

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Joint Optimization of Pavement Management and Reconstruction Policies for Segment and System Problems

Permalink

<https://escholarship.org/uc/item/7qk6f10d>

Author

Lee, Jinwoo

Publication Date

2015

Peer reviewed|Thesis/dissertation

Joint Optimization of Pavement Management and Reconstruction Policies
for Segment and System Problems

By

Jinwoo Lee

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Engineering – Civil and Environmental Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Samer M. Madanat, Chair

Professor Carlos F. Daganzo

Professor Robert C. Leachman

Summer 2015

Joint Optimization of Pavement Management and Reconstruction Policies
for Segment and System Problems

Copyright © 2015

by

Jinwoo Lee

Abstract

Joint Optimization of Pavement Management and Reconstruction Policies for Segment and System Problems

by

Jinwoo Lee

Doctor of Philosophy in Civil and Environmental Engineering

University of California, Berkeley

Professor Samer Madanat, Chair

This dissertation presents a methodology for the joint optimization of a variety of pavement construction and management activities for segment and system problems under multiple budget constraints. The objective of pavement management is to minimize the total discounted life time costs for the agency and the highway users by finding optimal policies. The scope of the dissertation is focused on continuous time and continuous state formulations of pavement condition. We use a history-dependent pavement deterioration model to account for the influence of history on the deterioration rate.

Three topics, representing different aspects of the problem are covered in the dissertation. In the first part, the subject is the joint optimization of pavement design, maintenance and rehabilitation (M&R) strategies for the segment-level problem. A combination of analytical and numerical tools is proposed to solve the problem. In the second part of the dissertation, we present a methodology for the joint optimization of pavement maintenance, rehabilitation and reconstruction (MR&R) activities for the segment-level problem. The majority of existing Pavement Management Systems (PMS) do not optimize reconstruction jointly with maintenance and rehabilitation policies. We show that not accounting for reconstruction in maintenance and rehabilitation planning results in suboptimal policies for pavements undergoing cumulative damage in the underlying layers (base, sub-base or subgrade). We propose dynamic programming solutions using an augmented state which includes current surface condition and age. In the third part, we propose a methodology for the joint optimization of rehabilitation and reconstruction activities for heterogeneous pavement systems under multiple budget constraints. Within a bottom-up solution approach, Genetic Algorithm (GA) is adopted. The complexity of the algorithm is polynomial in the size of the system and the policy-related parameters.

Contents

Contents	i
Acknowledgments.....	iii
1. Introduction	1
2. Literature Review	3
2.1 Pavement Design	3
2.2 Pavement Management Optimization	4
2.2.1 Segment-Level Problems	4
2.2.2 System-Level Problems	4
3. Joint Optimization of Pavement Design, Resurfacing and Maintenance Strategies with History-Dependent Deterioration Models	6
3.1 Optimization Problem Formulation	6
3.1.1 General formulation	6
3.1.2 Cost and Performance Models	8
3.2 Solution Approach	10
3.2.1 Optimality Conditions	10
3.2.1.1 Case of $s(t_i^+) > s_0$	11
3.2.1.2 Case of $s(t_i^+) = s_0$	14
3.2.2 Finding the Optimal Solution of (t, w) for Given $(SN, \Delta b)$	15
3.2.3 Algorithm to Find Optimal Solution of $(SN, \Delta b)$	16
3.3 Parametric Study	16
3.4 Discussion	22
4. Jointly Optimal Policies for Pavement Maintenance, Resurfacing and Reconstruction	24
4.1 Optimization Problem Formulation	24
4.1.1 General Formulation	24
4.1.2 Cost and Performance Models	26
4.2 Solution Approach	28
4.3 Parametric Examples	31

4.3.1 A Special Case : Memoryless Model	32
4.3.2 General Case: History-Dependent Model	35
4.3.3 Sensitivity Study to Uncertainty in the Deterioration Process	39
4.4 Discussion	40
5. A Joint Bottom-Up Solution Methodology for System-Level Pavement Rehabilitation and Reconstruction	42
5.1 Optimization Problem Formulation	42
5.1.1 General Formulation	42
5.1.2 Cost and Performance Models	44
5.2 Solution Approach	46
5.2.1 Segment-Level Optimization	46
5.2.2 System-Level Optimization	48
5.2.3 Discussion of the Case of Two Budget Constraints	50
5.3 Case Study	52
5.3.1 Parameters and Assumptions	52
5.3.2 Segment-Level Results	54
5.3.3 System-Level Results	56
5.4 Discussion	60
6. Conclusion	61
6.1 Contributions	61
6.2 Future Works	62
Bibliography	63
Appendix A Mathematical Details on the First Order Necessary Condition for Optimality of Rehabilitation	68
Appendix B Proof of Proposition 1	69
Appendix C Proof of $w^* = R(s(\tau\delta^-))$	71
Appendix D Segment-level Alternatives Enumeration Algorithms	73

Acknowledgements

The process of pursuing a doctorate and writing this dissertation is long and arduous but, thanks to the amazing people I met in the ITS program for their precious guidance and help I received. This dissertation would not have been possible without their great deal of help.

Most of all, I have been fortunate that my advisor, Prof. Samer Madanat, who has spent a lot of hours to help me develop the dissertation, who has helped me to solve all the practical, methodical and theoretical issues, who has guided me through the past four years since I enrolled in this program, with patience, understanding and tolerance. Prof. Madanat's encouragement and continuous guidance not only taught me how to become a researcher but also taught me how to become a more mature person. He is not just my advisor, also my life mentor.

I would also like to thank my committee members: Prof. Carlos Daganzo and Prof. Robert Leachman, I greatly appreciate them being incredibly accommodating with reviewing and providing invaluable feedback and reviews in alternative ways. Those thoughtful comments and suggestions have significantly improved the quality of my dissertation.

I would like to express my sincere appreciation and respect to the ITS faculty. This dissertation cannot finish without the academic knowledge and skills I learned from their professional and awesome classes. Learning from them was one of the most enjoyable and unforgettable experience I had at UC Berkeley.

Also, thanks to the SafeTREC members for being friendly, caring and helpful in various ways. I immensely enjoyed and learned from every meeting I had with them, not just research, also a variety of things of life.

I would also like to mention about my friends in ITS, especially in Room 116. I am lucky to have each and every of them being a part of my graduate school life.

A special thanks to Dr. Koohong Chung. Thanks for helping me all the time. Not only as a friend, but also, a big brother.

Last, but never least, thank my family, for their continuous and unconditional support that they have given to me. Thanks them for always believing me, helping me, encouraging me and being so strong.

Additionally, this research was partly funded by a University of California Transportation Center dissertation grant.

Chapter 1

Introduction

Governments in many countries need to construct and manage transportation infrastructure to meet excessive demand, under strict budget constraints. Insufficient transportation capacity forces heavier traffic onto restricted roadways, which accelerates the aging of the infrastructure. As a consequence, poor road conditions yield higher user costs such as vehicle operating costs and travel time delay costs. Due to the large construction, reconstruction and maintenance costs, as well as high user costs, it is necessary to address simultaneously the problems of different kinds of pavement projects.

In this dissertation, three sub-topics are addressed in three main parts respectively based on the size of the pavement system that we look into and the pavement projects considered as decision factors.

The first part of the dissertation addresses the segment-level joint optimization problem between pavement design and future management strategies. Pavement design refers to the selection of the thickness of the pavement layers, with consideration of the properties of the materials, environmental conditions and expected traffic load. After pavement facilities have been built, it is necessary to allocate limited resources for future management activities such as preventive maintenance and rehabilitation. Pavement deterioration rates are influenced by pavement design. More durably designed pavements require less frequent rehabilitation and a lower level of maintenance. Therefore, pavement design influences future management strategies after construction. Furthermore, pavement design directly affects the construction costs, which are very large and need to be invested all at one time. The purpose of pavement design optimization is to minimize discounted lifecycle costs, including user costs, and agency costs.

The second part of the dissertation presents the joint optimization of various pavement management treatments for segment-level problem. We include frequent maintenance activities aimed at slowing down the deterioration process, rehabilitation activities aimed at reducing pavement roughness and reconstruction activities aimed at complete renewal of pavement condition and its age. We develop a methodology to determine the optimal times and intensities to perform routine maintenance and rehabilitation and the optimal time to perform reconstruction to minimize the total discounted life time costs over an infinite time horizon.

In the last part of the dissertation, the problem of optimizing rehabilitation and reconstruction policies for large-scale pavement systems is formulated. We consider situations where rehabilitation and reconstruction projects are funded from separate budgets as well as those where they share the same budget. For example, there can be two independent budgets: a capital budget for construction and reconstruction projects, and a maintenance budget for maintenance and rehabilitation activities. We develop a bottom-up solution methodology for the system-level optimization. A bottom-up approach reflects pavement segment-specific characteristics, and provides individual optimal strategies for each segment. As a bottom-up approach, Genetic algorithms (GAs) are implemented, and the complexity of these algorithms is polynomial in the

size of the system and the policy-related parameters. The performance and deterioration models are deterministic and follow Markovian properties, but consider history-dependent deterioration process. This is achieved by using augmented condition states that include history variables, such as the age of the pavement or the cumulative traffic loading, in addition to the current pavement condition, represented by pavement roughness.

The remainder of the dissertation is organized as following. In Chapter 2, the literature review is presented. In Chapters 3, 4 and 5, three sub-topics are addressed respectively. Chapter 6 concludes the dissertation. The details of the mathematical solution are included in the Appendices.

Chapter 2

Literature Review

The literature review is organized with two parts: (i) pavement design approach (Section 2.1); (ii) and pavement management optimization (Section 2.2). The latter can be sub-divided into two groups, in terms of the number of segments considered: single segment problem (Section 2.2.1) and system-level problem (Section 2.2.2).

2.1 Pavement Design

Pavement design refers to the selection of the thickness of the pavement layers, with consideration of the properties of the materials, environmental conditions and expected traffic loading. After pavement segments have been built, it is necessary to allocate limited resources for maintenance and rehabilitation (M&R) activities. Pavement deterioration rates are influenced by pavement design. More durably designed pavements require less frequent rehabilitations and a lower level of maintenance. Therefore, pavement design influences future M&R strategies after construction. Furthermore, pavement design directly affects the construction costs, which are very large and need to be invested all at one time.

The most commonly used pavement design approach in practice is the AASHTO design method (AASHTO 1986). It is based on a regression analysis with experimental data which was collected from only one subgrade material and from only one type of climatic region as well. This method is based on statistical analysis of pavement failure in an accelerated loading experiment. The statistical method used did not appropriately account for censored observations, which led to statistical bias, as Small and Winston (1988) and Madanat et al. (2002) noted.

Small et al. (1988) and Madanat et al. (2002) corrected for censoring bias by improved statistical estimation of the coefficients in the deterioration model based on the Tobit and Weibull models, respectively. These authors also found solutions for the optimization problem of initial pavement design; given a pre-determined rehabilitation strategy.

Pryke et al. (2006) and Rajbongshi and Das (2008) presented mechanistic-empirical design methods, which included better representation of the physical properties of materials. McDonald and Madanat (2012) formulated an optimization problem for mechanistic-empirical pavement design to minimize the total life cycle costs of asphalt/concrete pavements. The tradeoffs between constructing a more durable pavement with a higher structural number at the start of the time horizon, and more frequent rehabilitations, were evaluated quantitatively. Optimal pavement design was determined by nonlinear mathematical programming. The maintenance method assumed in their analysis is limited only to rehabilitations.

2.2 Pavement Management Optimization

2.2.1 Segment-Level Problems

Single facility optimization problems are formulated in several ways. Much of the literature uses Markov Decision Processes (MDP), in which the condition state is discrete (Carnahan et al. 1987; Carnahan, 1988; Feighan et al., 1988; Gopal and Majidzadeh, 1991; Madanat, 1993; Madanat and Ben-Akiva, 1994). For problems formulated with continuous states, various solution approaches exist. Optimal control theory has been applied (Buttler and Shortreed, 1978; Friesz and Fernandez, 1979; Fernandez and Friesz, 1981; Markow and Balta, 1985; Markow et al., 1993; Tsunokawa and Schofer, 1994; Durango-Cohen and Sarutipand, 2007). Jido et al. (2008) extend the work of Madanat (1993) to a continuous-state framework. Li and Madanat (2002) determine the optimal pavement rehabilitation policies based upon the Markovian properties of the deterioration models which are continuous, deterministic and memoryless. Calculus of variations was utilized by Ouyang and Madanat (2006) to solve analytically for the optimal pavement rehabilitation schedule in a finite planning time horizon. Gu et al. (2012) improve on Ouyang and Madanat (2006) to consider multiple activities (maintenance and rehabilitation). In a similar manner, Rashid and Tsunokawa (2012) add decision variables for various treatments, building on Tsunokawa and Schofer (1994) which has a single decision variable (rehabilitation frequency).

Previous studies for infinite planning horizons that do not consider reconstruction policies are either unrealistic in general pavement cases (where reconstruction is mandatory because of accumulated permanent damage), or applicable only to the special case of perpetual pavements, where pavements are designed to have long lives (e.g. Nunn et al., 1997). As Guignier and Madanat (1999) have shown, joint optimization of reconstruction and periodic rehabilitation is beneficial in terms of lifecycle costs. In the latter paper, it is recognized that accounting for reconstruction in the maintenance and rehabilitation problem leads to steady-state solutions which consist of repeated cycles, where each cycle starts with a reconstruction and includes one or more rehabilitation events. To determine the timing of reconstruction, augmented state MDP models have been proposed. The history-dependent bridge maintenance and reconstruction optimization problem is solved in Robelin and Madanat (2007) with an augmented state MDP.

2.2.2 System-Level Problems

For system-level optimization, the objective is to solve for optimal Maintenance, Rehabilitation and Reconstruction (MR&R) policies which minimize the expected lifecycle cost of systems, or maximize the reliability of systems, under limited monetary budgets. Two approaches exist to solve this optimization problem: the top-down approach and the bottom-up approach. The most widely cited example of the top-down approach is the application in the state of Arizona (Golabi et al., 1982), which has been a notable success in achieving savings in road maintenance costs. The top-down approaches (Kuhn and Madanat, 2005; Durango-Cohen and Sarutipand, 2007) are computationally economical in comparison to the bottom-up approaches, but they are applicable only to homogenous systems, and they do not capture facility-specific features. The bottom-up approaches are beneficial in terms of incorporating heterogeneity among facilities in the system, but are more difficult from a computational perspective.

Robelin and Madanat (2008) solve the reliability-based problem for bridge system management with a bottom-up solution methodology. Sathaye and Madanat (2011, 2012) and Chu

and Chen (2012) propose bottom-up approaches for system-level optimization of pavement management with threshold-based decision variables; decision variables are defined as trigger values for one or multiple treatments. The threshold structure allows for mathematical simplifications and makes the optimization approach applicable to the system-level problem. Reger et al. (2014) extend the work of Sathaye and Madanat (2012) to simultaneously incorporate the trade-off between costs and GHG emissions into the system-level pavement management problem.

Genetic Algorithms (GAs) are widely used in infrastructure management (Fwa et al., 1994; Chan et al., 1994; Fwa et al., 1996; Cheu et al. 2004; Chootinan et al., 2006; Yeo et al. 2013) as bottom-up solution methodologies. In GAs, solution algorithms consist of iterative generation of offspring genotypes based upon the parent genotypes until a stopping criterion is satisfied. They take discrete activities as decision variables rather than continuous decision variables.

Chapter 3

Joint Optimization of Pavement Design, Rehabilitation and Maintenance

This chapter is focused on the joint optimization problem of pavement design and M&R strategies. These include frequent maintenance activities aimed at slowing down the deterioration process, and rehabilitation activities aimed at reducing pavement roughness immediately. The effects from maintenance activities and from rehabilitations are assumed to be independent. The scope of this chapter is focused on the case of continuous time and continuous state of pavement condition. The pavement deterioration model of Paterson (1987) is history-dependent. The deterioration rate depends not only on the current condition of the pavement, but also on its history of rehabilitation activity. Thus, a recently resurfaced pavement has a slower deterioration rate than a pavement resurfaced previously, even though both current surface conditions are identical. Additionally, pavement deterioration depends on traffic loading and the structural number of the pavement (i.e., heavier traffic causes faster deterioration and thicker pavements retard deterioration). The expected traffic loading is assumed to be fixed and exogenous in the optimization problem, and not responsive to pavement condition.

In Section 3.1, a mathematical formulation of the optimization problem is presented. In Section 3.2, a methodology for solving the joint optimization problem of design and M&R is presented. The solution methodology is utilized in a set of parametric studies in Section 3.3. The discussion is presented in Section 3.4.

3.1 Optimization Problem Formulation

In this section, the optimization problem is formulated with general formulation (Section 3.1.1) and cost and performance models (Section 3.1.2).

3.1.1 General Formulation

The objective of the optimization is to minimize the net present value of discounted lifecycle costs over an infinite time horizon, including initial construction costs, agency investments for both maintenance activities and rehabilitations, and user costs as shown in (3.1a). The decision variables consist of the structural number, SN , in terms of pavement thickness, selected at $t = 0$ and assumed to be constant throughout the planning horizon, the intensities (i.e. thickness), $\bar{w} = \{w_1, w_2, \dots\}$, and timings, $\bar{t} = \{t_1, t_2, \dots\}$, of rehabilitations consecutively numbered as $i = 1, 2, 3, \dots$, and the level of maintenance activities, Δb , which is expressed as a reduction in the deterioration rate: the deterioration rate b decreases as the intensity of maintenance activities increases. The assumption of constant pavement thickness is reasonable, because thickness is constant if the agency applies a rehabilitation after removing the top-most layer of asphalt concrete, which is the practice in California and in many countries. Given the above the objective function and decision variables, the problem can be formulated as following;

$$\min_{SN, \bar{t}, \bar{w}, \Delta b} J = K[SN] + \sum_{i=0}^{\infty} \int_{t_i^+}^{t_{i+1}} (C[s(t)] + C_M[s(t), \Delta b]) e^{-rt} dt + \sum_{i=1}^{\infty} M[w_i] e^{-rt_i} \quad (3.1a)$$

Subject to

$$s(t_i) - s(t_i^+) = G(w_i, s(t_i)), \quad \forall i = 1, 2, \dots \quad (3.1b)$$

$$s(t) = F(s(t_{i-1}^+), t - t_{i-1}, b, SN), \forall t \in [t_{i-1}^+, t_i], i = 1, 2, \dots \quad (3.1c)$$

$$s(t_i^+) \geq s_0, \quad \forall i = 1, 2, \dots \quad (3.1d)$$

$$0 \leq w_i \leq R(s(t_i)), \quad \forall i = 1, 2, \dots \quad (3.1e)$$

$$t_0 = 0, t_i \geq t_{i-1}, \quad \forall i = 1, 2, \dots \quad (3.1f)$$

$$b = b_0 - \Delta b \quad (3.1g)$$

$$b_{min} \leq b \leq b_0 \quad (3.1h)$$

$$SN_{min} \leq SN \leq SN_{max} \quad (3.1i)$$

$$s(0) = s_{new} \quad (3.1j)$$

where,

J : net present value of lifecycle costs

$s(t)$: pavement roughness at time t , and a higher value of $s(t)$ refers to worse pavement condition

$K[SN]$: construction costs at the initial point in time, an increasing function of SN

$C[s(t)]$: user costs per unit time, an increasing function of $s(t)$

$C_M[s(t), \Delta b]$: agency costs for routine maintenance activities per unit time, an increasing function of both $s(t)$ and Δb

$M[w_i]$: agency costs for the i^{th} rehabilitation, an increasing function of w_i

r : discount rate, assumed to be constant

t_i^+ : time right after the i^{th} rehabilitation; $t_i^+ \equiv t_i + \epsilon$ for small ϵ

$G(w_i, s(t_i))$: effectiveness of the i^{th} rehabilitation as a function of intensity w_i and the current pavement roughness, $s(t_i)$

s_0 : the best achievable roughness after rehabilitation

$F(s(t_{i-1}^+), t - t_{i-1}, b, SN)$: non-Markovian deterioration model

$R(s(t_i))$: upper bound of the effective rehabilitation intensity, a function of current pavement roughness

b_0 : deterioration rate without maintenance activities

b_{min} : minimum achievable level of the deterioration rate by applying maintenance activities

SN_{min}, SN_{max} : lower and upper bounds of structural number

s_{new} : pavement roughness achieved by construction; $s_0 > s_{new} > 0$ due to technological reasons

By the definition of t_i^+ , we can define the pavement roughness as an injective function expanded in the continuous time horizon. Constraint (3.1b) describes the rehabilitation effectiveness as a reduction in roughness. Constraint (3.1c) describes the non-Markovian deterioration process. The term, $t - t_{i-1}$, defines the history-dependent properties. The best achievable level of pavement condition is constrained in (3.1d). Constraint (3.1e) describes the range of effective rehabilitation intensity, w_i . The improvement in pavement condition, $G(w_i, s(t_i))$, does not increase with w_i beyond $R(s(t_i))$ (Ouyang and Madanat, 2004). Constraint (3.1f) indicates the sequential order of rehabilitations and $t_0 = 0$. The deterioration rate is defined by (3.1g) and constrained by (3.1h). Constraint (3.1i) describes the range of the structural number. The initial pavement roughness is given in (3.1j), which is always lower than the best achievable pavement roughness by rehabilitation, s_0 , because of permanent loss.

3.1.2 Cost and Performance Models

In this section, we describe the pavement deterioration, cost and effectiveness models. We use the pavement deterioration model developed by Paterson (1987), shown in (3.2). The pavement roughness function is continuous between consecutive rehabilitations.

$$F(s(t_{i-1}^+), t - t_{i-1}, b, SN) = (s(t_{i-1}^+) + a \cdot l \cdot (t - t_{i-1}) \cdot (1 + SN)^q) e^{b(t-t_{i-1})}, \forall t \in [t_{i-1}^+, t_i], i = 1, 2 \dots \quad (3.2)$$

In (3.2), a and q are constant parameters. From empirical research, it is known that q has negative value, so a higher structural number, SN , slows down the deterioration process. The traffic loading per unit time is denoted as l , and $l \cdot (t - t_{i-1})$ indicates the cumulative traffic loading from the time of last rehabilitation. We assume that l is constant, and independent of pavement condition and M&R activities. In practice, the traffic loading is influenced by pavement condition, as well as the deterioration process depends on the traffic loading. In addition, an implementation of M&R activities on pavement causes significant user costs from travel time delay due to the loss in network capacity. These dynamic interactions among the traffic loading, pavement condition and M&R strategies are influenced by the configuration of a pavement

network, but this research is focusing on a single segment, so we assume the constant traffic loading along the planning horizon to simplify the problem. We introduce the term $A(SN)$, a function of the structural number, to denote $a(1 + SN)^q$.

The construction cost per unit length of a new pavement section is expressed as (3.3). It is a linear function of the structural number, SN . This is realistic because the construction cost increases as larger quantities of materials are used to obtain a higher structural number. The constant term, k_1 , is the unit cost of pavement materials per unit of SN . The constant term, k_2 , does not influence the optimal solution, so it is assumed to be 0.

$$K[SN] = k_1 \cdot SN + k_2 \quad (3.3)$$

User costs are proportional to traffic volume, v , and mainly consist of vehicle operating costs and travel time delays. Travel time delay is significant when a roadway is being rehabilitated, but assumed to be independent of pavement roughness. Free flow speed and highway capacity may decrease with pavement roughness, but these effects can be ignored except in the case of extremely bad pavement condition. These assumptions can be easily relaxed without adding complexity to the solution approach. Vehicle operating costs are assumed to be linearly related to pavement roughness to be consistent with our most recent work on the subject (Gu et al. 2012) to facilitate the comparison of our results, even though this assumption may not be always true (Watanabe et al. 1987; Barnes and Langworthy 2004; Zabaar and Chatti 2010). Thus, the vehicle operating cost rate per unit time is formulated as a linear function of pavement roughness as shown in (3.4), where c_1 and c_2 are positive parameters. The relative weight between user cost and agency cost is considered by multiplying weight factor, $W_{u/a}$. The user cost commonly overwhelms the agency cost, but we assume $W_{u/a} = 1$ to be consistent with the state of the art (e.g. Gu et al, 2012).

$$C[s(t)] = W_{u/a} \cdot v \cdot (c_1 \cdot s(t) + c_2), \forall t \in [0, \infty) \quad (3.4)$$

The rehabilitation cost and effectiveness models, shown in (3.5) and (3.6), are taken from Ouyang and Madanat (2004), where m_1 , m_2 , μ_1 , μ_2 and μ_3 are positive constant parameters. The intensity of rehabilitations, w_i , is shown in (3.7). The component of user costs that corresponds total travel time delay caused by partial road closure is denoted by $C_2[v]$. A practical formula to calculate travel time delay is developed by the U.S Bureau of Public Roads (BPR). In this research, we assume that rehabilitation activities occur only during the nighttime, as is the practice in California, so $C_2[v] = 0$. This assumption can be easily relaxed.

$$M[w_i] = m_1 w_i + m_2, \quad \forall i = 1, 2, \dots \quad (3.5)$$

$$G(w_i, s(t_i)) = \frac{\mu_1 w_i s(t_i)}{\mu_2 s(t_i) + \mu_3}, \quad \forall i, \quad 0 \leq w_i \leq R(s(t_i)) = \mu_2 s(t_i) + \mu_3 \quad (3.6)$$

$$w_i = \frac{1}{\mu_1} (s(t_i) - s(t_i^+)) \left(\mu_2 + \frac{\mu_3}{s(t_i)} \right) \quad (3.7)$$

The maintenance cost per unit time is taken from Gu et al. (2012) as shown in (3.8), where $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1$ and γ_2 are positive constant parameters. The maintenance cost per unit time is positively related to Δb , and it is a linear function of pavement roughness. As shown in the literature (Ponnia and Kennepohl 1996; Labi and Sinha 2003), routine maintenance activities slow down the deterioration of pavement and the deterioration rate decreases as the maintenance intensity increases. Thus the maintenance intensity is denoted as the reduction in the deterioration rate. The maintenance cost per unit time is exponentially related to Δb and it is a linear function of pavement roughness.

$$C_M[s(t), \Delta b] = \begin{cases} C_{M,1} \cdot s(t) + C_{M,2}, & \text{if } \Delta b > 0 \\ 0, & \text{if } \Delta b = 0 \end{cases} \quad (3.8)$$

where,

$$C_{M,1} = \alpha_1 e^{\beta_1 \Delta b(\tau)} + \gamma_1 \text{ and } C_{M,2} = \alpha_2 e^{\beta_2 \Delta b(\tau)} + \gamma_2$$

3.2 Solution Approach

The optimization problem formulated in this chapter is a mixed-integer (binary) nonlinear program (MINLP), and it has four categories of decision variables: SN , $\bar{t}$, $\bar{w}$ and Δb . It is difficult to solve this problem by using commercial solvers such as CPLEX, because the formulations of this problem do not follow the standard forms, which commercial solvers are applicable. In the solution methodology, we define the problem as a bi-level optimization problem. In Section 3.2.2, We find the lower-level optimal solution of rehabilitation policies, $(\bar{t}, \bar{w})$, with fixed $(SN, \Delta b)$, according to the proposed algorithm shown in Section 3.2.1. If we find the lower-level optimal rehabilitation policies given $(SN, \Delta b)$ by this algorithm, SN and Δb are left as the decision variables of the upper-level optimization problem. The solution algorithm to determine SN and Δb is stated in Section 3.2.3.

3.2.1 Optimality Conditions

In order to find the optimal rehabilitation policies given SN and Δb , the algorithm uses infinitesimal perturbations to t_i and w_i as in Ouyang and Madanat (2006). Infinitesimal perturbations for each rehabilitation yield a locally optimal policy. Globally optimal rehabilitation policy is obtained from locally optimal policies for all rehabilitations in the planning horizon. From now on, we use a simple expression of the deterioration function, $F(s(t_{i-1}^+), t - t_{i-1})$, instead of $F(s(t_{i-1}^+), t - t_{i-1}, b, SN)$ in (3.2), because Δb and SN are fixed in the lower-level problem. The optimal times and intensities of rehabilitations will vary with different traffic loading, and the values of parameters. Two cases are considered:

1. The case where local optimum of pavement roughness achieved by rehabilitation i is not the best achievable pavement roughness, $s(t_i^+) > s_0$.
2. The case of $s(t_i^+) = s_0$.

$$\begin{aligned}
\frac{dJ}{d\eta_i} = & \left\{ c' \cdot \int_0^{\tau_{i+1}} \frac{d}{d\eta_i} F(F(s(t_i), \tau_i) - G(w_i + \eta_i, s(t_i)), u) e^{-ru} du \right. \\
& + \frac{d}{d\eta_i} M(w_i + \eta_i) \\
& + \frac{d}{d\eta_i} M \left(w_{i+1} - \eta_i F'_{s(t_i^+)}(s(t_i^+), t_{i+1} - t_i) \frac{G'_{w_i}(w_i, s(t_i))}{G'_{w_{i+1}}(w_{i+1}, s(t_{i+1}))} \right) e^{-r\tau_{i+1}} \Big\} \\
& \cdot e^{-rt_1}
\end{aligned} \tag{3.10}$$

where,

$$c' \equiv C'_{M_{s(t_i^+)}}[s(t), \Delta b] + C'_{s(t_i^+)}[s(t)]$$

$$\tau_i \equiv t_i - t_{i-1}, \forall i = 1, 2, \dots$$

$\frac{dJ}{d\eta_i}$ in (3.10) has a similar expression to $\frac{dJ}{d\eta_i}$ in Ouyang and Madanat (2006). Only $C'_{s(t_i^+)}[s(t)]$ used in Ouyang and Madanat (2006) is replaced by c' , but this does not affect the validity of the proposition. That is, the optimal intensities of all rehabilitations, w_i , are the maximum intensities, $R(s(t_i))$, as shown in (3.11).

$$w_i = R(s(t_i)), \forall i = 1, 2, \dots \tag{3.11}$$

Suppose t_i is perturbed to $t_i + \varepsilon_i$. An alternative rehabilitation policy [S – A – B' – C'] is constructed, and w_i and w_{i+1} are perturbed simultaneously. Because we suppose that [S – A – B – C] is part of an optimal rehabilitation policy, so w_i at t_i is $\mu_2 s(t_i) + \mu_3$, and w_i at $t_i + \varepsilon_i$ is $\mu_2 s(t_i + \varepsilon_i) + \mu_3$. The point B' is not on the curve [B – C], because $\frac{dF(s(t_i^+), \varepsilon_i)}{d\varepsilon_i} > (1 - \mu_1) \cdot \frac{dF(s(t_{i-1}^+), \tau_i + \varepsilon_i)}{d\varepsilon_i}$. At the $i + 1^{th}$ rehabilitation, the rehabilitation policy changes to [C' – D] from [C – D]. In (7), $s(t_{i+1}^+)$ is fixed, but only $s(t_{i+1})$ is perturbed. A first necessary condition of the locally optimal curve [S – A – B' – C'] is (3.12) where $\varepsilon_i = 0$. Except the case of $\frac{dJ}{d\varepsilon_i} = 0$, the optimal rehabilitations, constructed with $\frac{dJ}{d\varepsilon_i} \geq 0$ or $\frac{dJ}{d\varepsilon_i} \leq 0$, occur only when the initial condition is poor or at and the end point of the finite planning horizon. This cannot happen in the case of an infinite time horizon and a newly constructed segment. Therefore, $\frac{dJ}{d\varepsilon_i} = 0$ is the local optimality condition. We obtain the first order necessary condition for optimality of the i^{th} rehabilitation as (3.13). $\frac{dJ}{d\varepsilon_i}$ is a function of $s(t_{i-1}^+)$ (point “S” in Figure 1), τ_i and τ_{i+1} (see (3.13) and Appendix A). The second order condition, $\frac{d^2J}{d\varepsilon_i^2} \geq 0$, is satisfied, when it is calculated with generally used parameters.

$$\frac{dJ}{d\varepsilon_i} \begin{cases} \geq 0, \text{ if } t_{i-1} = t_i \\ = 0, \text{ if } t_i \in (t_{i-1}, t_{i+1}) \\ \leq 0, \text{ if } t_i = t_{i+1} \end{cases} \quad (3.12)$$

$$\begin{aligned} \frac{dJ}{d\varepsilon_i} = & \left\{ c' \frac{d}{d\varepsilon_i} \left\{ \int_0^{\varepsilon_i} (F(s(t_{i-1}^+), \tau_{i+1} + u) - F(s(t_i^+), u)) e^{-ru} du \right. \right. \\ & + \int_{\varepsilon_i}^{\tau_{i+1}} (F((1 - \mu_1)F(s(t_{i-1}^+), \tau_i + \varepsilon_i), u - \varepsilon_i) \\ & \left. \left. - F(s(t_i^+), u)) e^{-ru} du \right\} \right. \end{aligned} \quad (3.13)$$

$$\begin{aligned} & + \frac{d}{d\varepsilon_i} \{ M[\mu_2 F(s(t_{i-1}^+), \tau_i + \varepsilon_i) + \mu_3] e^{-r\varepsilon_i} \} \\ & + \frac{d}{d\varepsilon_i} \{ M[\mu_2 F((1 - \mu_1)F(s(t_{i-1}^+), \tau_i + \varepsilon_i), \tau_{i+1} - \varepsilon_i) + \mu_3] e^{-r\tau_{i+1}} \} \} e^{-rt_i} = 0 \end{aligned}$$

Steady state conditions are defined as $s(t_i) = s(t_{i+1})$ and $\tau_i = \tau_{i+1}$. We define two new terms: τ^* is the optimal time between two consecutive rehabilitations and s^* is the optimal trigger roughness when we assume that the optimal rehabilitation policy reaches a steady state.

Proposition 1: On an optimal roughness rehabilitation policy, $\tau_i \geq \tau_{i+1} \geq \tau^*$ and $s(t_i) \leq s(t_{i+1}) \leq s^*$, $\forall i$ if $s(t_1^+) > s_0$.

Proof: See Appendix B.

If the rehabilitation policy reaches a steady state, then equation (3.14) is derived based on the deterioration process (3.2) and the optimality condition (3.11). In (3.14), because the left hand side and s^* are positive, $1 - (1 - \mu_1)e^{b\tau^*}$ must be positive, so τ^* is restricted by (3.15). There are two points of $\tau > 0$, satisfying (3.14), but only the smaller one is satisfying (3.15). Consequently, we can find a unique τ^* by numerical methods such as the Newton method. The optimal trigger roughness is found by (3.16).

$$\tau^* \cdot A(SN) \cdot l \cdot e^{b\tau^*} = (1 - (1 - \mu_1)e^{b\tau^*})s^*, \forall i = 1, 2, \dots \quad (3.14)$$

$$\tau^* \in \left(0, \frac{-\log(1 - \mu_1)}{b} \right), \forall i = 1, 2, \dots \quad (3.15)$$

$$s^* = \frac{\tau^* \cdot A(SN) \cdot l \cdot e^{b\tau^*}}{1 - (1 - \mu_1)e^{b\tau^*}} \quad (3.16)$$

3.2.1.2 Case of $s(t_i^+) = s_0$

Suppose we have a part of an optimal rehabilitation policy, $[S - A - B - C]$, as shown in Figure 3.2. If t_i is perturbed into $t_i + \varepsilon_i$, then we obtain an alternative rehabilitation policy $[S - A' - B' - C']$.

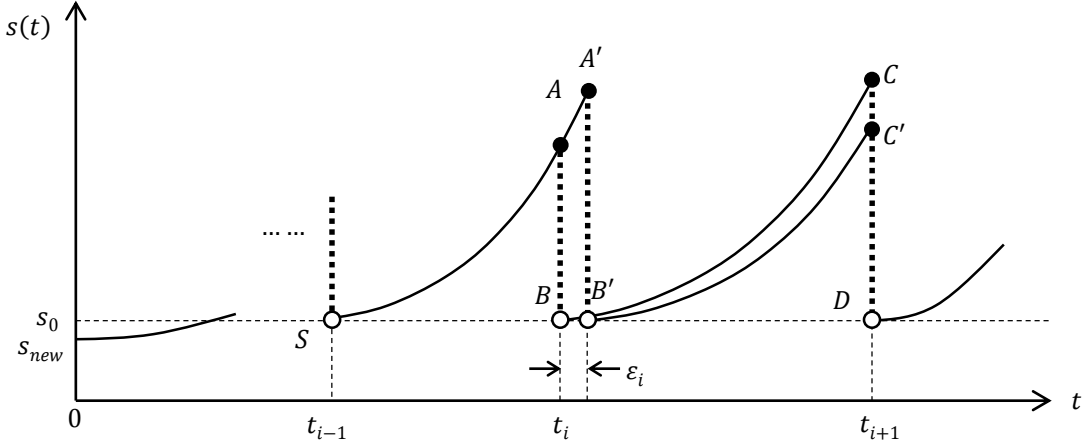


FIGURE 3.2 Alternative rehabilitation policy when $s(t_i^+) = s_0$

Note that the inequality constraint (3.1d) is binding as $s(t_i^+) = s_0$, and an optimal rehabilitation intensity, w_i , is less than R_i . The first order derivative of w_i with respect to $s(t_i)$ is expressed as (3.17), which is different from the previous case of $s(t_i^+) > s_0$. By Substituting (3.17) for both $\frac{dw_i}{ds(t_i)}$ and $\frac{dw_{i+1}}{ds(t_{i+1})}$ into the first order necessary condition for optimality of the i^{th} rehabilitation, shown in (3.18), we can formulate the first order necessary condition as a function of $s(t_{i-1}^+)$, τ_i and τ_{i+1} .

$$\frac{dw_i}{ds(t_i)} = \frac{1}{\mu_1} \left(\mu_2 + \frac{s_0 \mu_3}{s(t_i)^2} \right) \quad (3.17)$$

$$\frac{dJ}{d\varepsilon_i} = \left\{ c' \cdot \frac{d}{d\varepsilon_i} \left\{ \int_0^{\varepsilon_i} (F(s(t_{i-1}^+), \tau_{i+1} + u) - F(s_0, u)) e^{-ru} du \right. \right. \\ \left. \left. + \int_{\varepsilon_i}^{\tau_{i+1}} (F(s_0, u - \varepsilon_i) - F(s_0, u)) e^{-ru} du \right\} \right\} \quad (3.18)$$

$$\left. + M'_{w_i}(w_i) \frac{dw_i}{ds(t_i)} \frac{ds(t_i)}{d\varepsilon_i} - r \cdot M(w_i) + M'_{w_{i+1}}(w_{i+1}) \frac{dw_{i+1}}{ds(t_{i+1})} F'_{\varepsilon_i}(s_0, \tau_{i+1} - \varepsilon_i) e^{-r\tau_{i+1}} \right\} e^{-rt_i} \\ = 0$$

We can find the same properties of convergence for an optimal solution in the previous section: $\tau_i \geq \tau_{i+1} \geq \tau^*$ and $s(t_i) \leq s(t_{i+1}) \leq s^*$ for all i . We can also find an optimal solution in a steady state numerically based on (3.18) when (3.1d) is binding.

3.2.2 Finding the Optimal Solution of (t, w) for Given $(SN, \Delta b)$

In this section, we propose a simple method to obtain the optimal solution of (t, w) based on the local optimality presented in section 3.2.1.1 and 3.2.1.2. Equation (3.13) and (3.18) have two unknown variables, τ_i and τ_{i+1} and there is a unique τ_{i+1} correspondent to a certain τ_i , where $\tau_i, \tau_{i+1} \geq \tau^*$. By $s(t_0^+) = s_{new}$, we can find a narrow range of τ_1 satisfying $\tau_1 \geq \tau_2 \geq \tau^*$. For each possible value of τ_1 in the feasible range, we can find $\tau_2, \tau_3, \dots$ successively. We omit a candidate of τ_1 when τ_i and $s(t_i)$, derived from the candidate value of τ_1 , violate the properties of convergence. We assume that a steady state starts from the n^{th} rehabilitation when (3.19) is satisfied.

$$\frac{\tau^* - \tau_{n-1}}{\tau^*} > \theta > \frac{\tau^* - \tau_n}{\tau^*} (> 0) \quad (3.19)$$

where,

θ is defined as a positive allowable error level

We use the bisection method to find τ_1 which satisfies $\tau_i \geq \tau_{i+1} \geq \tau^*, \forall i < n$. $\tau_{1,k}$ and $\Delta\tau_{1,k}$ indicate τ_1 and the interval respectively for the k^{th} iteration in the bisection method. For the first iteration, $\tau_{1,1}$ is the mid-point of the range, satisfying $\tau_1 \geq \tau_2 \geq \tau^*$, and the interval of this range is $\Delta\tau_{1,1}$. For each iteration before reaching a steady state, if $\tau_{i,k} < \tau_{i+1,k}$, go to the next iteration and $\tau_{1,k+1} = \tau_{1,k} - \Delta\tau_{1,k+1}$. If $\tau_{i+1,k} < \tau^*$, go to the next iteration and $\tau_{1,k+1} = \tau_{1,k} + \Delta\tau_{1,k+1}$. If convergence (3.19) is satisfactory, return τ_1 and stop iterating. Finally we obtain the lower-level optimal rehabilitation plan $(\bar{t}, \bar{w})$ for given $(SN, \Delta b)$. The cost functions and performance functions described in the previous section are substituted into the objective function (3.1a). The upper-level objective function, $J(SN, \Delta b)$, is given as (3.20).

$$J(SN, \Delta b) = k_1 \cdot SN + \sum_{i=1}^n e^{-rt_{i-1}} J_i + \frac{e^{-rt_n}}{1 - e^{-r\tau^*}} J^* \quad (3.20)$$

where,

$$J_i = \int_0^{\tau_i} \left((c_1 l + C_{M,1}) \cdot F(s(t_{i-1}^+), u, b, SN) + C_{M,2} \right) e^{-ru} du + M[w_i^*] e^{-r\tau_i}$$

$$J^* = \int_0^{\tau^*} \left((c_1 l + C_{M,1}) \cdot F(\max((1 - \mu_1)s^*, s_0), u, b, SN) + C_{M,2} \right) e^{-ru} du + M[w^*] e^{-r\tau^*}$$

3.2.3 Algorithm to Find Optimal Solution of $(SN, \Delta b)$

In this section, an algorithm to find the optimal solution of $(SN, \Delta b)$ is presented. In the first step, we find optimal solutions $(SN, \Delta b)$ for both cases when maintenance is applied and not applied. The total lifecycle cost corresponding the convex feasible set of $(SN, \Delta b)$ is not differentiated by SN and Δb , because objective value with lower-level optimal future rehabilitation strategies given a set of $(SN, \Delta b)$ is numerically obtained. Note that the objective value is the sum of construction costs, maintenance costs, user costs and rehabilitation costs. For the case when maintenance is applied, the construction cost is linearly related to N and independent of Δb , so it is convex on $(SN, \Delta b)$. In the special case when the maintenance cost is independent of pavement roughness (i.e. $C_{M,1} = 0$), the maintenance cost is strictly convex (see (3.8)), so the sum of the construction cost and the maintenance cost is strictly convex. As SN and Δb increase, trigger roughness of each rehabilitation decreases and elapsed time between consecutive rehabilitations increases due to slow deterioration process according to the lower-level optimality. Therefore, the user cost and the rehabilitation cost are continuous and strictly decreasing along any positive direction of $(SN, \Delta b)$, so is the sum of user cost and rehabilitation cost. The sum of user cost and rehabilitation cost is numerically found, and it generally happens to have a convex shape. The total lifecycle cost is the sum of convex and strictly convex functions, so it is guaranteed to have a global optimal. This global optimization problem is solved by heuristic methods, such as fitness approximation method, simulated annealing and particle swarm optimization. In the general case when $C_{M,1} \neq 0$, the maintenance cost, the user cost and the rehabilitation cost may not be convex or strictly convex, but we can observe that $J(SN, \Delta b)$ happens to have a convex shape according to the random samples from the parametric study in Section 4, so heuristic methods for the global optimization problem are applicable in this case. For the case when maintenance is not applied (i.e. $C_{M,1} = C_{M,2} = 0$), it has a global optimal because it is a subset of the case of $C_{M,1} = 0$. In the second step, we compare the values of the objective function for both cases and determine whether applying maintenance is optimal or not.

3.3 Parametric Study

This section demonstrates the proposed algorithm to solve the joint optimization problem with pavement design and M&R strategies for different traffic loading by a parametric study. The parameters, b_0 and $\underline{b}$, in the non-Markovian deterioration model are selected to have a similar deterioration speed to the Markovian models used in previous works (Ouyang and Madanat, 2006; Gu et al., 2012). The parameters of maintenance costs are taken from Gu et al. (2012). We use parameter values for a single pavement with 1 kilometer long of one lane highway as shown in Table 3.1. Traffic volume is simply assumed to be proportional to traffic loading. If necessary, this assumption can be easily relaxed. The solution algorithms are programmed in MATLAB, and the optimization problems are solved on a Windows 7 Professional OS with a 2.93 GHz processor and 3 GB RAM.

Parameter	Value	Units	Parameter	Value	Units
c_1	5.91	\$/IRI/AADT	γ_2	500	\$/km/lane/yr
c_2	0	\$/AADT	b_0	0.04	-
m_1	11,000	\$/mm /km/lane	b_{min}	0.02	-
m_2	150,000	\$/km/lane	k_1	15,512	\$/SN/ km /lane
μ_1	0.66	-	SN_{min}	7	-
μ_2	7.15	mm/IRI	SN_{max}	11	-
μ_3	18.3	mm	s_{new}	1.0	IRI
r	0.07	-	s_0	1.2	IRI
α_1	156	\$/IRI /km/lane/yr	a	725	-
β_1	240	-	q	-5	-
γ_1	780	\$/IRI /km/lane/yr	l	0.4~0.8	million ESAL /km/lane/yr
α_2	100	\$/km/lane/yr	v	1467~2933	AADT
β_2	240	-			

∴ dimensionless or omitted; *IRI*: International Roughness Index (m/km); *ESAL*: Equivalent Single Axle Load; AADT: Annual Average Daily Traffic

TABLE 3.1 Parameters used in the parametric study

The optimal pavement designs and M&R strategies for different l are summarized in Table 3.2. An allowable error, θ , is set as 10^{-2} . In every case, an optimal rehabilitation policy reaches the steady state at most in five rehabilitations in the lower-level optimization within 10 iterations by using the bisection method, and it is not significantly sensitive to the magnitude of the allowable error. The table shows that higher structural number, lower trigger roughness of rehabilitation and higher maintenance level are jointly needed as the traffic loading increases, in general. However, as shown in the cases of $l = 0.5$ and $l = 0.6$, lower structural number and longer rehabilitation period are optimal in heavier traffic, because routine maintenance is applied only when $l = 0.6$. In lighter traffic situations ($l = 0.4$ and 0.5), applying maintenance is suboptimal.

l	J ($10^5\$$)	SN	τ_1	τ^*	s^*	Δb	Const. costs ($10^5\$$)	User costs ($10^5\$$)	Maint. costs ($10^5\$$)	Rehab. costs ($10^5\$$)
0.4	4.11	8.5	37.2	25.8	5.83	0	1.32	2.22	0	0.569
0.5	4.71	9.1	33.5	25.5	4.37	0	1.41	2.70	0	0.602
0.6	5.26	9.0	34.1	29.7	3.52	0.0072	1.40	2.88	0.520	0.465
0.7	5.79	9.4	33.7	29.3	3.45	0.0073	1.46	3.35	0.525	0.460
0.8	6.31	9.8	32.7	21.3	3.30	0.0074	1.52	3.80	0.529	0.455

TABLE 3.2 Optimal pavement design and M&R strategies obtained by the proposed methodology

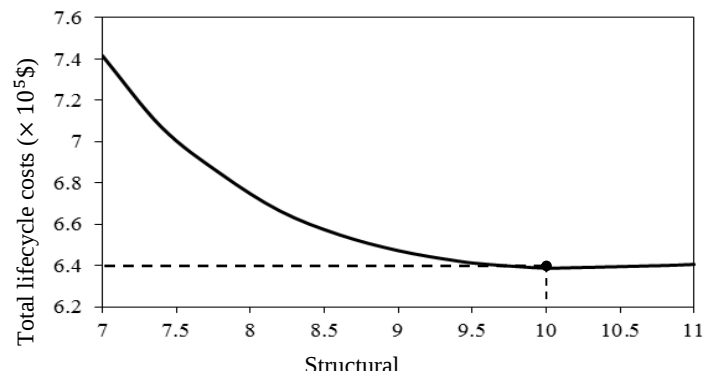


FIGURE 3.3 Total discounted lifecycle costs versus structural number when maintenance is not applied ($l = 0.8$)

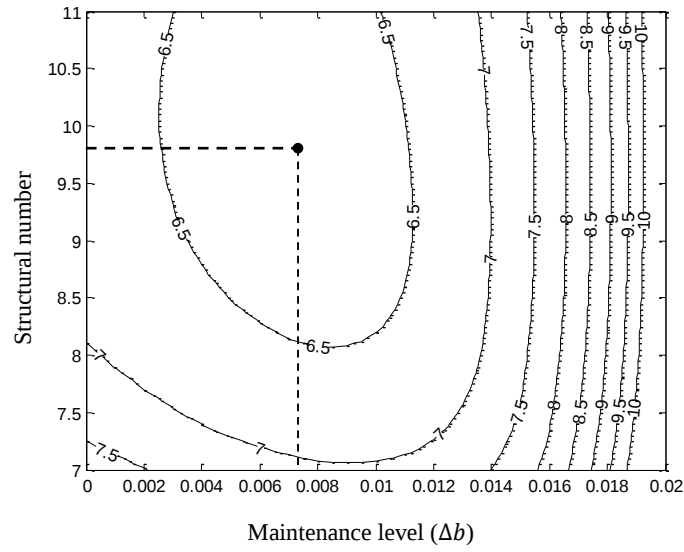


FIGURE 3.4 Total discounted lifecycle costs ($\times 10^5 \$$) when maintenance is applied ($l = 0.8$)

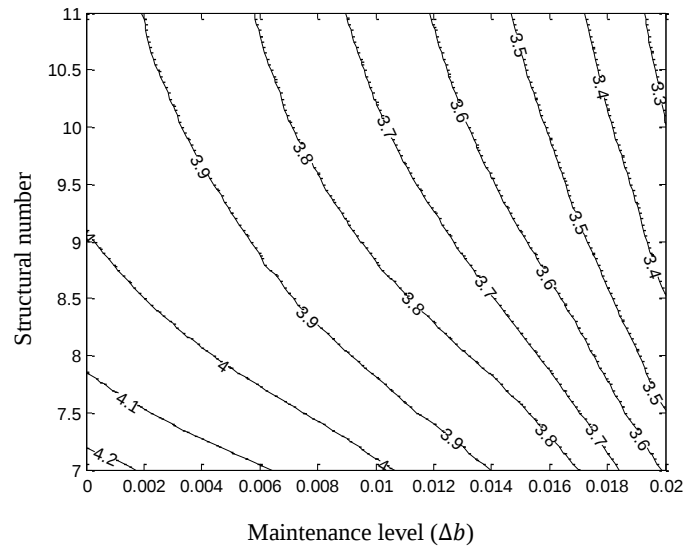


FIGURE 3.5(a) User costs ($\times 10^5 \$$) when maintenance is applied ($l = 0.8$)

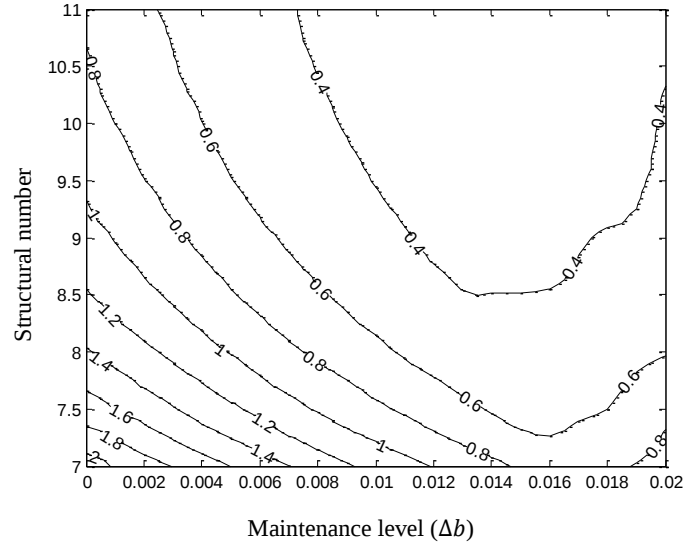


FIGURE 3.5(b) Rehabilitation costs ($\times 10^5 \$$) when maintenance is applied ($l = 0.8$)

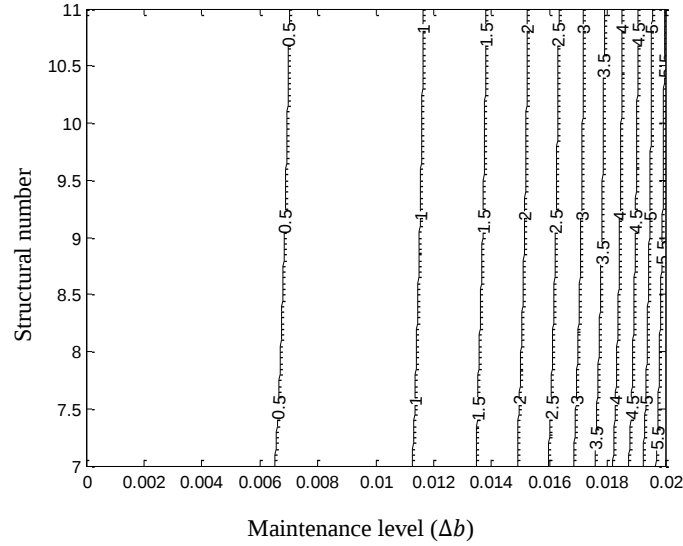


FIGURE 3.5(c) Maintenance costs ($\times 10^5 \$$) when maintenance is applied ($l = 0.8$)

Figure 3.3 shows that the structural number of 10.0 is the optimal pavement design when $l = 0.8$ and maintenance is not applied, and the minimum total discounted lifecycle costs are 6.39×10^5 dollars. Figure 3.4 exhibits the contour lines of total discounted lifecycle costs versus maintenance level and structural number when $l = 0.8$ and maintenance is applied. The optimal pavement design is 9.8, and the maintenance level, Δb , is 0.0074. In this case, the minimum discounted lifecycle costs are 6.31×10^5 dollars, so applying maintenance is optimal. In heavier traffic loading ($l = 0.6, 0.7$ and 0.8), the desirable pavement roughness after rehabilitation in a steady state reaches the best achievable state, s_0 . Figures 3.5(a), 3.5(b) and 3.5(c) are contour plots of user costs, rehabilitation costs and maintenance costs respectively. Figure 3.5(a) shows how maintenance level and pavement design influence user costs. More durably designed pavement,

which is managed higher level of routine maintenance, keeps its average roughness in lower level, so user costs decrease. Figure 3.5(b), and 3.5(c) reveal how rehabilitation and maintenance costs change with pavement design and maintenance level. Maintenance costs slightly decrease with structural number but these are not very sensitive. As traffic goes heavier, the agency should perform rehabilitations more frequently with lower overlay thicknesses and apply a higher level of maintenance to avoid significant user costs, because user costs proportionally increase with traffic loading.

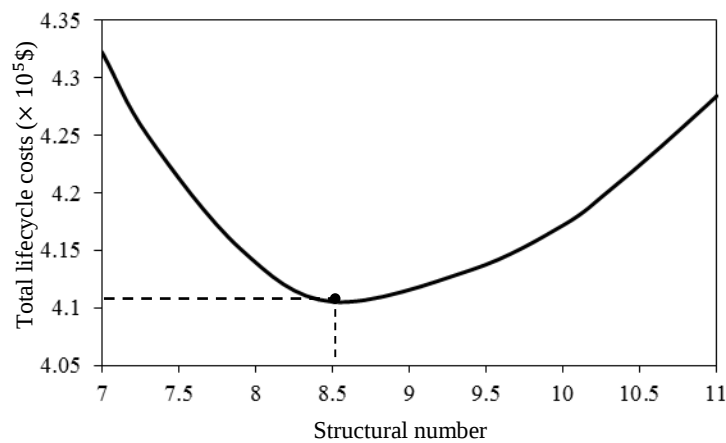


FIGURE 3.6 Total discounted lifecycle costs versus structural number when maintenance is not applied ($l = 0.4$)

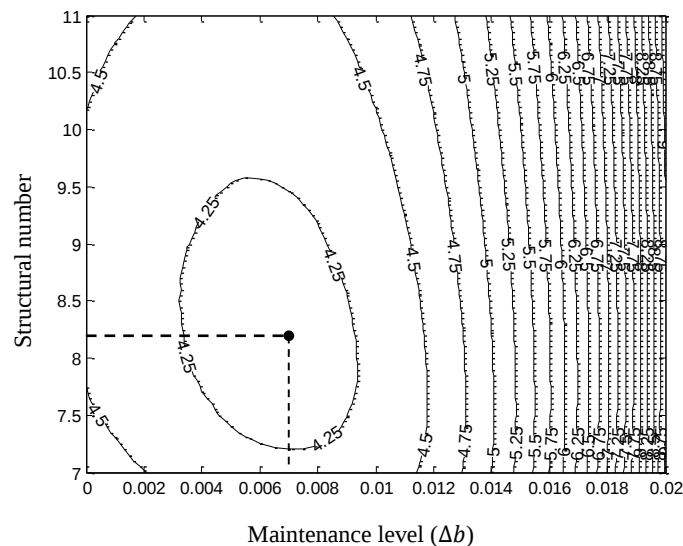


FIGURE 3.7 Total discounted lifecycle costs ($\times 10^5 \$$) when maintenance is applied ($l = 0.4$)

In light traffic ($l = 0.4$), Figure 3.6 and Figure 3.7 show the influence of the decision variables on total discounted lifecycle costs for both cases: maintenance is applied and not applied.

When maintenance is not applied, the minimum total discounted lifecycle costs are 4.11×10^5 dollars with structural number, $SN = 8.5$. With maintenance, the minimum total discounted lifecycle costs are 4.17×10^5 dollars with maintenance level, $\Delta b=0.0070$, and the lower structural number, $SN = 8.2$. In the case of light traffic, applying maintenance is suboptimal.

The pavement deterioration and improvement models used in this research are deterministic, but error-free estimation of pavement deterioration is impossible. Therefore, we need to examine the sensitivity of the optimal solution to particular parameters in the model in order to check the robustness of the solution approach that is built based on the assumption of deterministic behaviors. In this research, we assess the effects of uncertainty in two parameters q and a which are part of the function for structural number, $A(SN)$, on the objective value. Here, parameter q varies within the range of $[-5.5, -4.5]$, and parameter a varies within the range of $[580, 870]$. The results for the lighter traffic case ($l = 0.4$) are shown in Figure 3.8. The objective value varies with both parameters almost linearly, but the maintenance policy and the threshold structure of rehabilitation policy with the convergence properties are revealed to be still valid.

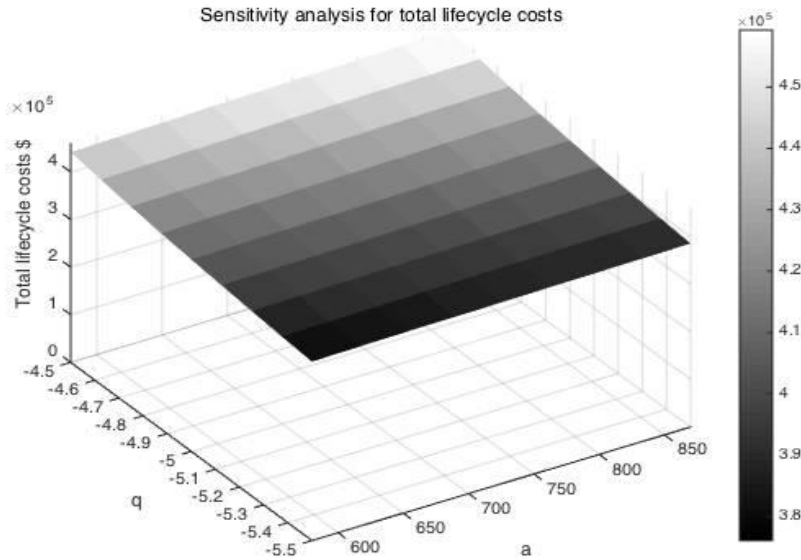


FIGURE 3.8 Sensitivity analysis on total lifecycle cost with changes in constant parameters a and q

3.4 Discussion

The solution methodology presented in this chapter is a promising approach to the joint optimization problem of pavement design and M&R strategies when pavement deterioration is history-dependent. The method uses the calculus of variations to find optimal rehabilitation policies for both cases: when the constraint of the best achievable pavement roughness after rehabilitation is not binding, and when it is binding. It is observed from the results of the parametric study that the optimal solution, under heavier traffic loading, generally requires a higher structural number for pavement design, higher levels of maintenance and a lower pavement-roughness trigger. Another finding is that rehabilitation planning converges to the steady state even if the

pavement deterioration model has history-dependent properties. In previous works, the ‘steady state in optimality’ was guaranteed by the use of a Markovian deterioration model. A non-Markovian deterioration model is more realistic, because it recognizes that deterioration is influenced by the history of previous rehabilitation plans. In practice, convergence to the steady state gives us a simple form for the optimal solution in the case of history-dependent deterioration model, similar to the optimal solution in the memory-less deterioration model case. We showed that the converging threshold structure of rehabilitation policy will still hold if uncertainty of deterioration model arises.

Chapter 4

Jointly Optimal Policies for Pavement Maintenance, Rehabilitation and Reconstruction

This chapter addresses the joint optimization of pavement MR&R activities in a continuous state space. We include frequent maintenance activities aimed at slowing down the deterioration process, rehabilitation activities aimed at reducing pavement roughness and reconstruction activities aimed at complete renewal of pavement condition and its age. The effects of different activities are assumed to be independent of each other. We develop a methodology to determine the optimal times and intensities to perform routine maintenance and rehabilitation and the optimal time to perform reconstruction to minimize the total discounted life time costs over an infinite time horizon. The deterioration model is history-dependent (i.e. the pavement deterioration rate depends on both current condition and the history of past management activities).

4.1. Optimization Problem Formulation

The optimization problem is formulated with general formulation (Section 4.1.1) and cost and performance models (Section 4.1.2).

4.1.1 General Formulation

The objective is to minimize the total discounted life time costs for the agency and the highway users. The relative weight between user cost and agency cost needs to be considered. This can be done by multiplying the agency costs by a weight factor. In the present research, we have assumed the weight factor to be equal to one so that our results can be directly compared to those in our most recent work (Gu et al 2012). This cost minimization problem is formulated as the following mathematical program:

$$\min \sum_{\tau=0}^{\infty} \left\{ \int_{\tau\delta^+}^{(\tau+1)\delta^-} \{C[S(u)] + C_M[S(u), x(\tau)]\} e^{-ru} du + \{M[x(\tau)] + RC[x(\tau)]\} e^{-r\tau\delta} \right\} \quad (4.1a)$$

Subject to

$$S(0) = S_0 (= \{s_0, h_0\}) \quad (4.1b)$$

$$s_{new} \leq s(t) \leq s_{max} \quad (4.1c)$$

$$0 \leq h(t) \leq h_{max} \quad (4.1d)$$

$$S(\tau\delta^+) = F_1[S(\tau\delta^-), x(\tau)] \quad (4.1e)$$

$$S(\tau\delta + u) = F_2[S(\tau\delta^+), u, x(\tau)], \forall u \in (0, \delta] \quad (4.1f)$$

where,

τ : discrete time index, $\tau = 0, 1, \dots$

δ : length of discrete time period

t^+, t^- : time moments after and before time t respectively

$s(t)$: surface condition of pavement at time t (e.g. roughness, or distress-based composite indices)

$h(t)$: age at time t : elapsed time from most recent reconstruction to time t

$S(t)$: augmented state of pavement at time t , including both current surface condition and age: $S(t) = \{s(t), h(t)\}$

$x(\tau)$: MR&R policy in the τ^{th} period: set of decision variables

$C[S(u)]$: user cost (vehicle operating cost) rate per unit time as a function of pavement condition

$C_M[S(u), x(\tau)]$: routine maintenance cost rate per time as a function of pavement condition

$M[x(\tau)]$: rehabilitation costs in the τ^{th} period

$RC[x(\tau)]$: reconstruction costs in the τ^{th} period

r : discount rate; we use the continuous discount factor e^{-rt}

s_{new} : surface condition after reconstruction

s_{max} : the worst allowable surface condition

h_{max} : maximum age

$F_1[S(\tau\delta^-), x(\tau)]$: effectiveness of rehabilitation and reconstruction on pavement surface condition, represented as surface condition after action is performed

$F_2[S(\tau\delta^+), u, x(\tau)]$: pavement deterioration function between two consecutive discrete time points

To account for the influence of the history of past management activities on pavement deterioration, the age of the pavement is considered as well as the current surface condition. Routine maintenance activities and rehabilitation are implemented on the surface layer, and thus cumulative damage in the underlying layers cannot be recovered except by reconstruction. Cumulative damage in the underlying layers accelerates the deterioration process, which means

that older pavements deteriorate faster even after rehabilitation. Considering age together with current surface condition provides a more accurate picture of the overall state of a pavement segment. In the objective function (4.1a), it is assumed that rehabilitation or reconstruction is performed at the starting point of the τ^{th} period, if either action is to be applied. On the other hand, routine maintenance, if selected, is continuously carried out along the time horizon. Constraint (4.1b) represents the initial pavement surface condition at the beginning of the planning horizon. Constraint (4.1c) describes the technical boundary of pavement surface condition. Constraint (4.1d) defines the range of age. Constraint (4.1e) sets effects of rehabilitation and reconstruction on pavement surface condition; it is a function of the before-action pavement condition and the action taken from three options; do nothing; perform rehabilitation; reconstruct. Constraint (4.1f) gives the continuous history-dependent pavement deterioration model, a function of current maintenance action and age. The cost functions for MR&R activities— $C_M[S(u), x(\tau)]$, $M[x(\tau)]$ and $RC[x(\tau)]$ — have zero value if the corresponding action is not taken.

4.1.2 Cost and Performance Models

The optimization problem presented in section 2 is very general, so various cost and performance models with different indicators can be used. Even though the formulation is based on deterministic models, it is easily modified to be applicable to stochastic models. We use cost and performance models adopted from the related literature as shown in this sub-chapter. The scope of pavement types in our research is flexible highway pavements.

In this research, we use roughness as an indicator of pavement surface condition, because user cost is directly related to pavement roughness. We adopt the pavement deterioration model developed by Paterson (1987), shown in (4.2). The term, $A \cdot l$, in (4.2) represents the aging effect on deterioration, which is not recovered by maintenance and rehabilitation activities applied on the surface layer. Hawas (2004) empirically shows that traveler route choices can be influenced by pavement condition. The change of pavement condition depends on traffic demand as shown in (4.2). This interdependency between traffic demand and pavement deterioration exists in reality, but we assume l to be exogenous and independent of pavement condition, because this chapter focuses on a single segment, and it is not possible to determine how l is affected by pavement condition without accounting for the network configuration. The maintenance levels are allowed to vary with time.

$$\frac{ds(t)}{dt} = b_c(t) \cdot s(t) + A \cdot l \cdot e^{b_c(t) \cdot h(t)} \quad (4.2)$$

where,

A : constant inversely related to the pavement structural number ($= a(SN + 1)^q$ in equation (3.2))

l : the annual traffic loading

$b_c(t)$: deterioration rate at time t , determined by maintenance level at time t

In the absence of rehabilitation or reconstruction, pavement roughness increases between two consecutive discrete time points; it is a function of $S(\tau\delta^+)$ and $b(\tau)$. We assume that $b(\tau)$ is fixed in the τ^{th} period.

$$\begin{aligned} S(\tau\delta + u) &= F_2[S(\tau\delta^+), u, x(\tau)] \\ &= \{s(\tau\delta^+)e^{b(\tau)u} + A \cdot l \cdot u \cdot e^{b(\tau)(h(\tau\delta^+)+u)}, u + h(\tau\delta^+)\}, \forall u \in (0, \delta] \end{aligned} \quad (4.3)$$

where,

$b(\tau)$: deterioration rate in the τ^{th} period, determined by maintenance level in the τ^{th} period

User vehicle operating costs are given as (4.4). The vehicle operating cost rate per unit time is a linear function of pavement roughness, traffic volume (or truck traffic volume), v , and weight factor, $W_{u/a}$.

$$C[s(t)] = W_{u/a} \cdot v \cdot (c_1 \cdot s(t) + c_2) \quad (4.4)$$

X_1, X_2 and X_3 are binary decision variables of MR&R activities respectively, and these are elements of $x(\tau)$. For example, if maintenance is applied in the τ^{th} period, $X_1 = 1$. Otherwise, $X_1 = 0$. The maintenance cost per unit time is given as (4.5). Here, $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1$ and γ_2 are positive constant parameters. The maintenance cost per unit time is positively related to Δb , and it is a linear function of pavement roughness. Thus the maintenance intensity is denoted as the reduction in the deterioration rate, $\Delta b(\tau) = b_0 - b(\tau)$, where b_0 is defined as the original deterioration rate without any routine maintenance activity. The range of $b(\tau)$ is defined by $b_{min} \leq b(\tau) \leq b_0$.

$$C_M[s(t), x(\tau)] = X_1 \cdot (C_{M,1} \cdot s(t) + C_{M,2}), \forall t \in (\tau\delta, (\tau + 1)\delta], \quad (4.5)$$

where,

$$C_{M,1} = \alpha_1 e^{\beta_1 \Delta b(\tau)} + \gamma_1 \text{ and } C_{M,2} = \alpha_2 e^{\beta_2 \Delta b(\tau)} + \gamma_2$$

The rehabilitation cost and effectiveness models are shown in (4.6) and (4.7) respectively. μ_1, μ_2 and μ_3 are positive parameters. The original pavement roughness $s(\tau\delta^-)$ is reduced by $G(w(\tau), s(\tau\delta^-))$ a function of $s(\tau\delta^-)$ and the intensity of rehabilitation, w_i . Generally, any rehabilitation is suboptimal if the intensity is larger than the upper limit, $R(s(\tau\delta^-))$, since it does not produce additional improvement in pavement roughness. We modify the original effectiveness model of rehabilitation of Ouyang and Madanat (2004) to (4.7) by introducing the concept of the best achievable roughness level, s_0 , which defines the technical limitation of the rehabilitation activity (i.e. the best achievable roughness after rehabilitation). As shown from Paterson (1990), the rehabilitation effect, which is a function of overlay thickness and the roughness before action, reaches zero as the roughness before action decreases, and thus the roughness after rehabilitation cannot be smaller than the best achievable roughness level. The user travel time delay costs caused

by partial road closure, $C_2[v]$, are assumed to be zero by the same reason presented in Chapter 3. The range of rehabilitation intensity is represented as (4.8).

$$M[x(\tau)] = X_2 \cdot [m_1 \cdot w(\tau) + m_2 + w_{u/a} \cdot C_2[v]] \quad (4.6)$$

$$G(w(\tau), s(\tau\delta^-)) = \max\left(0, \min\left(s(\tau\delta^-) - s_0, \frac{\mu_1 \cdot w(\tau)}{\mu_2 \cdot s(\tau\delta^-) + \mu_3} \cdot s(\tau\delta^-)\right)\right) \quad (4.7)$$

$$\begin{aligned} 0 \leq w(\tau) &\leq R(s(\tau\delta^-)) \\ &= \frac{\mu_2 + \mu_3/s(\tau\delta^-)}{\mu_1} \\ &\cdot (s(\tau\delta^-) - \min(s(\tau\delta^-), \max(s_0, (1 - \mu_1) \cdot s(\tau\delta^-)))) \end{aligned} \quad (4.8)$$

The reconstruction cost per unit length of a highway pavement segment is a function of the structural number. In this chapter, pavement design is not a decision variable, so the reconstruction cost function is assumed to be independent of pavement design including materials, structural number and etc., expressed as (4.9). The user cost due to complete road closure, $C_3[v]$, is added, which is caused by 100% loss in capacity during reconstruction. We simply assume that the detour cost, $C_3[v]$, is proportional to v , because it is not possible to predict the redistribution of traffic demands without considering the network configuration. Practitioners may use alternative $C_3[v]$ for different traffic environments. After reconstruction, the pavement condition is improved to a new one, i.e., $S(\tau\delta^+) = \{s_{new}, 0\}$.

$$RC[x(\tau)] = X_3 \cdot (k_1 + W_{u/a} \cdot C_3[v]) \quad (4.9)$$

where,

$$C_3[v] = k_2 \cdot v$$

4.2 Solution Approach

We enumerate six possible MR&R policies in Table 4.1, which are indexed as $I(x) = 1, 2, \dots, 6$ (refer to the first column). Policy, $x(\tau)$, is a set of binary decision variables of MR&R activities and intensities of both routine maintenance and rehabilitation (shown in the third column). As defined previously, X_1, X_2 and X_3 represent whether to perform routine maintenance, rehabilitation and reconstruction respectively. Even if we have 6 discrete policies as listed in Table 1, the decision variables, Δb and w , are still continuous. We reduce the dimensionality of the decision variables, because high dimensionality is computationally expensive. Using the maximum level of w , $R(s(\tau\delta^-))$, is proved to be optimal in Appendix C, and it has been shown to be optimal for both the infinite-horizon and finite-horizon problems in related literature (Li and Madanat, 2002; Ouyang and Madanat, 2006; Ouyang 2007). The effect of each policy on pavement condition is shown in the last two columns. Only reconstruction resets the age of segment to zero, as shown in the last column. If neither rehabilitation nor reconstruction is performed, $S(\tau\delta^+) = S(\tau\delta^-)$. We exclude unrealistic policies, such as: applying multiple rehabilitations or both

rehabilitation and reconstruction at the same time. Therefore, for each period, at most one rehabilitation or reconstruction can be selected. The last option is whether to perform routine maintenance, or not.

$I(x)$	Description	$x(\tau) = \{X_1, X_2, X_3, \Delta b(\tau), w(\tau)\}$	$S(\tau\delta^+) = F_1[S(\tau\delta^-), x(\tau)]$	
			$s(\tau\delta^+)$	$h(\tau\delta^+)$
0	Do nothing	$\{0,0,0,0,0\}$	$s(\tau\delta^-)$	$h(\tau\delta^-)$
1	Rehabilitation only	$\{0,1,0,0, R(s(\tau\delta^-))\}$	$s(\tau\delta^-) - G(R(s(\tau\delta^-)), s(\tau\delta^-))$	$h(\tau\delta^-)$
2	Reconstruction only	$\{0,0,1,0,0\}$	s_{new}	0
3	Maintenance only	$\{1,0,0, \Delta b(\tau), 0\}$	$s(\tau\delta^-)$	$h(\tau\delta^-)$
4	Rehabilitation and maintenance	$\{1,1,0, \Delta b(\tau), R(s(\tau\delta^-))\}$	$s(\tau\delta^-) - G(R(s(\tau\delta^-)), s(\tau\delta^-))$	$h(\tau\delta^-)$
5	Reconstruction and maintenance	$\{1,0,1, \Delta b(\tau), 0\}$	s_{new}	0

TABLE 4.1 Combinations of activities and their effects on pavement condition

The problem formulated in (4.1a) to (4.1f) is reformulated in (4.10a) to (4.10f) in terms of Dynamic Programming. It is a deterministic discounted infinite-horizon optimization problem with continuous pavement condition and discrete time. In this formulation, $x(\tau)$ is re-defined to be $x(S)$ which is a function of S instead of τ .

$$J^*(S) = \min_{x(S)} \{g(S, x) + e^{-r\delta} \cdot J^*(S')\}, \forall S \quad (4.10a)$$

Subject to

$$g(S, x) = \int_{0^+}^{\delta} \{C[F_2[F_1[S, x], u, x]] + C_M[F_2[F_1[S, x], u, x], x]\} e^{-ru} du + M[x] + RC[x] \quad (4.10b)$$

$$S(0) = S_0 (= \{s_0, h_0\}) \quad (4.10c)$$

$$s_{new} \leq s \leq s_{max} \quad (4.10d)$$

$$0 \leq h \leq h_{max}(s) \quad (4.10e)$$

$$S' = \{F_2[F_1[S, a], \delta, a], h + \delta\} \quad (4.10f)$$

where,

$J^*(S)$: optimal discounted lifetime costs

The dynamic program (4.10) can be solved by various numerical methods. In this research, we use a parametric approximation method in a similar way to related papers in the pavement management area (Ouyang 2007, Medury and Madanat 2013). We approximate $J^*(S)$ as $V_{\theta}(S)$, which is a linear combination of a relatively small number of K differentiable basis functions: $\{\varphi^1(s, h), \dots, \varphi^K(s, h)\}$. For example, ordinary or Chebyshev polynomials or poly-log can be basis functions. The time horizon is already discretized by the period, δ , so age, h , is also discretized by δ . The properties of continuity of h are considered in the function, $g(S, x)$ as shown in (4.10b). To simplify the methodology, we incorporate h into the basis functions by treating h as a continuous variable, which we discretize by δ again for sampling.

$$J^*(S) \approx V_{\theta}(S) = \sum_{k=1}^K \theta^k \varphi^k(s, h), \text{ where } \boldsymbol{\theta} = \{\theta^1, \dots, \theta^K\} \quad (4.11)$$

If true $J^*(S)$ is smooth and not too irregular, this approximation will be a suitable approach for appropriately selected value of K . For state-space sampling, we discretize the continuous state of pavement surface condition, s , into Z points, satisfying $Z > K$. Small K and Z do not guarantee convergence to $J^*(S)$ (Benitez-Silva et al. 2000), but the algorithm with properly set values of K and Z converges to the true optimal objective costs. The parameter vector $\boldsymbol{\theta}$ is initialized as a specific vector $\boldsymbol{\theta}_0$. Within the i^{th} iteration, policy $x_i(S)$ is found by the following:

$$x_i(S) = \underset{x \in X(S)}{\operatorname{argmin}} g(S, x) + e^{-r\delta} \sum_{k=1}^K \theta_{i-1}^k \varphi_{i-1}^k(s, h), \forall i \quad (4.12)$$

The decision variable $x \in X(S)$, where $X(S)$ is the set of all possible policies selected from Table 2 according to S . $X(S)$ varies with S to satisfy the constraints. If $s' > s_{\max}$ for a specific S , $X(S)$ does not contain policies $I(x) = 0$ and 3 to satisfy the constraint (4.10d). The intensity of rehabilitation for policies $I(x) = 1$ and 4 is found analytically. The intensity of maintenance for policies $I(x) = 3, 4$ and 5 can be found by numerical methods such as the gradient method because both $g(S, x)$ and $V_{\theta_i}(S)$ are differentiable by Δb as shown in (4.13).

$$\begin{aligned}
& \frac{\partial g(S, x)}{\partial \Delta b} + e^{-r\delta} \cdot \frac{\partial V_{\theta_i}(S)}{\partial (\Delta b)} \\
& = \int_{0^+}^{\delta} \left\{ c_1 \cdot \frac{\partial F_2[s^+, u, x]}{\partial \Delta b} + X_1 \right. \\
& \quad \cdot \left[C_{M,1} \cdot \frac{\partial F_2[s^+, u, x]}{\partial \Delta b} + \frac{\partial C_{M,1}}{\partial \Delta b} \cdot F_2[s^+, u, x] + \frac{\partial C_{M,2}}{\partial \Delta b} \right] \Big\} e^{-ru} du \\
& \quad + e^{(b-r)\delta} \cdot \frac{\partial J_{V_{\theta_i}}(S)}{\partial s'} \cdot \frac{\partial F_2[s^+, \delta, x]}{\partial (\Delta b)}, \text{ if } I(x) = 3, 4 \text{ or } 5
\end{aligned} \tag{4.13}$$

After selecting $x_i(S)$ from (4.12), we linearly regress $g(S, x_i(S))$ on $\{\varphi^k(s) - e^{-r\delta} \cdot \varphi^k(s', h')\}$ to update a parameter vector θ_i . The proposed algorithm repeatedly updates at all sampling points until both parameter vectors and policies converge.

4.3 Parametric Examples

In this section, we demonstrate the solution methodologies from two parametric examples: a case with the memory-less deterioration model and a general case with the history-dependent deterioration model. In the parametric approximation method, we use the same basic functions of ordinary polynomials, and the same total number of parameters $K = 21$ used in Ouyang (2007). We discretize the roughness state by 1 *IRI* and age by 0.5 year, so $Z \gg K$. The algorithms converges to the close-to-optimal solution within at most 300 iterations. Common parameter values used in both cases for a single pavement 1 kilometer, for a one lane highway are shown in Table 4.2.

Parameter	Value	Units	Parameter	Value	Units
c_1	5.91	\$/IRI/AADTT	α_1	260	\$/IRI /km /lane/yr
c_2	0	\$/AADTT	γ_1	1300	\$/IRI /km /lane/yr
m_1	11,000	\$/mm /km/lane	α_2	167	\$/IRI /km /lane/yr
m_2	150,000	\$/km/lane	γ_2	833	\$/IRI /km /lane/yr
μ_1	0.66	-	k_1	900,000	\$/km/lane
μ_2	7.15	mm/IRI	k_2	250	\$ · yr/km /AADTT
μ_3	18.3	mm	δ	0.5	yr
r	0.07	-	$W_{u/a}$	1	-

:- dimensionless; *ESAL*: Equivalent Single Axle Load; AADTT: Annual Average Daily Truck Traffic

TABLE 4.2 Common parameters used

4.3.1 A Special Case : Memoryless Model

To allow for comparison to a previous study, we use same cost, performance functions and parameters used in Gu et al. (2012). They jointly optimize maintenance and rehabilitation to solve the optimization problem of pavement management and assume that the maintenance level, one of decision variables, is fixed over the planning horizon after it is selected as optimal. The proposed solution methodologies allow the maintenance level to vary, so our methodologies should provide a better solution than the methodology of Gu et al. (2012), in that the constraint of the fixed maintenance level is released. Moreover, their methodology includes multiple rehabilitation activities at the same time when the pavement condition is very poor, which is unrealistic. Our solution proposes that performing reconstruction is optimal in this case. This result is similar to the optimal policies proposed by Rashid and Tsunokawa (2012).

We present an example, when traffic loading $l = 0.6$ million *ESALs/yr/lane* and truck traffic volume $v = 2200$ AADTT. In the memory-less deterioration model, $A = 0$, and b_0 has a larger value than the memory-less model. The parameter values are only used for this special case study of the memory-less model shown in Table 4.3. s_{new} and s_0 are set as relatively small in comparison to realistic values because Gu et al. does not include constraints about these values.

h_{max} is assumed to be an arbitrary large number, because aging is not considered in Gu et al. (2012).

Parameter	Value	Units
b_0	0.08	-
b_{min}	0.035	-
β_1	120	-
β_2	120	-
s_{new}	0.75	<i>IRI</i>
s_0	0.8	<i>IRI</i>
s_{max}	12.0	<i>IRI</i>
A	0	<i>IRI · lane/million ESAL</i>
h_{max}	$\gg 0$	<i>yr</i>

TABLE 4.3 Parameters used in the memory-less example

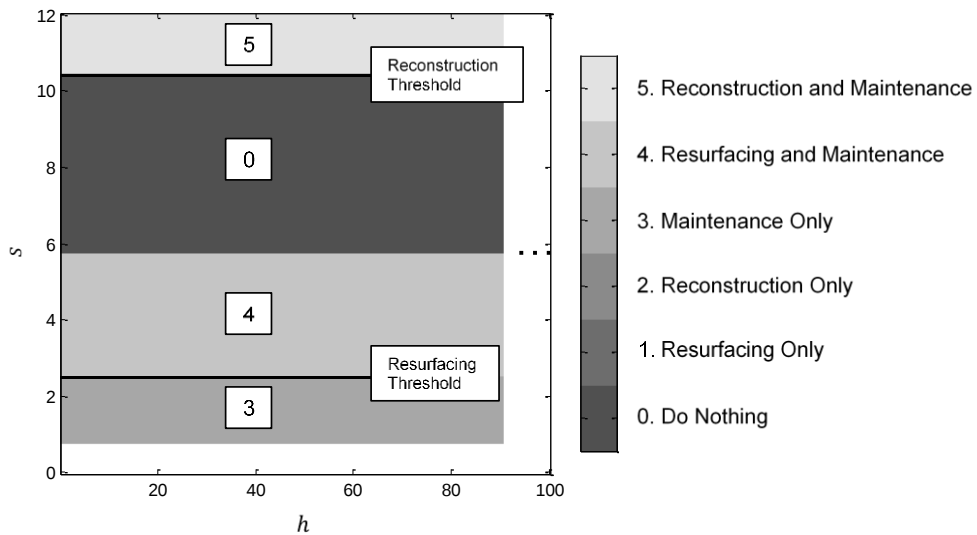


FIGURE 4.1 Optimal policies at convergence on the $h - s$ plane

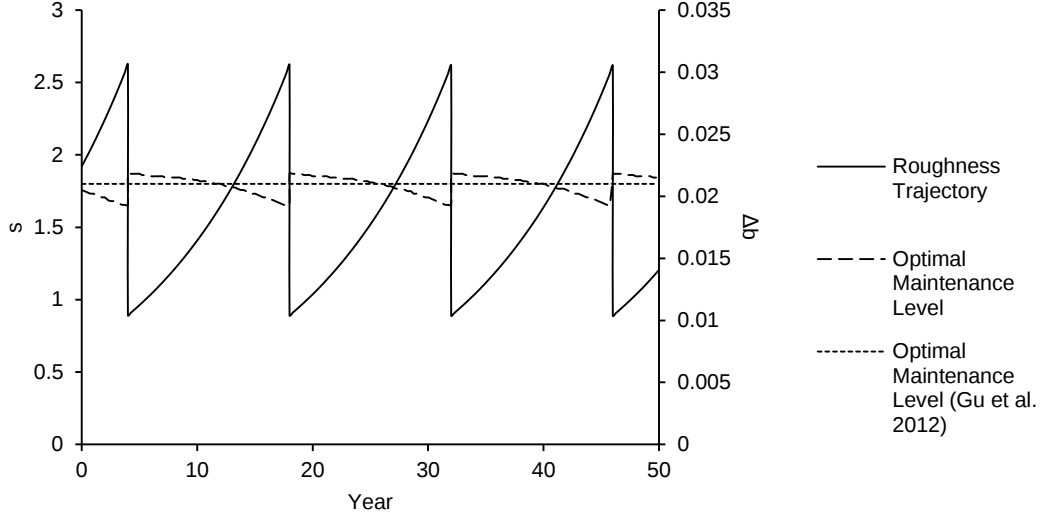


FIGURE 4.2 Optimal trajectory of roughness and the optimal maintenance policies

Figure 4.1 shows the optimal policies at convergence on the $h - s$ plane obtained from the proposed solution methodologies. It is observed that both optimal policies of rehabilitation and reconstruction have a threshold roughness structure, revealed by previous studies (Li and Madanat 2002, Ouyang and Madanat 2006). Also, Figure 1 shows that those thresholds are independent of age, because the deterioration model is memory-less. There is a horizontal threshold of reconstruction located near $s = 10.47$ above the rehabilitation threshold line. When a pavement condition reaches this threshold line, it is optimal to reconstruct. From the mapping of optimal MR&R policies shown in Figure 4.1, it is impossible to find the optimal intervals between reconstructions (i.e. the optimal lifecycle of the pavement), because the deterioration model does not account for aging on ($A = 0$). Figure 4.2 displays the optimal trajectory of pavement roughness and compares the optimal maintenance levels obtained by our approach to those of Gu et al. (2012) for $s_0 = 1.923$ IRI ($= 25QI$). It shows that applying levels of maintenance that vary with current pavement roughness is more beneficial ($J^* = 9.16 \times 10^5 \$$) than using a fixed maintenance level ($J^* = 9.36 \times 10^5 \$$). The maintenance costs are positively related to the current pavement roughness, so a lower level of maintenance is optimal when the deterioration rate increases.

Suppose the initial pavement roughness is 8.0 IRI. According to Gu et al., performing rehabilitation activities twice at the initial time point is optimal ($J^* = 2.09 \times 10^6 \$$), but, based on our methodology, it is optimal to reconstruct this pavement segment 3.5 years later ($J^* = 1.90 \times 10^6 \$$), when the state reaches the reconstruction threshold. As shown in Figure 4.1, the absence of maintenance is optimal during the first 3.5 years, because excessive maintenance costs are caused by high roughness. Essentially, the optimization problem of Gu et al. is equivalent to the optimization problem presented in this research with two additional constraints: (i) the maintenance level is constant, and (ii) the reconstruction cost is infinite (assumed to be a very large value). These additional constraints yield suboptimal results.

4.3.2 General Case: History-Dependent Model

History dependent deterioration models are more realistic than memory-less deterioration models. In this parametric study, this deterioration model is defined by $A > 0$. The parameter values only used for this parametric example of the history-dependent model are shown in Table 4.4. We consider two cases with different traffic loadings, $l = 1.2$ and 0.5 million *ESALs/yr/lane*, with different traffic volumes $v = 4400$ and 1830 *AADTT*.

Parameter	Value	Units
b_0	0.04	-
b_{min}	0.025	-
β_1	240	-
β_2	240	-
s_{new}	1.0	<i>IRI</i>
s_0	1.2	<i>IRI</i>
s_{max}	15	<i>IRI</i>
A	$2.54E-05^1$	<i>IRI \cdot lane/AADTT</i>
h_{max}	100	<i>yr</i>

¹: calculated based on Paterson (1987) with structural number $SN = 5$

TABLE 4.4 Parameters used in the general history-dependent example

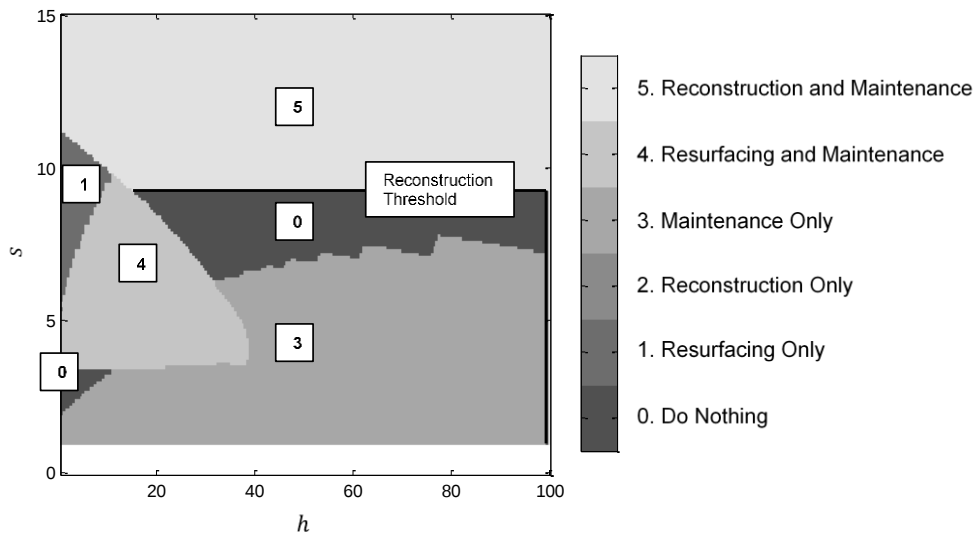


FIGURE 4.3(a) Optimal policies at convergence on the $h-s$ plane ($l = 1.2, v = 4400$)

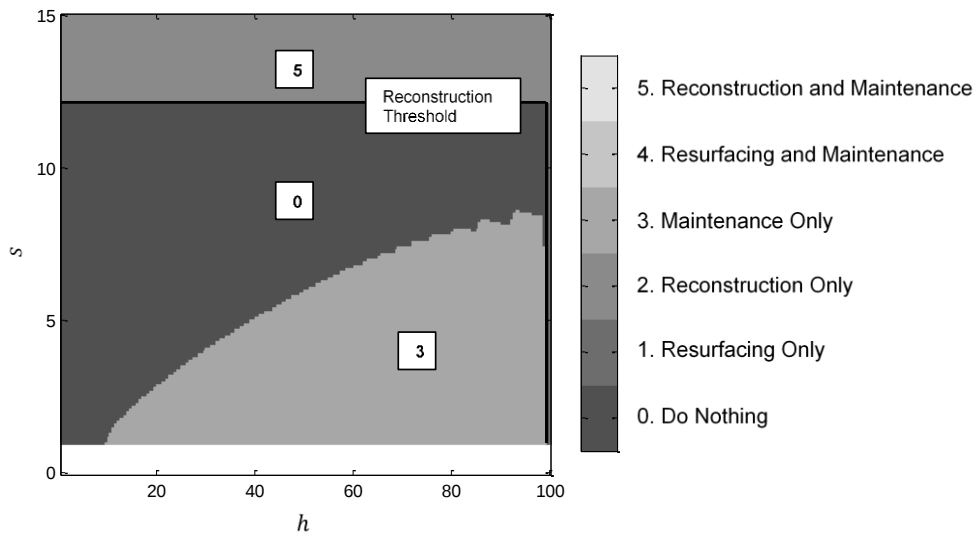


FIGURE 4.3(b) Optimal policies at convergence on the $h-s$ plane ($l = 0.5, v = 1830$)

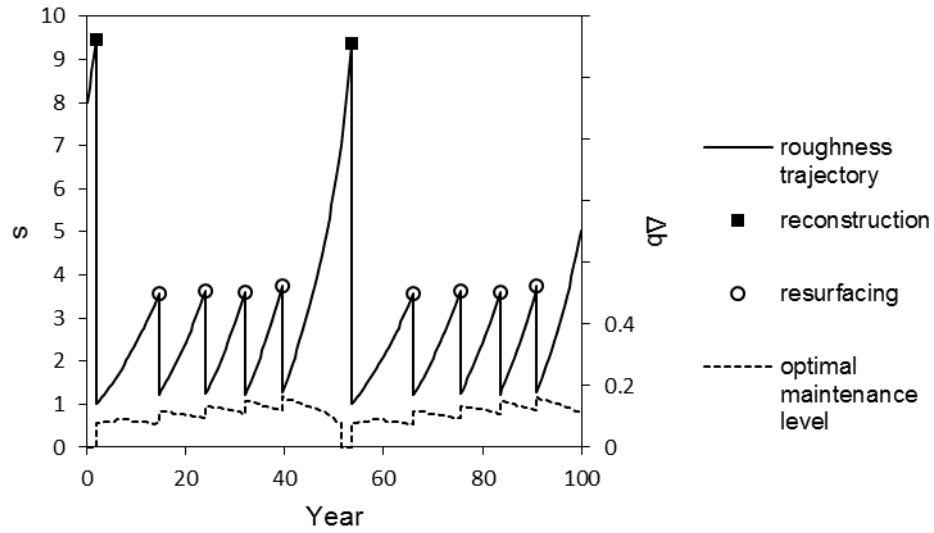


FIGURE 4.4(a) Optimal trajectory and optimal maintenance policies for a pavement segment starting from $S_0 = \{8.0 \text{ IRI}, 20 \text{ yr}\}$ ($l = 1.2, v = 4400$)

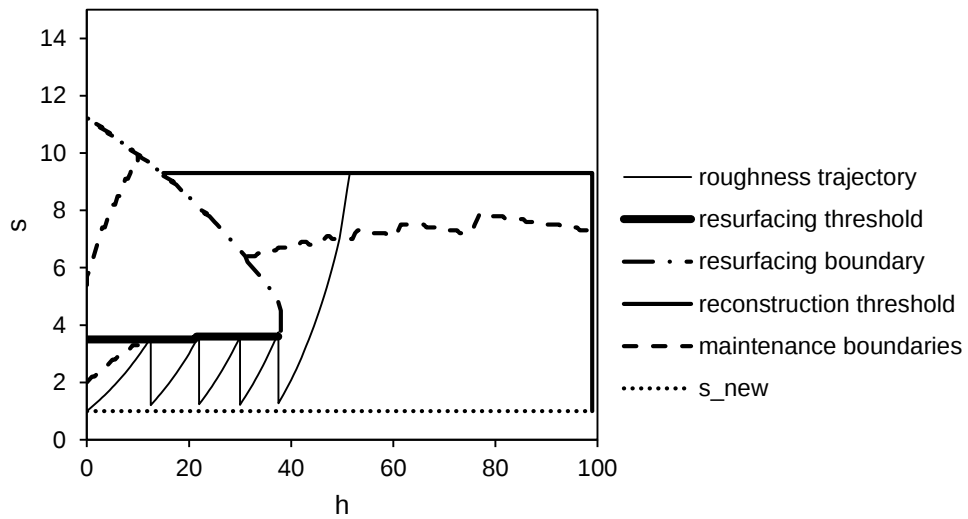


FIGURE 4.4(b) Optimal trajectory in the steady state superposed on the map ($l = 1.2, v = 4400$)

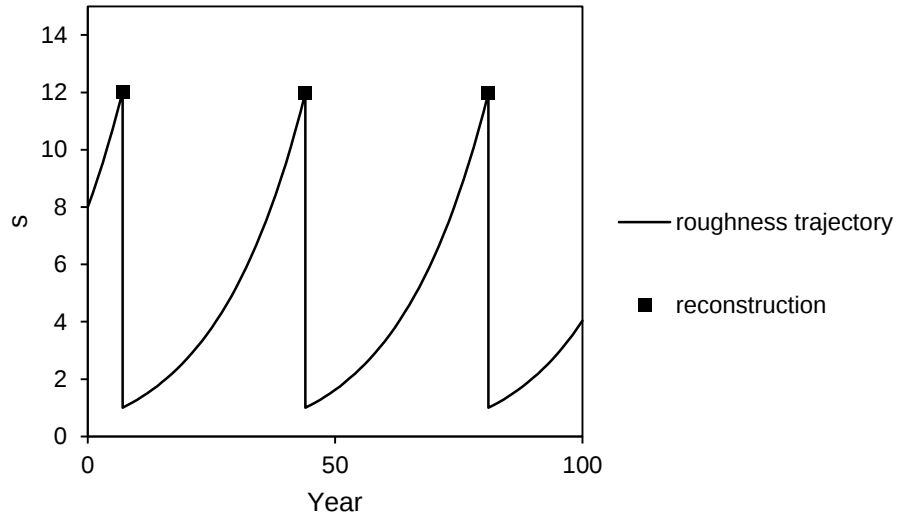


FIGURE 4.5(a) Optimal trajectory and optimal maintenance policies for a pavement segment starting from $S_0 = \{8.0 \text{ IRI}, 20 \text{ yr}\}$ ($l = 0.5, v = 1830$)

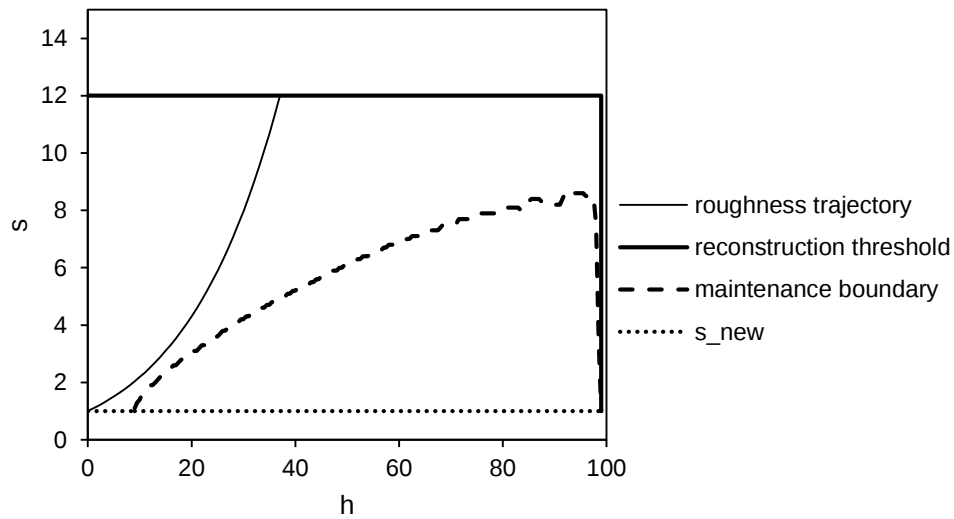


FIGURE 4.5(b) Optimal trajectory in the steady state superposed on the map ($l = 0.5, v = 1830$)

Figure 4.3(a) and 4.3(b) show the optimal policies of two cases with different traffic loadings at convergence. Figure 4.4(a) and 4.5(a) show the roughness trajectories under the optimal MR&R policies with different traffic loadings and volumes, where the trajectories start from the same pavement condition, $S_0 = \{8 \text{ IRI}, 20 \text{ yr}\}$ ($J^* = 3.39 \times 10^6$ for $\{l = 1.2, v = 4400\}$, and $J^* = 2.01 \times 10^6$ for $\{l = 0.5, v = 1830\}$). In a relatively heavier traffic situation, frequent rehabilitation activities between reconstructions are shown to be optimal. It can be seen that the optimal maintenance level decreases with the deterioration rate in the same rehabilitation cycle, and it increases with age for a given roughness level (refer to the dashed line in Figure

4.4(a)). Moreover, applying routine maintenance is suboptimal before trajectories reach the reconstruction thresholds, where performing reconstruction is optimal. This is graphically illustrated in Figure 4.4(b): the trajectory crosses the dashed line which is the boundary between policies 0 and 3. As shown in Figure 4.3(a), the rehabilitation threshold is not fixed (i.e. not a horizontal line), but increases with h . Every reconstruction produces the identical pavement condition, $\{s_{\text{new}}, 0 \text{ yr}\}$, so the steady state is guaranteed for all optimal reconstruction activities. In a lighter traffic case, figure 4.5(a) shows that performing rehabilitations is suboptimal in the steady state, and applying periodic reconstructions is more economical. This is because allowing a high roughness level does not yield significant user costs in lighter traffic, which are positively related to traffic loading as shown in (4.4), according to the assumption that travel time delays due to poor pavement conditions are ignored. Moreover, lower traffic loading produces slower deterioration, so it is unnecessary to apply frequent actions. Lastly, construction is relatively cheaper in lighter traffic (see (4.9)). In Figure 4.4(a) and 4.5(a), it is observed that performing reconstruction as the first activity is optimal for both situations, and heavier traffic situation requires reconstruction as the earliest action. In the lighter traffic situation, it is shown that the condition trajectory reaches the steady-state when the first reconstruction is performed, and MR&R policies including only reconstructions are optimal in the steady state as shown in Figure 4.5(b). The trigger roughness of reconstruction is 12 *IRI*; it decreases as the ratio of the weights of user costs to agency costs increases. Optimal lifecycles for the two cases are 51.5 years and 37 years respectively, as can be seen in Figures 4.4(b) and 4.5(b).

4.3.3 Sensitivity Study to Uncertainty in the Deterioration Process

Fifty simulations are conducted respectively for 13 cases of traffic loading. In every simulation, $S_0 = \{2 \text{ IRI}, 5 \text{ yr}\}$. Once an optimal MR&R strategy is found from the deterministic model (e.g. see Figure 3), the stochastic roughness trajectory is drawn from the values a , q and b , which are randomly generated in every time period from the pre-defined distributions shown in table 4.5, and it obeys the optimal policy yielded from the deterministic model. The objective value is calculated from the trajectory for each simulation. The simulation results are graphically represented in Figure 4.6. The bold line in Figure 4.6 refers to the object value that resulted from the deterministic deterioration with respect to traffic loading. Small circles stand for the results of simulations, and the mean of the objective values of fifty simulations, for each given traffic loading, is represented by the dashed line. As can be seen, the dashed line is very close to the bolded line. Also, the sample standard deviation of simulation results for given traffic loading is shown to be low compared to the mean value. The graph has clearly indicated that both lines are positively correlated with the change in ESALs. These results mean that we can select optimal MR&R policy based on deterministic deterioration models, because it is robust to the uncertainty in the models. This is similar to the conclusions concerning uncertainty in Sathaye and Madanat (2011, 2012) and Li and Madanat (2002).

Parameter	Distribution
a	$TN(725, 94400, 0, \infty)$
q	$TN(-4.99, 0.064, -\infty, 0)$
b	$TN(0.0153, 7.3 \times 10^7, 0, \infty)$

TABLE 4.5 Deterioration model parameters with their distributions

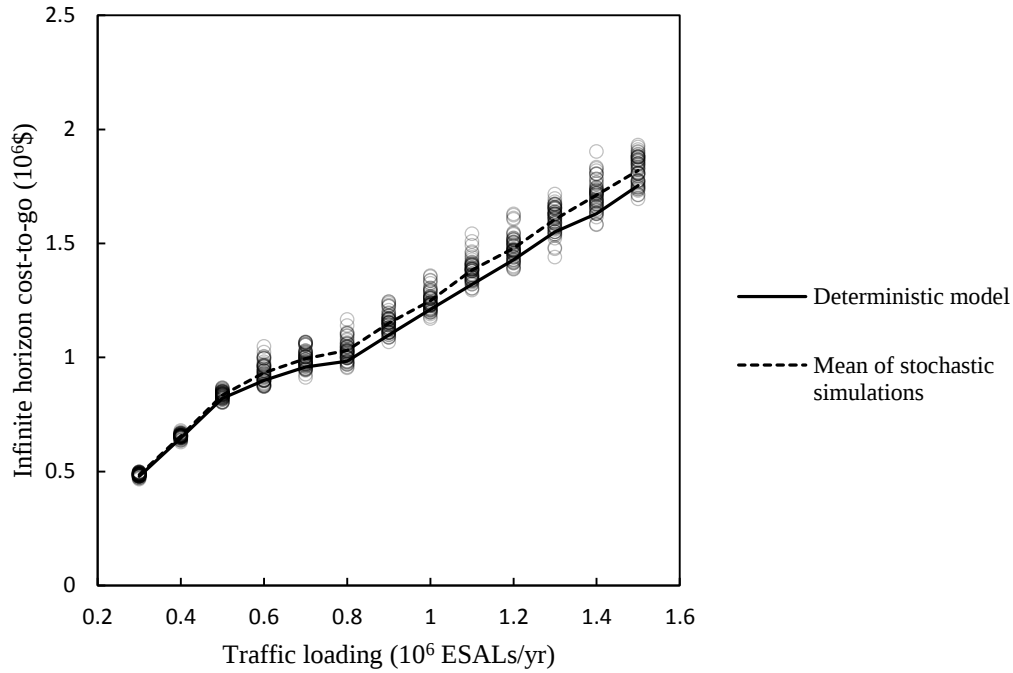


FIGURE 4.6 Sensitivity analysis on total lifecycle cost with changes in traffic loading

4.4 Discussion

This chapter describes a methodology for the joint optimization problem of MR&R activities for flexible highway pavements. The objective function is to minimize the total discounted life time costs over an infinite time horizon, consisting of agency investments for MR&R activities, as well as for user costs. Our approach optimizes reconstruction intervals jointly with maintenance and rehabilitation policies. We enumerate a possible set of combinations of different management treatments and clarify the effect of each combination on pavement condition. A numerical scheme to solve the dynamic programming is presented, which is parametric approximation method. This methodology is shown to be effective through parametric examples. The numerical results show that: a pavement-condition-dependent maintenance level is optimal; management policies without applying rehabilitation can be optimal in the case of light traffic loading; the optimal lifecycle of pavement segment can be obtained through the joint optimization of maintenance and

reconstruction; and the optimal results are robust to the uncertainty in the deterioration-related parameters.

The mathematical program and solution methodologies are easily applicable to various cost and performance models. The discrete approximation methodology can be used with deterministic continuous deterioration models, or stochastic discrete models. The parametric approximation methodology can be used to solve high-dimensional problems.

In this chapter, the proposed solution methodology is for a segment-level problem. The next step is to extend the single segment problem to the system level problem under budget constraints, with necessary modifications to ensure its general applicability.

Chapter 5

A Joint Bottom-Up Solution Methodology for System-Level Pavement Rehabilitation and Reconstruction

The subject of this chapter is the problem of optimizing rehabilitation and reconstruction policies for large-scale pavement systems. We address situations where rehabilitation and reconstruction projects are funded from separate budgets as well as those where they share the same budget. For example, there can be two independent budgets: a capital budget for construction and reconstruction projects, and a maintenance budget for maintenance and rehabilitation activities.

We develop a bottom-up solution methodology for the system-level optimization. A bottom-up approach reflects pavement segment-specific characteristics, and provides individual optimal strategies for each segment. The performance and deterioration models are deterministic and follow Markovian properties, but consider history-dependent deterioration process. This is achieved by using augmented condition states that include history variables, such as the age of the pavement or the cumulative traffic loading, in addition to the current pavement condition, represented by pavement roughness.

5.1 Optimization Problem Formulation

The optimization problem is formulated with general formulation (Section 5.1.1) and cost and performance models (Section 5.1.2).

5.1.1 General Formulation

The objective of the problem is to minimize the total discounted costs, subject to budget constraints. Consider a system of pavement which includes N segments numbered as $n \in \{1, \dots, N\}$. The discounted cost-to-go of segment n at time t , $f_n(S_n(t), x_n)$, is defined as a function of the augmented condition state, $S_n(t)$, and segment-level MR&R actions, x_n . The multi-dimensional condition state $S_n(t)$ contains the pavement condition state and a history variable such as age. The system-level pavement management system (PMS) problem is formulated in (5.1). We focus on the infinite horizon problem and find the optimal steady state strategies. Thus, the objective function (1a) is not influenced by the current condition of the segments. A set of decision variables, x , is defined as $\{x_1, \dots, x_N\}$, and a set of all possible x is noted as X . Two problems are separately defined by different budget constraints, (5.1b) and (5.1c). Monetary budget constraints for different types of activities (construction, rehabilitation, maintenance) are separately represented in (5.1b) and (5.1c) and indexed by j . The function $g_n^j(x_n)_\tau$ is the cost corresponding to budget j for segment n in the τ^{th} period when x_n is applied. For example, if B_1 is the available capital budget per unit period, then $g_n^1(x_n)_\tau$ is the reconstruction cost of the n^{th} segment in the τ^{th} budget expenditure period. The unit budget expenditure period, denoted by Γ , is the number of years that funds allocated in a time period τ can be spent before the unspent moneys are returned to the governing agency. For instance, if a budget is provided to manage pavement systems, and it

can be used in next five years, then Γ is five years. Constraint (5.1c) describes the case where all activities are funded from a single budget, B .

$$\min_{x \in X} \sum_n f_n(x_n) \quad (5.1a)$$

Subject to

$$\text{Problem 1. separate budgets: } \sum_n g_n^j(x_n)_\tau \leq B_j, \forall j, \tau \quad (5.1b)$$

$$\text{Problem 2. combined budget: } \sum_n \sum_j g_n^j(x_n)_\tau \leq B, \forall \tau \quad (5.1c)$$

In the steady state, the optimal trajectory of $S_n(t)$, which is controlled by x_n , consists of repeated cycles, where each cycle starts with a reconstruction and includes any number of rehabilitation events. Therefore, we approximate $g_n^j(x_n)_\tau$ by a constant $g_n^j(x_n)$, which is a time-average cost, for every period τ in the steady state period as shown in (5.2). This approximation is only valid at the system-level scale, and will not create large errors in the optimization, as long as the number of segments N is large enough and the budget expenditure period, Γ , is not too short. For each segment, investments in reconstruction and rehabilitation activities are made instantly and periodically rather than continuously on the time horizon. If a system is comprised of a few segments, and the budget expenditure period is too short, system-level agency costs for each period will be different, so the constraint (5.1b) or (5.1c) needs to be addressed for each period. Otherwise, system-level agency costs for each period converge to the average value as shown in (5.2). With this time-average costs approximation, the formulation follows the structure of a ‘resource allocation problem’, as defined by Ibaraki and Katoh (1988), and in recent related literature (Sathaye and Madanat, 2012); this approximation is adopted. For the extreme case where the ages of all segments are distributed in a narrow range (i.e. pavement systems are constructed or reconstructed in the same time period), the cost approximation is invalid.

$$\sum_{n=1}^N g_n^j(x_n) \cong \sum_{n=1}^N g_n^j(x_n)_\tau, \forall j, \tau \quad (5.2)$$

We consider two major activities and their respective costs in the PMS problem for flexible pavements (Asphalt Concrete pavements): reconstruction ($j = 1$) and rehabilitation ($j = 2$). Pavement reconstruction is the renewal of the entire structure of the existing pavement, and it is required when a pavement has either become functionally or structurally deficient, and where the damage extends below the surface layer (Caltrans 2012). A pavement life cycle length is the period between reconstructions. The length of this period depends on the traffic volume and loads, and the characteristics of the segment including structural design. Environmental factors and traffic demands are assumed to be time invariant in this paper. We assume that reconstruction consists of

replacing the existing structure by a similar structure. On the other hand, rehabilitation consists of the removal and replacement of the surface only. Rehabilitation does not affect the structural design, by overlaying the pavement with an asphalt concrete layer of the same thickness as the layer removed, and that it cannot improve the structural condition of underlying layers. Accordingly, structural damage below the surface layer can be restored only by reconstruction. Subsurface structural damage is positively related to the cumulative traffic loading, which is proportional to the age of the structure. Thus, to account for structural deterioration of the underlying layers, an augmented condition state $S_n(t)$ is defined: it contains age, $h_n(t)$, representing the number of years from the most recent reconstruction or construction to time t , as well as the pavement roughness, $s_n(t)$. Roughness is selected as an indicator of pavement surface condition instead of other distress indexes because it is the primary determinant of user costs. To consider multiple activities ($j = 1, 2, 3, \dots$) in the mathematical programming formulation (5.1), it is necessary to adopt appropriate performance and cost models for all different activities. In addition to reconstruction and rehabilitation, corrective and preventive maintenance activities can be taken into account. The solution methodology for this more general problem ($j = 1, 2, 3, \dots$) is a simple extension of that of the two-activity problem addressed in this paper ($j = 1, 2$).

5.1.2 Costs and Performance models

Our model treats state as continuous (roughness) and time as semi-continuous. The measurement unit of roughness is IRI (m/km) which is continuous. The meaning of semi-continuous time is that decisions are made at discrete time points, but pavement deterioration is represented in continuous time. The length of time between decisions, noted as δ , is typically one year. We use the roughness model developed by Paterson (1990), which is deterministic and history dependent. The continuous expression of pavement roughness, $F[S_n(t^+), u]$ ($= S_n(t + u)$), is represented as (5.3). The second term on the right-hand side of (5.3) indicates the effects of the traffic loading, l_n , and the structural number, SN_n , on the deterioration process. We assume l_n to be exogenous and independent of pavement condition. It is assumed that an activity is performed at the starting point of the t^{th} period. Here, the unit period length is not Γ but δ (it is set to one year in this research). At most one activity (reconstruction or rehabilitation) can be carried out on the single segment in the same time period.

$$F[S_n(t^+), u] = \{s_n(t^+)e^{bu} + a \cdot (SN_n + 1)^q \cdot l_n \cdot u \cdot e^{b(h_n(t^+)+u)}, u + h_n(t^+)\}, \forall u \in (0, 1] \quad (5.3)$$

where,

a : positive constant

q : negative constant

t^+ : $\lim_{u \rightarrow +0} t + u$

SN_n : structural number of segment n

l_n : the traffic loading per unit time on segment n

b : deterioration constant

User costs mainly consist of vehicle operating costs and travel time delays. The vehicle operating cost per unit time, $C_n^1[s_n(t)]$, is formulated as a function of pavement roughness as (5.4), where c_1 is the marginal operating cost of a car, c_2 is the marginal operating cost of a truck, c_3 is the fixed cost, $AADT_n$ is the time average traffic volume, and $AADTT_n$ is the time average truck traffic volume. The relative weight between user cost and agency cost is captured by the weight factor, $W_{u/a}$.

$$C_n^1[s_n(t)] = W_{u/a} \cdot ((c_1 \cdot (AADT_n - AADTT_n) + c_2 \cdot AADTT_n) \cdot s(t) + c_3) \quad (5.4)$$

Travel time delay arises when a roadway is resurfaced or reconstructed. The travel time delay cost caused by partial road closure during rehabilitation is denoted by C_n^2 , which is also positively related to traffic volume and weight factor, $W_{u/a}$. In this chapter, we assume that rehabilitation activities occur only during the nighttime with partial roadway closures, so $C_n^2 \cong 0$ for all segments. The travel delay user cost due to reconstruction, C_n^3 , results from complete roadway closure during reconstruction. We assume that the detour cost, C_n^3 , increases with traffic volume as shown in (5.5) with positive constant c_4 . Practitioners may use alternative forms of C_n^3 as needed.

$$C_n^3 = AADT_n \cdot c_4 \quad (5.5)$$

The rehabilitation cost and effectiveness models for segment n are shown in (5.6) and (5.7) respectively. The agency cost for the i^{th} rehabilitation of segment n , RH_n , is a linear function of overlay thickness, w_n^i , and it increases with number of lanes (including shoulders), D_n , as shown in (5.6), where m_1 , m_2 and m_3 are positive parameters. In (5.7), μ_1 , μ_2 and μ_3 are positive parameters. If rehabilitation is performed at time t , the original pavement roughness, $s_n(t^-)$, is reduced to $s_n(t^+)$ by $G(w_n^i, s_n(t^-))$, which is a function of the intensity of rehabilitation, w_n^i , and $s_n(t^-)$. The segment-level rehabilitation intensity is limited in the range as shown in (5.8).

$$RH_n[w_n^i] = D_n \cdot w_n^i \cdot m_1 + D_n \cdot m_2 + m_3 \quad (5.6)$$

$$G(w_n^i, s_n(t^-)) = \max\left(0, \min\left(s_n(t^-) - s_0, \frac{\mu_1 \cdot w_n^i}{\mu_2 \cdot s_n(t^-) + \mu_3} \cdot s_n(t^-)\right)\right) \quad (5.7)$$

$$\begin{aligned} 0 \leq w_n^i &\leq R(s_n(t^-)) \\ &= \frac{\mu_2 + \mu_3/s_n(t^-)}{\mu_1} \\ &\cdot (s_n(t^-) - \min(s_n(t^-), \max(s_0, (1 - \mu_1) \cdot s_n(t^-)))) \end{aligned} \quad (5.8)$$

The reconstruction cost per unit length of a highway pavement segment is a function of the structural number, SN_n , and the number of lanes, D_n . In this paper, the pavement design is not a

decision variable, so the reconstruction cost function (5.9) is fixed for every reconstruction of segment n . After the reconstruction, the pavement condition is improved to a new one, i.e., $S(\tau\delta^+) = \{s_{new}, 0\}$. In (5.9), m_4 , m_5 and m_6 are positive parameters.

$$RC_n = D_n \cdot SN_n \cdot m_4 + D_n \cdot m_5 + m_6 \quad (5.9)$$

5.2 Solution Approach

The formulation in (5.1) is re-organized into a problem consisting of the lower-level single segment problem and the upper-level system-level problem. The system-level problem is formulated as a constrained combinatorial optimization, and we will utilize GAs to solve this upper-level problem.

5.2.1 Segment-Level Optimization

The unconstrained segment-level optimization is formulated as a mathematical program shown in (5.10). The objective is to minimize the discounted cost-to-go of segment n , by selecting the best x_n , as shown in (5.10a). The surface condition after reconstruction is denoted by s_{new} , and the worst allowable roughness level is defined as s_{max} in (5.10b). Both the upper and lower allowable bounds (T_{min} and T_{max} respectively) of pavement life time are defined in (5.10c).

$$\min_{x_n} f_n(x_n) \quad (5.10a)$$

Subject to

$$s_{new} \leq s_n(t) \leq s_{max} \quad (5.10b)$$

$$T_{min} \leq T_n \leq T_{max} \quad (5.10c)$$

The segment-specific decision variable x_n is $\{T_n, H_n, \bar{t}_n\}$ including the life cycle length, T_n , defined as the time period between consecutive reconstructions, the number of rehabilitations in one life cycle, H_n , and their timings, $\bar{t}_n = \{t_n^1, \dots, t_n^{H_n}\}$, if $H_n > 0$, or $\emptyset$, otherwise. We assume that the intensity of a rehabilitation (i.e. overlay thickness) is the maximum in the effective range described in (5.8); this has been shown in previous research to be optimal (Li and Madanat, 2002; Ouyang and Madanat, 2006). The steady state properties are described in (5.11a) and (5.11b). As soon as the segment enters the steady state, it is sufficient to find decision variables only for one life cycle. Note that the steady state properties are guaranteed to be optimal even in the cases that some of decision variables are constrained (refer to the objective functions shown in (5.13) and (D.1)), as long as the models are Markovian and deterministic.

$$s_n(t) = s_n(t + pT_n), \forall \text{non negative integers } p \quad (5.11a)$$

$$h_n(t) = h_n(t + pT_n), \forall \text{non negative integers } p \quad (5.11b)$$

Optimal strategies for the constrained system-level problem cannot be derived only from the results of the unconstrained single segment optimization. Segment-level suboptimal alternatives, for which the agency costs are less than those of the unconstrained case, must be considered. To implement a GA to the system-level optimization, the number of sub-optimal alternatives for each segment should be finite.

Let $M_n = \{x_n^0, x_n^1, x_n^2, \dots, x_n^k, \dots\}$ be an alternative activity set for segment n in the steady state, where the elements are sorted by $f_n(\cdot)$ in ascending order. The zero in x_n^0 refers to the optimal activity from the unconstrained problem (5.10) and k represents the k^{th} alternative. $g_n^1(x_n)$ and $g_n^2(x_n)$ are defined as (5.12a) and (5.12b) respectively.

$$g_n^1(x_n) = \frac{RC_n \cdot \Gamma}{T_n} \quad (5.12a)$$

$$g_n^2(x_n) = \frac{\sum_{i=1}^{H_n} RH_n[w_n^i] \cdot \Gamma}{T_n} \quad (5.12b)$$

As H_n changes, $g_n^2(x_n)$ varies for a given T_n . $g_n^1(x_n)$ decreases as T_n increases for given H_n . Therefore, to satisfy the two-dimensional budget constraints shown in (5.1b) in the system-level optimization, alternatives for each segment need to be enumerated with respect to T_n and H_n independently. $g_n^2(x_n)$ is positively related to the number of rehabilitations, H_n , and their intensities (i.e. overlay thicknesses). As shown in (5.8), the intensity of a rehabilitation l , $R(s(\tau^-))$, depends on the pavement condition, $s(\tau^-)$, which depends on its timing t_n^i , so $g_n^2(x_n)$ is influenced by rehabilitation timings $\bar{t}_n = \{t_n^1, \dots, t_n^{H_n}\}$ for given H_n and T_n . An alternative rehabilitation timing set $\bar{t}_n'$, satisfying $g_n^2(T_n, H_n, \bar{t}_n') > g_n^2\left(T_n, H_n, \underset{\{\bar{t}_n\}|T_n, H_n}{\operatorname{argmin}} f_n(x_n)\right)$ for given T_n and H_n , should not be enumerated in M_n because its expected cost-to-go is higher than that of $\underset{\{\bar{t}_n\}|T_n, H_n}{\operatorname{argmin}} f_n(x_n)$. To find an alternative $\{H_n', \bar{t}_n'\}$, satisfying $g_n^2(T_n, H_n', \bar{t}_n') < g_n^2\left(T_n, \underset{\{H_n, \bar{t}_n\}|T_n}{\operatorname{argmin}} f_n(x_n)\right)$ for given T_n , we do not adjust the vector $\bar{t}_n$ for a given H_n to reduce the intensity of the rehabilitations, but instead reduce the number of rehabilitations H_n . This is reasonable because $g_n^2(x_n)$ is positively dependent on H_n due to the expensive fixed costs for each rehabilitation, $D_n \cdot m_2 + m_3$. If fewer rehabilitations are performed during a given life cycle, T_n , the roughness before each rehabilitation, $s_n(t^-)$, is larger, and thus the rehabilitation costs, RH_n^i , are larger. If $g_n^2(x_n)$ increases as H_n decreases from H_n of x_n^0 , the corresponding alternatives are filtered out from the enumeration by Algorithm 3 in Appendix D. If $g_n^2(x_n)$ decreases for a given T as H_n increases from H_n of x_n^0 , the corresponding alternatives should be considered to be enumerated. If $\exists H > H_n$ of x_n^0 , s.t. $g_n^2(x_n) > H \cdot (D_n \cdot R(s_{min}) \cdot m_1 + D_n \cdot m_2 + m_3)$, this H will be conservatively included as well. We assume that the timings of the rehabilitations are chosen to be optimal when T_n and H_n are given for all enumerated alternatives. The resulting single segment optimization problem for an alternative is shown in (5.13).

$$\min_{\{\bar{t}_n\}|T_n, H_n} f_n(x_n) \quad (5.13)$$

The number of alternatives included in M_n is $\in O(T_{max}^2)$. In the worst case, rehabilitation is performed every year, and $|M_n| \in O\left(\frac{T_{max}^2 - T_{min}^2 + T_{max} + T_{min}}{2}\right) \in O(T_{max}^2)$. We exclude all segment-level strategies x'_n from M_n , which satisfy $\exists x_n^k \in M_n, s. t. f_n(x_n^k) \leq f_n(x'_n), g_n^j(x_n^k) \leq g_n^j(x'_n), \forall j = 1, 2$, if the system-level constraints are (5.1b). If the system-level optimization is constrained by (5.1c), the segment-level strategy x'_n is excluded from M_n , if $\exists x_n^k \in M_n, s. t. f_n(x_n^k) \leq f_n(x'_n), g_n^1(x_n^k) + g_n^2(x_n^k) \leq g_n^1(x'_n) + g_n^2(x'_n)$. The details of the methodology for the segment-level alternatives enumeration are shown in Appendix D: algorithms 1a and 1b are for the problem formulated in (5.11); algorithm 2 is for the calculations of (5.12); algorithms 3, 4 and 5 are for the problem presented in (5.13).

5.2.2 System-Level Optimization

In this section, we focus on the solution algorithm for the system-level optimization constrained by (5.1b). From the segment-level optimization, the expected cost-to-go and time average expenses for two different activities are known for all elements in M_n , and for all n . The system-level problem formulated in (5.1) has $O(T_{max}^{2N})$ combinations of decision variables in the worst case. This exponential order of complexity is computationally infeasible for large N . As shown in the related literature (Chan et al. 1994; Yeo et al. 2013), GAs are effective in identifying close-to-optimal solutions in polynomial complexity for large-scale-system-level problems. The N -element set $x(\bar{K})$ is defined as (5.14) by using the solution vector $\bar{K}$.

$$x(\bar{K}) \equiv \{x_1(\bar{K}), \dots, x_N(\bar{K})\} = \{x_1^{k_1}, \dots, x_N^{k_N}\} \quad (5.14)$$

where,

$$x_n(\bar{K}) = x_n^{k_n} (\in M_n), \forall n$$

$$\bar{K} = \{k_1, \dots, k_N\}$$

The current solution vector $\bar{K}$ is initialized to a zero vector, $\mathbf{0}$. The zero vector means that $x(\mathbf{0})$ is the optimal system-level strategy of the unconstrained problem. A number of mutant offspring of $\bar{K}$ is generated randomly according to the pre-defined distribution, and each direction vector is denoted by $d\bar{K}$. The offspring pool is comprised of randomly generated $\bar{K} + d\bar{K}$ vectors and the current solution vector $\bar{K}$, where $d\bar{K} = \mathbf{0}$, and the next parent generation is chosen by evaluation and selection rules. For the initial solution vector $\bar{K} = \mathbf{0}$, $\sum_{n=1}^N g_n^1(x_n^0)$ and $\sum_{n=1}^N g_n^2(x_n^0)$ are calculated from the results derived from the segment-level enumeration process. We can check whether both $B_1 \geq \sum_{n=1}^N g_n^1(x_n^0)$ and $B_2 \geq \sum_{n=1}^N g_n^2(x_n^0)$ are satisfied together or not. If both are satisfied, the sum of expenses found from the system-level unconstrained optimization is less than the budget, any additional expenditure is suboptimal. In this case, the budget constraints are not binding. The situation described above may not occur in practice. However, if it does, it is unnecessary to implement a GA, and the following stages should be skipped.

The proposed GA has two steps: (i) moving toward the feasible region by offspring generation, evaluation and selection; and (ii) solution improvement. The input to the algorithm is $M_n = \{x_n^0, \dots, x_n^k, \dots\}, f_n(x_n^k), g_n^1(x_n^k)$ and $g_n^2(x_n^k) \forall n, k$, and the output is the close-to-optimal x .

Stage 1: Moving toward the feasible region by offspring generation, evaluation and selection

If the initial solution vector $\bar{K} = \mathbf{0}$ (zero vector) is infeasible in the original constrained problem, the vector is moved towards the feasible region, and it is possible to set the starting vector $\bar{K}$ randomly. The best moving direction is defined differently according to the current state. If budgets for both rehabilitation projects and reconstruction projects are fixed and known, we can categorize all possible scenarios of budget constraints according to $\sum_{n=1}^N g_n^1(x_n(\bar{K}))$ and $\sum_{n=1}^N g_n^2(x_n(\bar{K}))$ into four cases: (i) $B_j < \sum_{n=1}^N g_n^j(x_n(\bar{K})), \forall j = 1, 2$; (ii) $B_j < \sum_{n=1}^N g_n^j(x_n(\bar{K}))$ and $B_{j'} \geq \sum_{n=1}^N g_n^{j'}(x_n(\bar{K})), j \neq j'$; and (iii) $B_j \geq \sum_{n=1}^N g_n^j(x_n(\bar{K})), \forall j = 1, 2$. Until the parent generation of the solution vector reaches the feasible region (Case 4), the procedure is repeated. For the different cases, the following selection rules are proposed:

Case 1. $B_j < \sum_{n=1}^N g_n^j(x_n(\bar{K})), \forall j = 1, 2$

$$\min_{d\bar{K}} \sum_{n=1}^N f_n(x_n(\bar{K} + d\bar{K})) \quad (5.15a)$$

Subject to

$$\sum_{n=1}^N g_n^j(x_n(\bar{K} + d\bar{K})) \leq \sum_{n=1}^N g_n^j(x_n(\bar{K})), \forall i = 1, 2 \quad (5.15b)$$

Case 2. $B_j < \sum_{n=1}^N g_n^j(x_n(\bar{K}))$ and $B_{j'} \geq \sum_{n=1}^N g_n^{j'}(x_n(\bar{K})), j \neq j'$

$$\min_{d\bar{K}} \sum_{n=1}^N f_n(x_n(\bar{K} + d\bar{K})) \quad (5.16a)$$

Subject to (5.16b)

$$\begin{aligned} \sum_{n=1}^N g_n^j(x_n(\bar{K} + d\bar{K})) &\leq \sum_{n=1}^N g_n^j(x_n(\bar{K})) \\ \sum_{n=1}^N g_n^{j'}(x_n(\bar{K} + d\bar{K})) &\leq B_{j'} \end{aligned} \quad (5.16c)$$

Case 3. $B_j \geq \sum_{n=1}^N g_n^j(x_n(\bar{K})), \forall j = 1, 2$

The solution vector reaches the feasible region. Go to the next procedure.

Stage 2: Solution improvement

In this stage, the solution vector moves according to the following rules until it converges. Convergence does not guarantee the global optimality in GAs, so we need to check the existence of a better solution by repeating the solution improvement procedure with a different random distribution of $d\bar{K}$, or with different starting points until the solution converges.

$$\min_{dk} \sum_{n=1}^N f_n(x_n(\bar{K} + d\bar{K})) \quad (5.17a)$$

Subject to

$$\sum_{n=1}^N g_n^j(x_n(\bar{K} + d\bar{K})) \leq B_j, \forall j = 1, 2 \quad (5.17b)$$

5.2.3 Discussion of the Case of Two Budget Constraints

As noted by literature (Fwa et al. 1996; Guignier and Madanat 1999; Gu et al. 2012; Lee and Madanat 2014a), there exists a trade-off between different management activities. Guignier and Madanat (1999) use history-dependent deterioration models, and they numerically show that spending a larger annual average maintenance cost is optimal as the reconstruction frequency decreases. They emphasize that a joint optimization of different activities leads to significant savings compared to optimizing different strategies individually. Moreover, it has been shown that the total discounted cost-to-go increases as the available budget decreases in the system-level problem, when it is binding (Sathaye and Madanat 2012).

Let the optimal agency costs of the unconstrained problem $\min_{\sum_n g_n^1, \sum_n g_n^2} \sum_n f_n$ be denoted by $((\sum_n g_n^1)^{**}, (\sum_n g_n^2)^{**})$. This solution is plotted as the solid circle-shaped point in Figure 5.1. We can intuitively derive the following:

- Let $(\sum_n g_n^2)^* = \underset{\sum_n g_n^2 | \sum_n g_n^1 \leq B_1}{\operatorname{argmin}} \sum_n f_n \forall B_1 \leq (\sum_n g_n^1)^{**}$, which is defined by B_1 , and $(\sum_n g_n^2)^*$ non-increases as B_1 increases
- Let $(\sum_n g_n^1)^* = \underset{\sum_n g_n^1 | \sum_n g_n^2 \leq B_2}{\operatorname{argmin}} \sum_n f_n \forall B_2 \leq (\sum_n g_n^2)^{**}$, which is defined by B_2 , and $(\sum_n g_n^1)^*$ non-increases as B_2 increases

The trajectories of $(\sum_n g_n^1)^*$ and $(\sum_n g_n^2)^*$, from the one-budget-constrained problems, are conceptually represented as the bold line in Figure 5.1. However these trend lines are not always observed in the segment-level problems where $H_n = \underset{H_n | T_n}{\operatorname{argmin}} f_n$ and $T_n = \underset{T_n | H_n}{\operatorname{argmin}} f_n$. It is

possible that a shorter life cycle length can produce a lower expected cost-to-go for a given number of rehabilitations in a cycle. This is because of the discreteness of the decision variables. The feasible range of $\sum_n g_n^1$ that satisfies constraint (5.10c) is defined by the lower bound, $\sum_n RC_n/T_{max}$, and the upper bound, $\sum_n RC_n/T_{min}$. Therefore, a budget B_1 is infeasible if it is lower than $\sum_n RC_n/T_{max}$. Moreover, B_1 is always redundant, if it is higher than $\sum_n RC_n/T_{min}$. Based on the optimal properties, the following are true:

- The objective function value $\sum_n f_n$ corresponding $(\sum_n g_n^1, \sum_n g_n^2 \neq (\sum_n g_n^2)^*)$ is higher than $\sum_n f_n$ at $(\sum_n g_n^1, (\sum_n g_n^2)^*)$ for given $\sum_n g_n^1 \leq (\sum_n g_n^1)^{**}$.
- The objective function value $\sum_n f_n$ corresponding $(\sum_n g_n^1 \neq (\sum_n g_n^1)^*, \sum_n g_n^2)$ is higher than $\sum_n f_n$ at $((\sum_n g_n^1)^*, \sum_n g_n^2)$ for given $\sum_n g_n^2 \leq (\sum_n g_n^2)^{**}$.

The first bullet implies that spending more money on rehabilitation projects than $(\sum_n g_n^2)^*$ is suboptimal if the available budget for reconstruction projects is less than $(\sum_n g_n^1)^{**}$, that is, if only the budget constraint of B_1 is binding. This region of (B_1, B_2) is designated as Area A in Figure 5.1. For any point inside the Area A, such as point (a) in Figure 5.1, the optimal objective function value and policies are the same as those at the intersect of the bold line and the line perpendicular to the x -axis through the point (a). Similarly, if $B_1 \leq (\sum_n g_n^1)^{**}$ and $B_2 > (\sum_n g_n^2)^*$, which means that only the budget constraint for rehabilitation projects is binding, spending more funds on reconstruction than $(\sum_n g_n^1)^{**}$ is suboptimal. The region of (B_1, B_2) corresponding to this case is denoted in Figure 5.1 as Area B. Any point in area B, such as point (b) has the identical optimal solution to that of the point located at the end of the horizontal dotted arrow.

In the situation of $B_1 \geq (\sum_n g_n^1)^{**}$ and $B_2 \geq (\sum_n g_n^2)^{**}$ (Area C in Figure 5.1), the optimal solution is the result obtained from the unconstrained problem, where both budget constraints are not binding. This rarely happens in the real world, but we discuss optimal properties for all possible scenarios. The situation where $B_1 < (\sum_n g_n^1)^{**}$ and $B_2 < (\sum_n g_n^2)^{**}$ is the situation that is most common in the real world, where both constraints are binding. Finally, if $B_1 < \sum_n RC_n/T_{max}$, it is impossible to satisfy the conditions of maximum life cycle length, so it is infeasible (see the area INF).

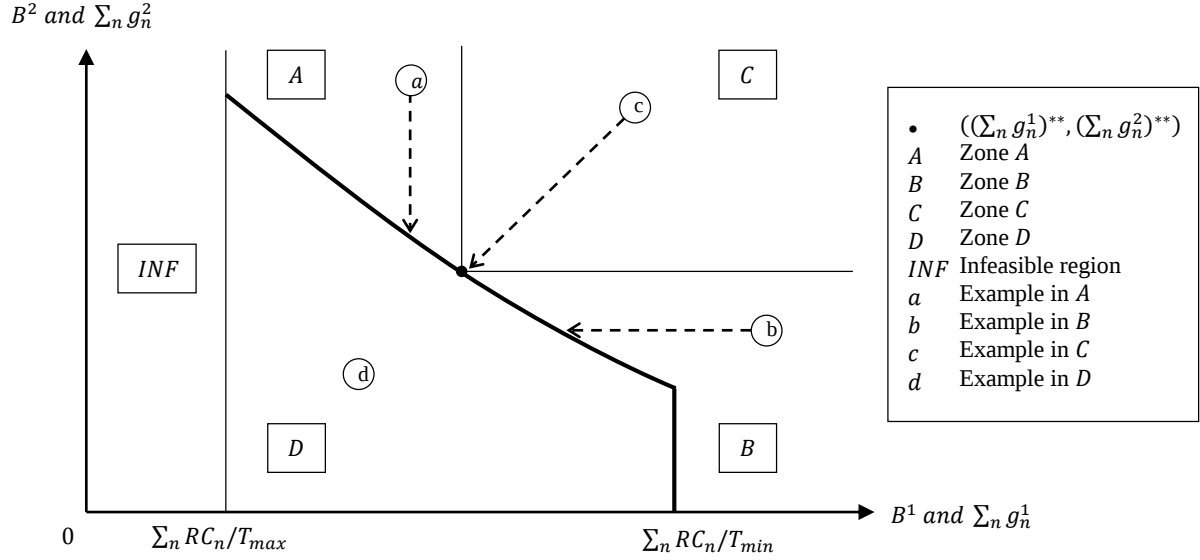


FIGURE 5.1 Graphical representation of the different regions for the optimal solution

5.3 Case Study

In this section, we demonstrate the solution algorithms from a real world pavement system example. The system used in the case study consists of 311 pavement segments, each 1 kilometer (one-directional) long, from California highways in Caltrans District 4. The selected highway system includes both urban and rural roads, and it has a range of pavement characteristics such as structural number, road width, traffic volume (AADT, AADTT), traffic loading (ESALs), etc. The solution algorithms are programmed in MATLAB, and the optimization problems are solved on a Windows 7 Professional OS with a 2.93 GHz processor and 3 GB RAM.

5.3.1 Parameters and Assumptions

Table 5.1 presents the cost model parameters used in the case study. We select a value of $W_{u/a}$ larger than 1, representing the case where user costs are more heavily weighed than agency costs.. A constant traffic loading is assumed along the planning horizon, so the parameter c_3 is set to zero, which does not influence the optimization results. We simply assume the costs parameters that are not related to the scale of the construction (e.g. thickness of overlay, number of lanes) to be zero.

Cost Parameter	Value	Units
r	0.07	-
$W_{u/a}$	2	-
c_1	0.001785	$\$/\text{IRI}/\text{car}/\text{km}$
c_2	0.004080	$\$/\text{IRI}/\text{truck}/\text{km}$
c_3	0	$\$/\text{km}/\text{yr}$
c_4	2.00	$\$/\text{AADT}/\text{km}$
m_1	413.28	$\$/\text{lane}/\text{mm}/\text{km}$
m_2	33012	$\$/\text{lane}/\text{km}$
m_3	0	$\$/\text{km}$
m_4	3414	$\$/\text{lane}/\text{SN}/\text{km}$
m_5	24000	$\$/\text{lane}/\text{km}$
m_6	0	$\$/\text{km}$

TABLE 5.1 Case study cost model parameter values

Table 5.2 presents the deterioration and improvement parameters as well as other policy parameters. The parameters determining the rate of the deterioration process are set higher than those estimated by Paterson (1990) because his data consisted of roadways in developing countries with lower traffic loads than those present in California.

Deterioration and Improvement Parameters			Policy Parameters		
Parameter	Value	Units	Parameter	Value	Units
b	0.025	-	s_{max}	6.0	IRI
a	725	-	T_{min}	40	yr
q	-4	-	T_{max}	80	yr
μ_1	0.66	-	δ	1	yr
μ_2	0.55	-	Γ	5	yr
μ_3	18.3	-			
s_{new}	1.0	IRI			
s_0	1.2	IRI			

TABLE 5.2 Case study deterioration, improvement policy parameter values

5.3.2 Segment-Level Results

Two one-way pavement segments with different traffic environments, located in Alameda County and Stanislaus County respectively, are selected as examples. The first segment carries relatively low traffic loading and volume, where AADTT is 382 trucks per day, AADT is 10900 vehicles per day, traffic loading is 0.054 million ESALs per year, structural number is estimated to 7.70, and the number of lanes including a shoulder is 4.69. The other selected pavement segment is on a relatively heavy traffic road, where AADTT is 1723 trucks per day, AADT is 56500 vehicles per day, traffic loading is 0.36 million ESALs per year, the structural number is estimated to 9.39, and the number of lanes including a shoulder in one-way is 4.69. The enumeration results from both examples are graphically presented in Figure 5.2 and Figure 5.3. Each bubble in Figures 5.3 or 5.4 indicates an alternative selected from the enumeration method described in Appendix D. The bubbles are shown on a two-dimensional plane according to their corresponding life cycle length on the x-axis and segment-level annual average rehabilitation costs $g_n^2(x_n^k)$ on the y-axis. The trend lines correspond to different occurrences of rehabilitations per unit life cycle. The double-lined circle in each figure stands for the optimal result from the unconstrained segment-level problem shown in (5.10). The bubble size of an alternative is proportional to its cost-to-go. Bigger bubbles stand for less cost-effective strategies.

As can be seen in Figure 5.2 (light traffic), in the unconstrained case, performing rehabilitation once in a life cycle is optimal, where the life cycle length is 75 years, and the optimal timing of the rehabilitation is 32 years after every reconstruction. Note that the expected cost-to-go increases as the life cycle length increases for a given number of rehabilitations in a cycle. This

is because the increase in user costs is larger than the decrease in agency costs, as the length of the life cycle increases and the number of rehabilitations is fixed.

The horizontal bottom line, consisting of alternatives without performing any rehabilitation, ranges from 47 years to 73 years. Alternatives corresponding to longer life cycle length than 73 years are not included in M_n , because the constraint of maximum allowable roughness (5.10b) cannot be satisfied without any rehabilitation activity. Alternatives corresponding to shorter life cycle length than 47 years produce higher objective values than the alternative with 47 years. Since a shorter lifecycle length means expensive annual reconstruction costs, those alternatives do not improve the system level optimization. As we can see, the number of possible alternatives for the single segment example, $|M_n| = 37 \in O(T_{max}^2)$, is finite and countable.

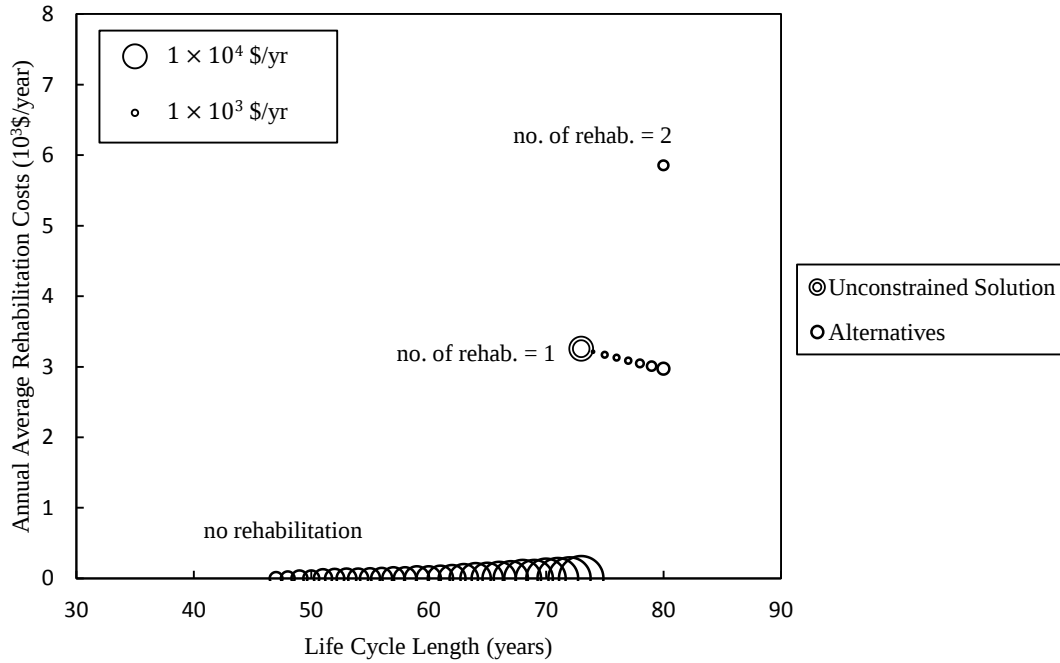


FIGURE 5.2 Enumeration of alternatives for the relatively light traffic highway

In Figure 5.3 (heavy traffic), the unconstrained optimal number of rehabilitations in one life cycle duration of 60 years, is found to be three. The timings of rehabilitations are 20, 34, and 47 years after every reconstruction. It is observed that the expected cost-to-go values decrease as number of rehabilitations increases for given reconstruction duration. Compared to the results from the lighter traffic highway, more frequent rehabilitations are required, which is consistent with the results of our previous research (Lee and Madanat, 2014a, etc.). The number of possible alternatives for the example of the heavy traffic highway segment, $|M_n|$, is 162.

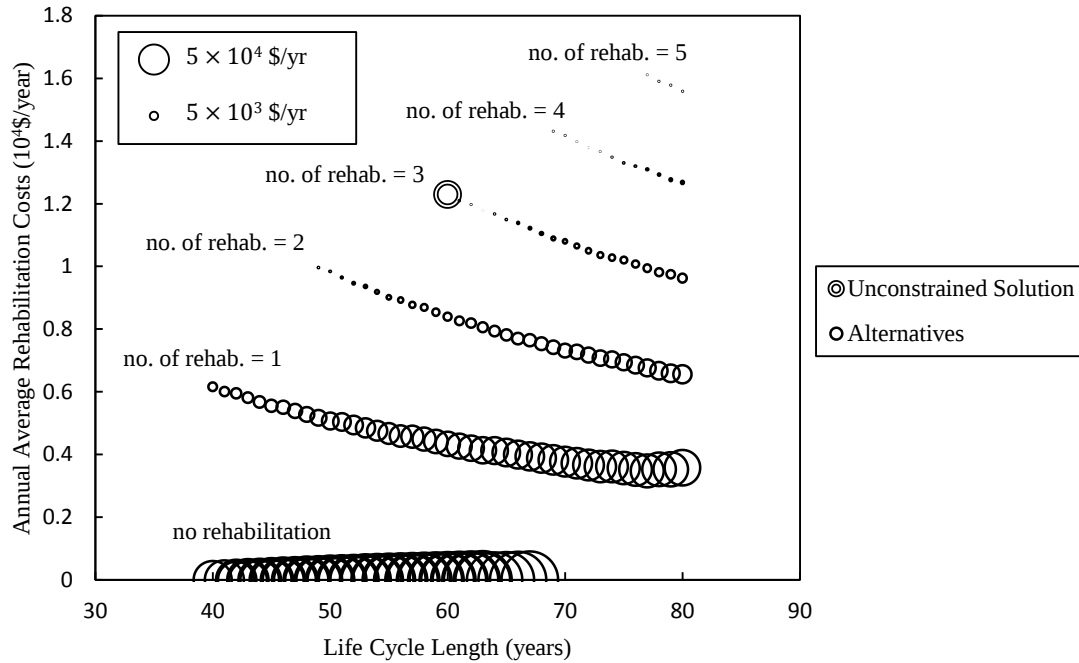


FIGURE 5.3 Enumeration of alternatives for the relatively heavy traffic highway

5.3.3 System-Level Results

In Figure 5.4, the square-shaped point in the middle of the two-dimensional plane indicates the optimal annual expenditures for both reconstruction and rehabilitation, produced from the budget-unconstrained problem. The expected cost-to-go at this point is 2.295×10^8 \$ with 1.921×10^6 \$/yr in reconstruction costs and 2.152×10^6 \$/yr in rehabilitation costs. Two lines meet at the unconstrained optimal point: a dotted line and a bold line. Along those lines, the objective values increase as we move farther from the unconstrained optimal point. The bold line represents the optimal expenditure for reconstruction when the budget for rehabilitation project is binding, but the budget for reconstruction is unlimited. The dotted line represents the optimal annual rehabilitation costs obtained when only reconstruction costs are limited. The two-dimensional plane can be partitioned into five regions as discussed in section 5.2.3, and each region is labeled as shown in Figure 5.4. All budget situations on the same vertical line in area A have the same optimal strategies and cost-to-go. Similarly, all points on a horizontal line in the area B have the same optimal strategies and objective value. The bold line intersects the x-axis at 2.311×10^6 \$/yr. The vertical boundary of the infeasible area will move to left, if the maximum allowable life T_{max} increases.

The average life cycle length of the budget-unconstrained results is 64.82 years, and the average number of rehabilitations for the 311 segments is 13.31 times per year. Along the dotted line starting from the unconstrained optimal point, the average life cycle length increases to 80 years which is the upper limit of pavement life cycle length, and the average number of rehabilitations increases to 15.65 times to compensate for the longer life cycle. The expected

objective value at the minimum reconstruction budget point on the dotted line, where all segments have maximum life cycle length, is $2.297 \times 10^8 \$$, and the sum of annual agency costs for both treatments are $4.017 \times 10^6 \$/yr$. A similar trend is observed along the bold line. As the budget for rehabilitation becomes binding, performing reconstruction is required more frequently. At the end point of the bold line intersecting the x-axis, only reconstruction is performed. The average life cycle length is 42.05 years which is slightly higher than the minimum life cycle length $T_{min} (= 40 yrs)$. Reconstruction is more expensive than rehabilitation because of its high equipment, material and labor costs and significant user costs caused by road closure, so performing reconstruction frequently is not optimal. The expected cost-to-go of the system is $2.505 \times 10^8 \$$ at this extreme point, which is worse than the extreme point of the opposite case, where the reconstruction budget is binding.

Most countries and states face limited budgets for all pavement management activities. This budget situation is represented by region D in Figure 5.4, where optimal expenses are always bounded by the budgets. We selected four scenarios, and those are shown on Figure 5.4 as circles. Scenarios are numbered from 1 to 4 according to the level of budget shortage. The increments of B_1 and B_2 between consecutive scenarios are $5 \times 10^5 \$/yr$ and $2 \times 10^6 \$/yr$ respectively; details are presented in Table 5.3.

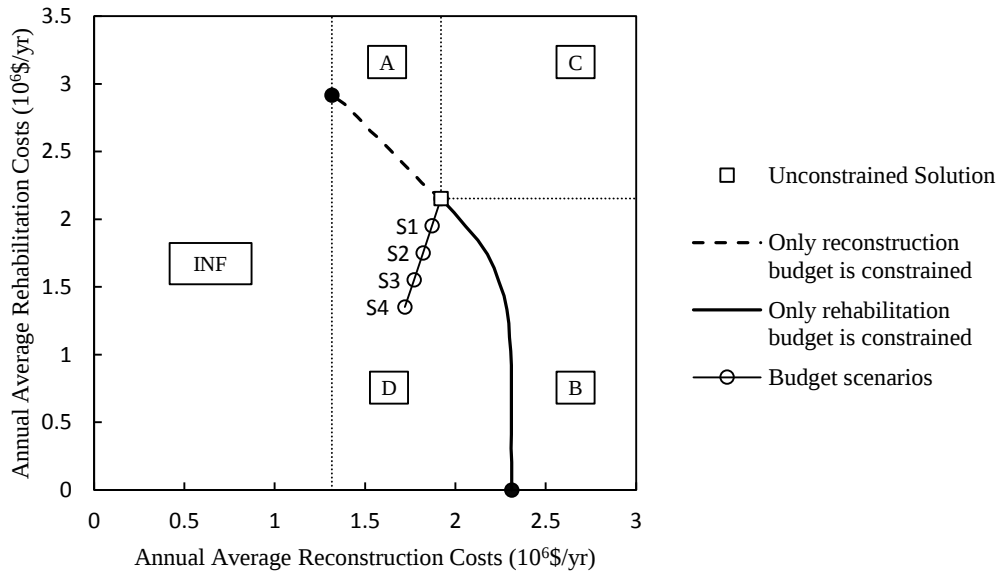


FIGURE 5.4 Results in the steady state and graphical representations of selected scenarios

Scenario	Budget values in (5.1b)		Combined single budget value in (5.1c) $B/\Gamma (= B_1/\Gamma + B_2/\Gamma)$
	B_1/Γ	B_2/Γ	
S1	1.871×10^6	1.952×10^6	3.823×10^6
S2	1.821×10^6	1.752×10^6	3.573×10^6
S3	1.771×10^6	1.552×10^6	3.373×10^6
S4	1.721×10^6	1.352×10^6	3.073×10^6

TABLE 5.3 Scenarios according to the constraint values

We solve two different optimization problems according to whether budget constraints are combined or separate as shown in (5.1b) and (5.1c). For each scenario, we compare the optimization results between the separate budget problem (B_1 and B_2) and the combined budget problem with the constraint value B , which is the sum of B_1 and B_2 . The results are shown in Figure 5.5. The square-shaped points refer to the results obtained in the separate budget problems, and the star-shaped points indicate those of the combined budget problems. As the budget shortage worsens, the objective values of both problems increase as shown by the bold lines in Figure 5.5. Due to the flexibility in allocating the budget, the optimal expected cost-to-go in the combined budget problem is less sensitive to the shortage. However, even in the combined budget problem, the discounted user costs increase significantly as the discounted agency costs decreases. The increments of budgets B_1 and B_2 between scenarios are constant, so the slopes of annual reconstruction costs and rehabilitation costs are constant as shown in two lines linking the square points, which are located in the lower part of Figure 5.5. In the combined budget problem, spending more money on rehabilitation and less on reconstruction is optimal. It can be seen that flexibility in allocation of funds among different activities is preferable, which is the same result reported in Guignier and Madanat (1999).

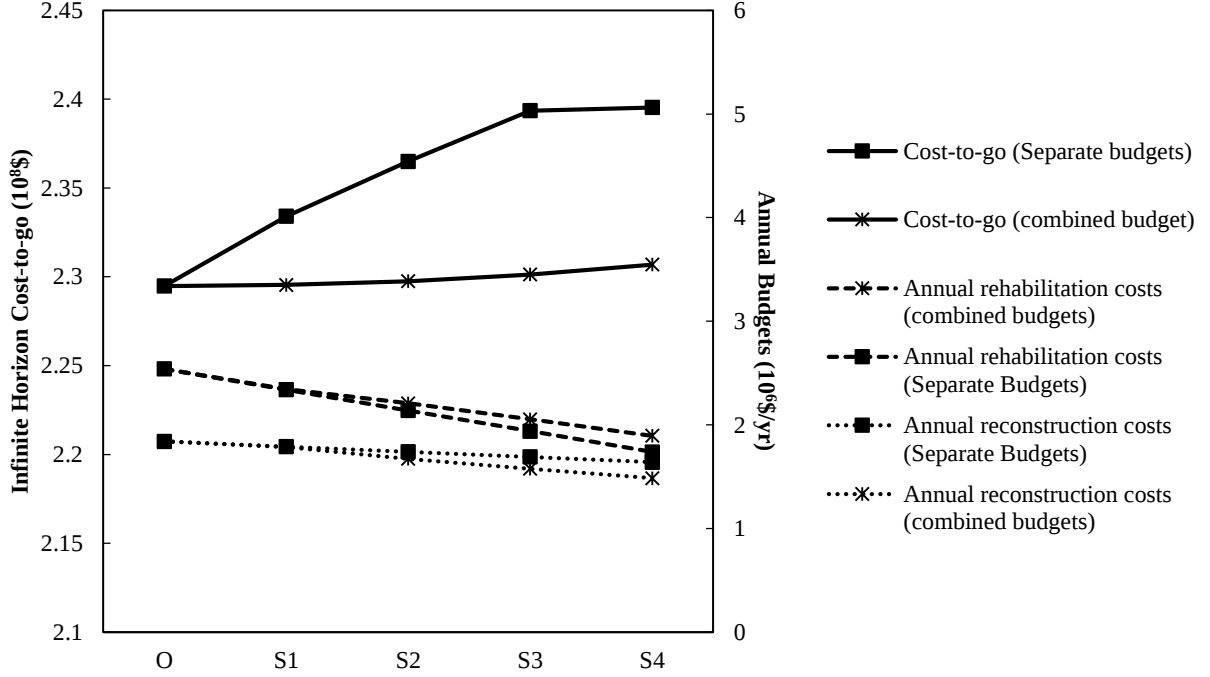


FIGURE 5.5 Optimization results of five scenarios for both the separate-budget problem and the combined-budget problem

To evaluate the optimality of the GA used in this paper, we compare the result found from the GA to the real optimal solution found by exhaustive search (i.e. brute-force search). Because the complexity of exhaustive search is exponential to the number of segments, it is computationally infeasible for the system used in the case study consisting of 311 segments. Instead, a simpler system, comprised of randomly selected 15 segments from the original system, is used. The real optimal costs is denoted by f^{OPT} . 100 experiments of the GA are carried out. Budgets are set to 80% of $(\sum_n g_n^j)^{**}$. The objective value found from each experiment is denoted by f^{GA} . The average value of f^{GA}/f^{OPT} is 1.005 and its sample standard deviation is 0.18 % of f^{OPT} .

To evaluate the execution speed of the proposed algorithm, 50 experiments are performed for systems consisting of 50, 100, 150, 200, 250 and 311 segments respectively. The average computation times are presented in Figure 5.6. For all cases, the sample standard deviation of f^{GA} is less than 0.05 % of the average of f^{GA} . As shown in Figure 5.6, the complexity of the algorithm has a polynomial trend with respect to the size of system.

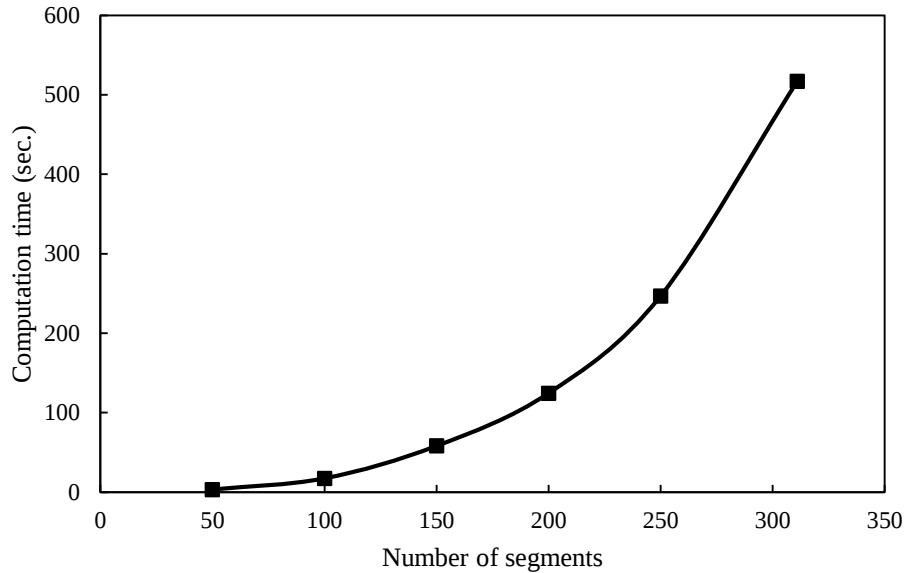


FIGURE 5.6 Average computation time for the proposed GA

5.4 Discussion

We have presented a mathematical formulation and solution for the system-level pavement rehabilitation and reconstruction problem considering segment-specific characteristics. The major contribution of this paper is that we incorporate the history-dependent aspects of pavement deterioration into the optimization of rehabilitation and reconstruction policies for a large system of pavements. The solution is based on MDP, and it focuses on steady state situations. The objective is to minimize the total discounted life time costs, consisting of agency costs for rehabilitation and reconstruction activities, as well as user costs, of a system of highway segments over an infinite time horizon.

We obtained optimal policies for two different budget situations: (i) reconstruction and rehabilitation projects are funded by different budget sources; and (ii) all activities are constrained by a single budget. Our case study presents the optimization results for both the segment-level problem and the system-level problem, and a range of optimal solutions as a function of the rehabilitation and reconstruction budgets.

The effect of uncertainty in the pavement deterioration models on the optimization results has been examined in previous research. For example, Sathaye and Madanat (2011 and 2012), who adopt the same deterioration models used in this paper, found that the optimal results are robust to the uncertainty in the deterioration-related parameters. Therefore, while many of the present PMS are based on probabilistic deterioration models, Sathaye and Madanat's results may indicate that our deterministic models provide sufficiently accurate results. Nevertheless, a systematic investigation of the effect of deterioration model uncertainty on the optimal policies and total costs for this problem is required to confirm this.

Chapter 6

Conclusion

This dissertation provides various optimization tools to find pavement construction, reconstruction and management strategies, which can be used for segment and system problems. As concluding remarks, the main contributions of this dissertation are summarized in Section 6.1, and possible research directions for future work are presented in Section 6.2.

6.1 Contributions

We incorporate the history-dependent aspects of pavement deterioration into the optimization of pavement design and management strategies to account for the influence of M&R history on the deterioration rate, which is physically more realistic. Three representations of optimization problems are separately addressed in terms of the size of pavement system and activities considered as decision factors. The objective is to minimize the total discounted life time costs, consisting of agency costs and user costs, over an infinite time horizon.

The subject of the first part of the dissertation is the joint optimization of pavement design and M&R strategies. For new construction or reconstruction, joint optimization is economical, because pavement deterioration rates are influenced by pavement design, so pavement design influences future M&R strategies after construction. The literature focuses on pavement design optimization and M&R strategies optimization separately. We find that optimal solution of the problem varies with traffic loading significantly. That is, under heavier traffic loading, a higher structural number for pavement design, higher levels of maintenance and a lower pavement-roughness trigger are required. Another important finding is that resurfacing planning converges to the steady state with a history-dependent deterioration model. In previous works, the steady state was guaranteed by the use of a memory-less deterioration model. In practice, convergence to the steady state gives us a simple form for the optimal solution in the case of a history-dependent deterioration model. The research results should be of use for developing countries in the process of expanding their highway networks in the face of multiple constraints.

In the second part of the dissertation, we present a methodology for the joint optimization of flexible pavement MR&R activities for the segment-level problem. The majority of research on the related subject do not optimize reconstruction jointly with maintenance and rehabilitation policies. We show that not accounting for reconstruction in maintenance and rehabilitation planning results in suboptimal policies for pavements undergoing cumulative damages in the underlying layers (base, sub-base or subgrade). We present dynamic programming solutions using an augmented state which includes current surface condition and age. The graphical representation of the optimal MR&R policy on the age-roughness plane allows decision-makers to easily find optimal current actions if they have information on the current roughness condition and records of recently performed treatments.

The third part of this dissertation presents a methodology for the joint optimization of various pavement management activities for heterogeneous pavement systems under multiple budget constraints. Unlike previous related works, we include reconstruction policies in the

system-level problem as well as less intensive activities such as rehabilitation. The solution is based on MDP, and it focuses on steady state situations. A bottom-up solution approach is proposed to reflect segment-specific characteristics. The complexity of the algorithm is polynomial in the size of the system and the policy-related parameters, so it is computationally feasible.

6.2 Future Work

The research done in this dissertation can be extended along several directions. Some examples of future research topics are presented as follows:

- In the first part of this dissertation, initial pavement design and M&R strategies are jointly optimized, but the reconstruction policy is not considered as a decision variable. In the second part, reconstruction policy is included in the future management policy, but its design is assumed to be identical to that of current facility. It is not guaranteed that current pavement has been optimally designed, so finding an economical pavement design, which will be applied in the next reconstruction, is necessary. The solution methodology of this problem will be the combination of the solution methodologies proposed in the first and second parts of this dissertation. Additionally, it will be more realistic if varying traffic demands along the time horizon are considered. We can intuitively expect that more durable design will be preferable if traffic demands increase. This future research will give a mathematical tool determining the optimal pavement design for the next reconstructions as well as future management policies which appropriately reflect the changes of traffic situations.
- The third sub-topic of the dissertation focuses only on steady state situations. The short-range optimal strategies for the present planning period depend on the current condition of the pavement segments in the system, so the optimal strategies for the transient period are different from the steady state strategies. Practitioners may be more interested in the short-term problem rather than the steady state strategies that will be reached in a distant future. Because short-term pavement MR&R optimization that do not account for future years may yield shortsighted and suboptimal policies, the preferred approach is the joint optimization of the transient period and steady state strategies.
- Reger et al. (2014) present an approach to assess the tradeoffs between life-cycle costs and GHG emissions in the system-level pavement rehabilitation problem. Their work can be improved by considering reconstruction as a decision factor in addition to rehabilitation. We have discovered the necessity and the importance of incorporating optimal reconstruction policy into the system-level pavement management problem from this dissertation. Since reconstruction yields huge amount of GHG emissions as well as significant agency costs due to its scale of process, it is quite important to find the best reconstruction policy which is feasible in both criteria: limited budgets and emission standards. The solution methodologies proposed in the third part of the dissertation can be utilized to solve this problem with an additional constraint on the total yearly emissions. These total yearly emissions consist of agency emissions and additional user emissions that result from the fuel consumption due to high level of roughness.

Bibliography

AASHTO, 1986. *Guide for design of pavement structures*. American Association of State Highway and Transportation Officials, Washington, D.C.

Barnes, G., Langworthy, P., 2004. Per mile costs of operating automobiles and trucks. *Transportation Research Record* 1864 (1), 71-77.

Bellman, R. E., 1955. Equipment replacement policy. *Journal of the Society of Industrial and Applied Mathematics* 8 (3), 133–146.

Benitez-Silva, H., Hall, G., Hitsch, G.J., Pauletto, G., Rust, J., 2000. *A comparison of discrete and parametric approximation methods for continuous-state dynamic programming problems*. Mimeo, SUNY at Stony Brook.

Buttler, H., Shortreed, J., 1978. Investment planning of a road link. *Transportation Research* 12, 357–366.

California Department of Transportation, 2012. *California Highway Design Manual California Department of Transportation*, Publication Distribution Unit, chapter 630 (www.dot.ca.gov/hq/opd/hdm/pdf/english/chp0630.pdf) (accessed on 7 March 2013).

Carnahan, J., 1988. Analytical framework for optimizing pavement maintenance. *Journal of Transportation Engineering* 114 (3), 307–322.

Carnahan, J., Davis, W., Shahin, M., Keane, P., Wu, M., 1987. Optimal maintenance decisions for pavement management. *Journal of Transportation Engineering* 113 (5), 554-572.

Chan, W., Fwa, T., Tan, C., 1994. Road maintenance planning using genetic algorithms. I: Formulation. *Journal of Transportation Engineering* 120 (5), 693-709.

Cheu, R., Wang, Y., Fwa, T., 2004. Genetic Algorithm-Simulation Methodology for Pavement Maintenance Scheduling. *Computer-Aided Civil and Infrastructure Engineering* 19 (6), 446-455.

Chootinan, P., Chen, A., Horrocks, M., Bolling, D., 2006. A multi-year pavement maintenance program using a stochastic simulation-based genetic algorithm approach. *Transportation Research Part A* 40 (9), 725-743.

Chu, J., Chen, Y., 2012. Optimal threshold-based network-level transportation infrastructure life-cycle management with heterogeneous maintenance actions. *Transportation Research Part B* 46 (9), 1123-1143.

Derman, C., 1962. On sequential decisions and Markov chains. *Management Science* 9 (1), 16–24.

Dreyfus, S., 1960. A generalized equipment replacement study. *Journal of the Society of Industrial and Applied Mathematics* 8 (3), 425–435.

Durango-Cohen, P., Sarutipand, P., 2007. Capturing interdependencies and heterogeneity in the management of multifacility transportation infrastructure systems. *Journal of Infrastructure Systems* 13 (2), 115–123.

- Feighan, K., Shahin, M., Sinha, K., White, T., 1988. Application of dynamic programming and other mathematical techniques to pavement management systems. *Transportation Research Record* 1200, 90–98.
- Fwa, T., Tan, C., Chan, W., 1994. Road maintenance planning using genetic algorithms. II: Analysis. *Journal of Transportation Engineering* 120 (5), 710–722.
- Fwa, T., Chan, W., Tan, C., 1996. Genetic-algorithm programming of road maintenance and rehabilitation. *Journal of Transportation Engineering* 122 (3), 246–253.
- Friesz, T., Fernandez, J., 1979. A model of optimal transport maintenance with demand responsiveness. *Transportation Research Part B* 13 (4), 317–339.
- Fernandez, J., Friesz, T., 1981. Influence of demand-quality interrelationships on optimal policies for stage construction of transportation facilities. *Transportation Science* 15 (1), 16–31.
- Ferreira, A., Antunes, A., Picado-Santos, L., 2002. Probabilistic segment-linked pavement management optimization model. *Journal of Transportation Engineering* 128 (6), 568–577.
- Gao, L., Zhang, Z., 2009. Approximate dynamic programming approach to network-level budget planning and allocation for pavement infrastructure. In: *Transportation Research Board 88th Annual Meeting*. Number 09-2344.
- George, K. P., Rajagopal, A. S., Lim, L. K., 1989. Models for predicting pavement deterioration. *Transportation Research Record* 1215, 1–7.
- Golabi, K., Kulkarni, R., Way, G., 1982. A statewide pavement management system. *Interfaces* 12 (6), 5–21.
- Gopinath, D., Ben-Akiva, M., Ramaswamy, R., 1994. Modeling performance of highway pavement. *Transportation Research Record* 1449, 1–7.
- Gopal, S., Majidzadeh, K., 1991. Application of Markov decision process to level-of-service-based maintenance systems. *Transportation Research Record* 1304, 12–17.
- Gu, W., Ouyang, Y., Madanat, S., 2012. Joint optimization of pavement maintenance and resurfacing planning. *Transportation Research Part B* 46 (4), 511–519.
- Guignier, F., Madanat, S., 1999. Optimization of infrastructure systems maintenance and improvement policies. *Journal of Infrastructure Systems* 5 (4), 124–134.
- Hawas, Y., 2004. Development and calibration of route choice utility models: Factorial experimental design approach. *Journal of Transportation Engineering* 130 (2), 159–170.
- Ibaraki, T., Katoh, N., 1988. *Resource Allocation Problems: Algorithmic Approaches*. MIT Press.
- Jido, M., Otazawa, T., Kobayashi, K., 2008. Optimal repair and inspection rules under uncertainty. *Journal of Infrastructure Systems* 14 (2), 150–158.
- Klein, M., 1962. Inspection–maintenance–replacement schedules under Markovian deterioration. *Management Science* 9 (1), 25–32.

- Kuhn, K., Madanat, S., 2005. Model uncertainty and the management of a system of infrastructure facilities. *Transportation Research Part C* 13 (5), 391-404.
- Kuhn, K.D., 2010. Network-level infrastructure management using approximate dynamic programming. *Journal of Infrastructure Systems* 16 (2), 103-111.
- Labi, S., Sinha, K.C., 2003. The effectiveness of maintenance and its impact on capital expenditures. Joint Transportation Research Program, Paper 208, Purdue University.
- Lee, J., Madanat, S., 2014a. Joint optimization of pavement design, resurfacing and maintenance strategies with history-dependent deterioration models. *Transportation Research Part B* 68, 141-153.
- Lee, J., Madanat, S., 2014b. Jointly optimal policies for pavement maintenance, resurfacing and reconstruction. *EURO Journal on Transportation and Logistics* 4 (1), 75-95.
- Li, Y., Madanat, S., 2002. A steady-state solution for the optimal pavement resurfacing problem. *Transportation Research Part A* 36 (6), 525-535.
- Madanat, S., 1993. Incorporating inspection decisions in pavement management. *Transportation Research Part B* 27 (6), 425-438.
- Madanat, S., Ben-Akiva, M., 1994. Optimal inspection and repair policies for infrastructure facilities. *Transportation Science* 28 (1), 55-62.
- Madanat, S., Bulusu S. and Mahmoud, A., 1995. Estimation of infrastructure distress initiation and progression models. *Journal of Infrastructure Systems* 1 (3), 146-150.
- Madanat, S., Prozzi, J., and Han, M., 2002. Effect of performance model accuracy on optimal pavement design. *Computer-Aided Civil and Infrastructure Engineering*, 17 (1), 22-30.
- Markow, M., Balta, W., 1985. Optimal rehabilitation frequencies for highway pavements. *Transportation Research Record* 1035, 31-43.
- Markow, M., Madanat, S., and Gurenich, D., 1993. Optimal rehabilitation times for concrete bridge decks. *Transportation Research Record* 1392, 79-89.
- Medury, A., Madanat, S., 2013. Incorporating network considerations into pavement management systems: A case for approximate dynamic programming. *Transportation Research Part C* 33, 134-150.
- McDonald, M., Madanat, S., 2012. Life-cycle cost minimization and sensitivity analysis for mechanistic-empirical pavement design, *Journal of Transportation Engineering* 138 (6), 706-713.
- Nunn, M., Brown, A., Weston, D., Nicholls, J., 1997. Design of long-life flexible pavements for heavy traffic. *TRL Report* 250.
- Ng, M., Zhang, Z., Waller, S., 2011. The price of uncertainty in pavement infrastructure management planning: an integer programming approach. *Transportation Research Part C* 19 (6), 1326-1338.
- Ouyang, Y., 2007. Pavement resurfacing planning for highway networks: parametric policy iteration approach. *Journal of Infrastructure Systems* 13 (1), 65-71.

- Ouyang, Y., Madanat, S., 2004. Optimal scheduling of rehabilitation activities for multiple pavement facilities: exact and approximate solutions. *Transportation Research Part A* 38 (5), 347–365.
- Ouyang, Y., Madanat, S., 2006. An analytical solution for the finite-horizon pavement resurfacing planning problem. *Transportation Research Part B* 40 (9), 767–778.
- Paterson, W.D.O., 1987. *Road deterioration and maintenance effects: Models for planning and management*. Johns Hopkins University Press, Baltimore, MD.
- Paterson, W. D. O., 1990. Quantifying the effectiveness of pavement maintenance and rehabilitation. In: *Proceedings at the 6th REAAA Conference*, Kuala Lumpur, Malaysia
- Ponniah, J.E., Kennepohl, G.J., 1996. Crack sealing in flexible pavements: a life-cycle cost analysis. *Transportation Research Record* 1529, 86–94.
- Pryke, A., Evdorides, H., and Abu Ermaileh, R., 2006. Optimization of pavement design using a genetic algorithm. Proc., *IEEE Congress on Evolutionary Computation*, IEEE, 1095–1098.
- Rajbongshi, P., and Das, A., 2008. Optimal asphalt pavement design considering cost and reliability. *Journal of Transportation Engineering* 134 (6), 255–261.
- Rashid, M., Tsunokawa, K., 2012. Trend curve optimal control model for optimizing pavement maintenance strategies consisting of various treatments. *Computer-Aided Civil and Infrastructure Engineering* 27 (3), 155–169.
- Reger, D., Madanat, S., Horvath, A., 2014. Economically and environmentally informed policy for road resurfacing: tradeoffs between costs and greenhouse gas emissions. *Environmental Research Letters* 9 (10), 104020.
- Robelin, C., Madanat, S., 2007. History-dependent bridge deck maintenance and replacement optimization with Markov decision processes. *Journal of Infrastructure Systems* 13 (3), 195–201.
- Robelin, C., Madanat, S., 2008. Reliability-based system-level optimization of bridge maintenance and replacement decisions. *Transportation Science* 42 (4), 508–513.
- Shahin, M. Y., Stock, C., Beckberger, L., 1994. Comparing pavement performance and its effect on maintenance and rehabilitation cost. *Transportation Research Board Conference Proceedings* 1 (1).
- Sathaye, N., Madanat, S., 2011. A bottom-up solution for the multi-facility optimal pavement resurfacing problem. *Transportation Research Part B* 45 (7), 1004–1017.
- Sathaye, N., Madanat, S., 2012. A bottom-up optimal pavement resurfacing solution approach for large-scale networks. *Transportation Research Part B* 46 (4), 520–528.
- Small, K., Winston, C., 1988. Optimal highway durability. *The American economic review*, 560–569.
- Tsunokawa, K., Schofer, J., 1994. Trend curve optimal control model for highway pavement maintenance: case study and evaluation. *Transportation Research Part A* 28 (2), 151–166.

Watanabe, T., Dhareshwar, A., Rezende Lima, P., 1987. *Vehicle speeds and operating costs. Models for Road Planning and Management*, The World Bank.

Yeo, H., Yoon, Y., Madanat, S., 2013. Algorithms for bottom-up maintenance optimisation for heterogeneous infrastructure systems. *Structure and Infrastructure Engineering* 9 (4), 317-328.

Zaabar, I., Chatti, K., 2010. Calibration of HDM-4 models for estimating the effect of pavement roughness on fuel consumption for U.S. conditions. *Journal of the Transportation Research Board*, 105-116.

Appendix A

Mathematical Details on the First Order Necessary Condition for Optimality of Rehabilitation

$\frac{dJ}{d\varepsilon_i}$ consists of four parts formulated at $\varepsilon_i = 0$: (A.1), (A.2), (A.3) and (A.4). All these parts are the functions of $s(t_{i-1}^+)$, τ_i and τ_{i+1} .

$$\begin{aligned} \frac{d}{d\varepsilon_i} \int_0^{\varepsilon_i} (F(s(t_{i-1}^+), \tau_{i+1} + u) - F(s(t_i^+), u)) e^{-ru} du &= \mu_1 \cdot s(t_i) \\ &= \mu_1 \{s(t_{i-1}^+) + \tau_i \cdot A(SN) \cdot l\} e^{b\tau_i} \end{aligned} \quad (\text{A.1})$$

$$\frac{d}{d\varepsilon_i} \int_{\varepsilon_i}^{\tau_{i+1}} (F((1 - \mu_1) \cdot F(s(t_{i-1}^+), \tau_i + \varepsilon_i), u - \varepsilon_i) - F(s(t_i^+), u)) e^{-ru} du \quad (\text{A.2})$$

$$= \begin{cases} -\frac{A(SN) \cdot l}{b - r} \left\{ (1 - e^{(b-r)\tau_{i+1}}) \cdot \left(\frac{r}{b - r} + e^{b\tau_i}(1 - \mu_1) \right) + b \cdot \tau_{i+1} \cdot e^{(b-r)\tau_{i+1}} \right\}, & \text{if } b \neq r \\ -A(SN) \cdot l \cdot \tau_{i+1} \{1 - e^{b\tau_i}(1 - \mu_1)\} - \frac{1}{2} \tau_{i+1}^2 \cdot b \cdot A(SN) \cdot l, & \text{if } b = r \end{cases} \quad (\text{A.3})$$

$$\begin{aligned} &\frac{d}{d\varepsilon_i} \{M(\mu_2 \cdot F(s(t_{i-1}^+), \tau_i + \varepsilon_i) + \mu_3) e^{-r\varepsilon_i}\} \\ &= m_1 \mu_2 \{A(SN) \cdot l \cdot e^{b\tau_i} + b(s(t_{i-1}^+) + \tau_i \cdot A(SN) \cdot l) e^{b\tau_i}\} \\ &\quad - r \{m_1(\mu_2(s(t_{i-1}^+) + \tau_i \cdot A(SN) \cdot l) e^{b\tau_i} + \mu_3) + m_2\} \\ &\frac{d}{d\varepsilon_i} \{M(\mu_2 \cdot F((1 - \mu_1) \cdot F(s(t_{i-1}^+), \tau_i + \varepsilon_i), \tau_{i+1} - \varepsilon_i) + \mu_3) e^{-r\tau_{i+1}}\} \\ &= -m_1 \cdot \mu_2 \cdot A(SN) \cdot l \cdot e^{(b-r)\tau_{i+1}} \{1 - (1 - \mu_1) e^{b\tau_i} + b\tau_{i+1}\} \end{aligned} \quad (\text{A.4})$$

Appendix B

Proof of Proposition 1

The initial roughness is known as s_{new} ($= s(0) = s(t_0^+)$), so (3.13) becomes an equation which has two unknown variables, τ_1 and τ_2 , because (3.13) for the i^{th} rehabilitation is the function of $s(t_{i-1}^+)$, τ_i and τ_{i+1} . Iteratively, we have a general equation which has two unknown variables as τ_i and τ_{i+1} , because $s(t_{i-1}^+)$ is function of τ_{i-1} and $s(t_{i-2}^+)$ for all $i \geq 2$. From the first order derivative of (3.13) with respect to τ_i , we can find $\frac{d\tau_{i+1}}{d\tau_i} \gg 1$ by calculation with the empirical parameters generally used in practice. This relation shows that τ_{i+1} is significantly sensitive to τ_i . The optimal solution of $\tau_i = \tau_{i+1} = \tau^*$ and $s(t_i) = s(t_{i+1}) = s^*$ for all i is only for the case of $s_{new} = s_0 = (1 - g_1)s^*$ where there is no permanent loss.

We need to show:

- i) $s(t_i) < s^*, \forall i$
- ii) $s(t_i) < s(t_{i+1}), \forall i$
- iii) $\tau_i > \tau_{i+1}, \forall i$
- iv) $\tau_i > \tau^*, \forall i$

Proof of i):

For a given $s(t_{i-1}^+)$, we define τ_i^0 and τ_i^1 as

$$\tau_i^0 \equiv \{\tau_i | F(s(t_{i-1}^+), \tau_i) \cdot (1 - \mu_1) = s(t_{i-1}^+), \tau_i \geq 0, \}$$

$$\tau_i^1 \equiv \left\{ \tau_i \left| \frac{dJ}{d\tau_i} = 0, \tau_i = \tau_{i+1} \geq 0, \right. \right\}$$

From (14), $\frac{d\tau_i^1}{ds(t_{i-1}^+)} < 0$. In the case of the first rehabilitation, $i = 1$, $s(t_{i-1}^+) = s_{new} < (1 - \mu_1)s^*$. For a rehabilitation given $s(t_{i-1}^+) < (1 - \mu_1)s^*$, $\tau_i^1 > \tau_i^0$ is found. Assume that there exists a rehabilitation i , $s(t_{i-1}^+) \geq (1 - \mu_1)s^*$ and $s(t_{i-1}^+) < (1 - \mu_1)s^*$. Because $\tau_{i-1}^1 < \tau^* < \tau_{i-1}$, it is clear that $\tau_i^1 < \tau_{i-1}^1 < \tau^* < \tau_{i-1} < \tau_i$. Therefore, the roughness threshold increases rapidly and $\frac{dJ}{d\tau_i} = 0$ is not solvable, at least with generally used parameters. It cannot be optimal, and the assumption turns to be false.

Proof of ii) and iii):

Assume that $\tau_i \leq \tau_{i+1}$. Therefore, $s(t_i^+) > s(t_{i+1}^+)$, when $s(t_{i-1}^+) < (1 - \mu_1)s^*$. From $\frac{d\tau_i^1}{ds(t_{i-1}^+)} < 0$, we can find $\tau_i^1 > \tau_{i+1}^1$. However, $\tau_{i+1}^1 < \tau_i^1 \leq \tau_{i+1}$ and $\frac{d\tau_{i+2}}{d\tau_{i+1}} \gg 1$, so $\tau_{i+1}^1 \ll \tau_{i+2}$. Therefore, the roughness threshold increases rapidly and $\frac{dJ}{d\tau_i} = 0$ is not solvable, at least with generally used parameters. The assumption of $\tau_i \leq \tau_{i+1}$ is not true, so $\tau_i > \tau_{i+1}$ and $s(t_i) < s(t_{i+1}), \forall i$.

Proof of iv) :

From i), ii) and iii), $\tau_i > \tau^*, \forall i$

Appendix C

Proof of $w^* = R(s(\tau\delta^-))$

Both w and $\frac{\partial J(S)}{\partial w}$ are equal to 0 except for policies that include rehabilitation. For policies with rehabilitation, the partial derivative of $J(S)$ in terms of w is (C.1).

$$\begin{aligned} \frac{\partial J(S)}{\partial w} &= \frac{\partial g(S, x)}{\partial w} + e^{-r\delta} \cdot \frac{\partial J(S')}{\partial w} \\ &= m_1 + \int_{0^+}^{\delta} \left\{ c_1 + (\alpha_1 e^{\beta_1 \Delta b} + \gamma_1) \cdot \frac{-\mu_1}{\mu_2 + \mu_3/s} \right\} e^{-ru} du + \frac{\partial J(S')}{\partial s'} \cdot e^{(b-r)\delta} \cdot \frac{-\mu_1}{\mu_2 + \mu_3/s} \\ &\quad , \text{if } I(a) = 1 \text{ or } 4 \end{aligned} \quad (C.1)$$

All three terms on the right-hand-side are independent of w , so $\frac{\partial J(S)}{\partial w}$ is independent of w . We consider three cases according to the sign of $\frac{\partial J(S)}{\partial w}$.

i) Case 1: $\frac{\partial J(S)}{\partial w} < 0$

$J(S)$ decreases as the intensity of rehabilitation increases. It is obvious that the maximum intensity of rehabilitation is optimal. If current surface condition is lower than the best achievable level of surface condition after rehabilitation, $s < s_0$, we assume that there is no improvement after rehabilitation. Thus, $w = 0$ is optimal, and $w > 0$ yields suboptimal results. For policies $I(x(S)) = 1$ and 4, $w = 0$ is always suboptimal in comparison to policies $I(x(S)) = 0$ and 3 respectively because S' is the same, and $g(S, x)$ is larger. Consequently, any value of w in a feasible range is suboptimal in the case of $s < s_0$. If $s_0 \leq s < 1/(1 - \mu_1)$, w is restricted by $\frac{\mu_2 + \mu_3/s}{\mu_1} \cdot (s - \max(s_0, (1 - \mu_1) \cdot s))$. If $1/(1 - \mu_1) < s$, w is binding by $\mu_2 s + \mu_3$. As a result, for policies $I(x(S)) = 1$ and 4, w is given by (C.2).

$$w \leftarrow \frac{\mu_2 + \mu_3/s}{\mu_1} \cdot (s - \min(s, \max(s_0, (1 - \mu_1) \cdot s))), \text{if } \frac{\partial J(S)}{\partial w} < 0 \quad (C.2)$$

ii) Case 2: $\frac{\partial J(S)}{\partial w} = 0$

Any value of w in a feasible range is optimal, so the intensity w can be set as $\frac{\mu_2 + \mu_3/s}{\mu_1} \cdot (s - \min(s, \max(s_0, (1 - \mu_1) \cdot s)))$.

iii) Case 3: $\frac{\partial J(S)}{\partial w} > 0$

$J(S)$ increases with the intensity of rehabilitation. It is obvious that $w = 0$ is optimal. When $w = 0$ is optimal for policies $I(x(S)) = 1$ and 4, policies $I(x(S)) = 0$ and 3 are always optimal compared to policies $I(x(S)) = 1$ and 4. Therefore, any value of

w is suboptimal. If we set w as $\frac{\mu_2 + \mu_3/s}{\mu_1} \cdot (s - \min(s, \max(s_0, (1 - \mu_1) \cdot s)))$, it does not affect the final optimal solution because policies $I(x(S)) = 1$ and 4 are not selected as an optimal policy regardless of the intensity of rehabilitation.

In every iteration, it is difficult to find the sign of $\frac{\partial J(S)}{\partial w}$ in general because of the difficulty in computing $\frac{\partial J(S')}{\partial s'}$ in (C.1) without approximating $J(S)$ as a differentiable function of s . However, based on i), ii) and iii), we can set w as (C.3) without considering of the sign of $\frac{\partial J(S)}{\partial w}$.

$$w \leftarrow \frac{\mu_2 + \mu_3/s}{\mu_1} \cdot (s - \min(s, \max(s_0, (1 - \mu_1) \cdot s))), \text{ if } I(x(S)) = 1 \text{ or } 4, \forall S \quad (\text{C.3})$$

Appendix D

Segment-level Alternatives Enumeration Algorithms

As a prior step for solving the segment-level problem in the steady state formulated in (5.13), the following mathematical program, formulated in (D.1), is solved. The objective is to minimize the expected cost-to-go by determining the optimal number of rehabilitations and their timings in one life cycle period, for a given H_n . In the steady state, the expected cost-to-go, $f_n(x_n)$, is a sum of geometric progression as shown in (D.2). The definition of $J_n(x_n, S_{new})$ is the discounted cost-to-go for the finite-horizon problem, where planning horizon is defined as one life cycle in the steady state. Because T_n is given, the problem (D.1) is equivalent to the minimizing $J_n(x_n, S_{new})$ by determining optimal decision variables H_n and $\bar{t}_n$.

$$\min_{\{H_n, \bar{t}_n\} | T_n} f_n(x_n) \quad (D.1)$$

$$f_n(x_n) = \frac{J_n(x_n, S_{new})}{1 - e^{-rT_n}} \quad (D.2)$$

The problem formulated in (D.1) is reformulated as (D.3) in Dynamic Programming, where $a_n(S)$ is a binary decision variable. $a_n(S) = 1$ means that rehabilitation is performed if pavement condition is S . It is a deterministic discounted finite-horizon optimization problem. δ is set as 1 year.

$$J_n^*(S) = \min_{a_n(S) \in [0,1]} \{g_n(S, a_n(S)) + e^{-r} \cdot J_n^*({s}', h + 1)\}, \forall s, \forall h \leq T_n - 1 \quad (D.3a)$$

s.t.

$$J_n^*({s}, T_n) = RC_n, \forall s \quad (D.3b)$$

$$g(S, X_n(S)) = \int_0^1 C_n^1[F[\{s(t^-) - a_n(S) \cdot G(R(s), s), h\}, u]]e^{-ru} du + a_n(S) \cdot RH_n[R(S)] \quad (D.3c)$$

$$s_{new} \leq s \leq s_{max} \quad (D.3d)$$

$$s' = F[\{s - a_n(S) \cdot G(R(s), s), h\}, 1] \quad (D.3e)$$

In appendix A, 5 algorithms are implemented to find M_n . To solve the above problem, we will find the approximation function of $J^*(S)$ for each $h \in \{0, \dots, T_n - 1\}$, which is continuous with respect to $s \in [s_{new}, s_{max}]$, and it is denoted by $V_n(s|h)$. We partition the range of pavement roughness $[s_{new}, s_{max}]$ into $Z - 1$ grids, and discretized roughness, s^z , is defined as $s_{new} +$

$\frac{(z-1)}{z} \cdot (s_{max} - s_{new}), \forall z \in \{1, 2, \dots, Z\}$. Two possible algorithms to solve the problem (D.3) with two different approximation methods of $V_n(s|h)$ are given in Algorithm 1a and Algorithm 1b respectively. More details on Algorithms 1a and 1b are found in Lee and Madanat (2014b).

Algorithm 1a: Piecewise-linear approximation of $V_n(s|h)$

Input: $T \in \{T_{min}, \dots, T_{max}\}$

Output: $V_n(s|h), \forall s \in [s_{new}, s_{max}], \forall h \in \{0, \dots, T\}$

1. $V_n(s|T) \leftarrow RC_n, \forall s, h \leftarrow T - 1$

2. Iteration process

While $h \geq 0$

$z \leftarrow 1$

while $z \leq M$

$$V_n(s^z|h) \leftarrow \min_{a_n(s^z, h) \in [0, 1]} \{g(\{s^z, h\}, a_n(s^z, h)) + e^{-r} \cdot a_n(s^{z'}|h + 1)\}, \forall z, s. t. s^{z'} \leq s_{max}$$

$z \leftarrow z + 1$

End while

$z \leftarrow 1$

while $z \leq Z - 1$

$$V_n(s|h) \leftarrow (s^{z+1} - s) \cdot V_n(s^z|h) + (s - s^z) \cdot V_n(s^{z+1}|h) \quad \forall s \in [s^z, s^{z+1}]$$

End while

$h \leftarrow h - 1$

End while

Algorithm 1b: Parameter approximation of $J_n(s|h)$

Input: $T \in \{T_{min}, \dots, T_{max}\}$

Output: $V_n(s|h), \forall s \in [s_{new}, s_{max}], \forall h \in \{0, \dots, T\}$

0. Define $V_n(s|h) \equiv \sum_{q=1}^Q \beta_n^q(h) \varphi^q(s)$

Where,

$\varphi^q(s)$: a continuous and differentiable basis function such as polynomials or log-poly

$\beta_n(h)$: set of parameters, $\{\beta_n^1(h), \dots, \beta_n^Q(h)\}$

1. $V_n(s|T) \leftarrow RC_n, \forall s, h \leftarrow T - 1$

2. Iteration process

While $h \geq 0$

$z \leftarrow 1$

While $z \leq Z$

$$V_n(s^z|h) \leftarrow \min_{a_n(s^z, h) \in [0, 1]} \{g(\{s^z, h\}, a_n(s^z, h)) + e^{-r} \cdot V_n(s^{z'}|h + 1)\}, \forall z, s. t. s^{z'} \leq s_{max}$$

$z \leftarrow z + 1$

End while

Estimate a set, $\beta_n(h)$, by the regression based on the Z samples of $V(s^z|h), \forall z$

$$V_n(s|h) \leftarrow \sum_q \beta_n^q(h) \varphi^q(s)$$

$h \leftarrow h - 1$

End while

At the end of the implementation of either of algorithms, $J_n(S_{new})$ is approximated as $V(s = s_{new}|h = 0)$, and $f_n(x_n)$ is found consequently by (D.2). The complexity of Algorithm 1a is $O(ZT_{max}^2)$ and that of Algorithm 1b is $O(Q^2 ZT_{max}^2)$. Those are polynomial in policy-related parameters and the parameters determining the fineness of discretization and the precision of the approximation. We denote the optimal number of rehabilitation for given life cycle period, T , resulting from the mathematical program (D.1), as $H_n(T)$. A solution methodology to find $H_n(T)$ is proposed in Algorithm 2.

Algorithm 2: Finding $H_n(T)$

Input: $V_n(s|h)$ achieved for given $T, \forall T \in \{T_{min}, \dots, T_{max}\}$

Output: $H_n(T)$,

1. $H_n(T) \leftarrow 0, h \leftarrow 0$ and $s \leftarrow s_{new}$

2. Iteration process:

While $h \leq T$

If $\operatorname{argmin}_{a_n(S) \in [0,1]} \{g(S, a_n(S)) + e^{-r} \cdot V(\{s', h\})\} = 1, s. t. s' \leq s_{max}$ then
 $H_n(T) \leftarrow H_n(T) + 1$ and $s \leftarrow F[\{s - G(R(s), s), h\}, 1]$ else $s \leftarrow F[\{s, h\}, 1]$

$h \leftarrow h + 1$

End while

For all $T \in \{T_{min}, \dots, T_{max}\}$, every alternative that has a higher number of rehabilitations in one life cycle of T than $H_n(T)$, is not selected as an element of M_n , because it yields both higher cost-to-go and higher period-average rehabilitation costs compared to those of the strategy corresponding to $H_n(T)$ and T . The output of the algorithm is M_n . The complexity of the algorithm for the worst case is $O(T_{max}^4)$.

Algorithm 3: Finding M_n

Input: $H_n(T), \forall T \in \{T_{min}, \dots, T_{max}\}$

Output: $M_n = \{x_n^1, \dots, x_n^k, \dots\}$ and their $g_n^1(x_n^k)$ and $g_n^2(x_n^k)$

1. $M_n \leftarrow \emptyset$

2. Iteration process 1:

$T \leftarrow T_{min}$

While $T \leq T_{max}$

$i \leftarrow \max\{0, H_n(T) - 1\}$

While $i \geq 0$

If $i = 0$ then $\bar{t} \leftarrow \emptyset$

Else run the GA (Algorithm 4) to find optimal $t, g_n^1(\cdot)$ and $g_n^2(\cdot)$ for given T and H

End if

If the constraint (10b) is satisfied then $x \leftarrow \{T, i, \bar{t}\}$ and $M_n \leftarrow M_n + \{x\}$

Else exit while loop

End if

$i \leftarrow i - 1$

End while

$T \leftarrow T + 1$

End while

3. Ascending sort all components in a set M_n according to $f_n(x_n)$ and numbered by $k = 0, 1, \dots$, where x_n^0 is the solution of the unconstrained segment-level problem.

4. Iteration process 2:

$k \leftarrow 1$

While $k \leq |M_n|$

If $\exists x_n^{k'} \in M_n$ s.t. $g_n^j(x_n^{k'}) \leq g_n^j(x_n^k) \forall j = 1, 2$ if $k' < k$ then $M_n \leftarrow M_n / \{x_n^k\}$ else $k \leftarrow k + 1$

End while

* Algorithms 1 and 2 can be skipped when M_n is found but the computing time significantly increases. To skip Algorithms 1 and 2, change “ $i \leftarrow \max\{0, H_n(T) - 1\}$ ” in step 2 in Algorithm 3 to “ $i \leftarrow T$ ”

For each T , we want to find a set of the optimal timings of rehabilitations for given i . This problem is formulated in (5.13), and the possible solution algorithm is Algorithm 4.

Algorithm 4: finding the optimal set of timings of rehabilitations for fixed number and life cycle length

Input: i and T

Output: t and f_n, g_n^1, g_n^2 of $x = \{T, i, \bar{t}\}$

1. Initialize $\bar{t}$

2. Stage 2.1 and stage 2.2 are repeated until $\bar{t}$ converges (i.e. there is no solution improvement)

2.1 Generate mutant offspring pool of $\bar{t}$ as $\bar{t} + d\bar{t}$ where moving direction $d\bar{t}$ is randomly generated movement vector satisfying $t^i + dt^i < t^{i+1} + dt^{i+1} < T \forall i \in \{1, \dots, i-1\}$ if $i \geq 2$. The number of offspring generated decide how precise search is

2.2 Offspring evaluation (Algorithm 5) and selection of the best mutant genotype $\rightarrow t$

3. Because the solution convergences in GAs do not guarantee the global optimality, check the optimality by changing the distribution of the random vector $d\bar{t}$ (or changing the initial $\bar{t}$) until there is no solution improvement, and the final results are $\bar{t}$ and $f_n(\{T, i, \bar{t}\})$

The algorithm 5 is calculating $f_n(\{T, i, \bar{t}\})$ and $g_n^j(\{T, i, \bar{t}\})$ for given $\{T, i, \bar{t}\}$, and its worst complexity is $O(T_{max})$.

Algorithm 5: evaluation of each mutant offspring

Input: $\{T, i, \bar{t}\}$

Output: $f_n(\{T, i, \bar{t}\})$, $g_n^1(\{T, i, \bar{t}\})$ and $g_n^2(\{T, i, \bar{t}\})$

1. $f_n(\{T, i, \bar{t}\}) \leftarrow 0$, $g_n^1 \leftarrow \frac{\theta \cdot RC_n}{T}$, $g_n^2(\{T, i, \bar{t}\}) \leftarrow 0$ $u \leftarrow 0$ and $s \leftarrow s_{new}$

2. Iteration process:

While $u \leq T$

2.1 If $u \in \bar{t}$ then $f_n(\{T, i, \bar{t}\}) \leftarrow f_n(\{T, i, \bar{t}\}) + \frac{e^{-r\delta u}}{1-e^{-rT}} \cdot g(S, a_n(S) = 1)$,
 $g_n^2(\{T, i, \bar{t}\}) \leftarrow g_n^2(\{T, i, \bar{t}\}) + \frac{\theta \cdot RH_n[R(s)]}{T}$ and $s \leftarrow F[\{s - G(R(s), s), u\}, \delta]$

Else $f_n(\{T, i, \bar{t}\}) \leftarrow f_n(\{T, i, \bar{t}\}) + \frac{e^{-r\delta u}}{1-e^{-rT}} \cdot g(S, X_1 = 0)$ and $s \leftarrow F[\{s, u\}, \delta]$

$u \leftarrow u + 1$

End while