation Research Board, Washington, D.C.

Mishalani, R.G., and Gong, L. (2008). Evaluating the Impact of Pavement Condition Sampling Advances on Life-cycle Management. *Transportation Research Record*, No. 2068, TRB, National Research Council, Washington, D.C., pp. 1-9.

Mishalani, R.G., and Gong, L. (2009). Optimal Infrastructure Condition Sampling over Space and Time for Maintenance Decision-making under Uncertainty. *Transportation Research Part B*, Vol. 43, pp. 311-324.

Mishalani, R.G., and Koutsopoulos, H.N. (1995). Uniform Infrastructure Fields: Definition and Identification. *Journal of Infrastructure Systems*, ASCE, Vol. 1, No. 1, pp. 44-55.

Mishalani, R.G., and Koutsopoulos, H.N. (2002). Modeling the Spatial Behavior of Infrastructure Condition. *Transportation Research Part B*, Vol. 36, pp. 171-194.

Smilowitz, K., and Madanat, S. (2000). Optimal Inspection and Maintenance Policies for Infrastructure Networks. *Computer-Aided Civil and Infrastructure Engineering*, Vol. 15, No. 1, pp. 5-13.