

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Incorporating Network Considerations into System-level Pavement Management Systems

Permalink

<https://escholarship.org/uc/item/4x2073v0>

Author

Medury, Aditya

Publication Date

2013

Peer reviewed|Thesis/dissertation

Incorporating Network Considerations into System-level Pavement Management Systems

by

Aditya Medury

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

in

Civil and Environmental Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Samer Madanat, Chair
Professor Alexander Skabardonis
Professor Phil Kaminsky
Professor Zuo-Jun Max Shen

Fall 2013

**Incorporating Network Considerations into System-level Pavement
Management Systems**

Copyright © 2013
by
Aditya Medury

Abstract

Incorporating Network Considerations into System-level Pavement Management Systems

by

Aditya Medury

Doctor of Philosophy in Civil and Environmental Engineering

University of California, Berkeley

Professor Samer Madanat, Chair

The objective of transportation infrastructure management is to provide optimal maintenance, rehabilitation and replacement (MR&R) policies for a system of facilities over a planning horizon. While most approaches in the literature have studied it as a finite resource allocation problem, the presence of an underlying network configuration has been largely ignored. The recognition of the network configuration introduces several challenges, as well as opportunities, for system-level MR&R decision-making, which cannot be adequately handled by the existing decision-making frameworks.

This dissertation focuses on furthering the development of Markov decision process (MDP)-based system-level MR&R decision-making frameworks. In particular, two problems of interest are identified. The first problem concerns itself with identifying an optimal system-level optimization approach for solving budget allocation problems. The second problem of interest involves moving beyond traditional budget allocation problems to incorporate network considerations into system-level decision-making.

In the first part of the dissertation, a revised MDP-based optimization framework is proposed for solving the budget allocation problem. The framework, referred to as simultaneous network optimization (SNO), combines the salient features of the different MDP-based optimization approaches in infrastructure management literature, and provides optimal facility-specific MR&R policies for budget allocation problems. The proposed methodology is then compared with the other state-of-the-art MDP methodologies using a parametric study involving varying system sizes. The results of the study indicate the SNO outperforms the other MDP-based optimization frameworks.

In the second part of the dissertation, it is argued that while SNO is optimal for solving budget allocation problems, it can produce sub-optimal policies upon introducing network constraints. Consequently, the use of an approximated dynamic programming (ADP) framework is motivated to solve system-level MR&R decision-making problems involving network constraints. ADP facilitates the modeling of complex problem formulations by overcoming the curse of dimensionality associated with traditional dynamic programming frameworks.

To assess the suitability of ADP for system-level infrastructure management, two scenarios involving network considerations are investigated. In the first scenario, an approximate dynamic programming framework is proposed, wherein capacity losses due to construction activities are subjected to an agency-defined network capacity threshold. A parametric study is conducted on a stylized network configuration to infer the impact of network-based constraints on the decision-making process. The results indicate that ADP performs better than SNO when the network capacity constraints are binding on the decision-making process.

In the second scenario, the impact of introducing economies of scale (EOS) within budget allocation problems is investigated. Herein, incorporating network considerations leads to economic interdependence, wherein potential cost savings can be achieved by combining MR&R activities across adjacent road sections. Using parametric case studies, it is observed that the performances of ADP and SNO are comparable, with ADP improving upon the results of SNO under low budget and high EOS settings.

In conclusion, the findings from this dissertation indicate that ADP is a robust modeling framework for MDP-based infrastructure management problems. While previous research illustrates the use of ADP in solving system-level budget allocation problems, it is shown here that ADP is more relevant for modeling problems involving complex inter-facility dynamics. In particular, ADP is most beneficial in scenarios wherein finding optimal policies using analytical frameworks is not feasible.

To *Amma* and *Appa*.

Contents

Contents	ii
List of Figures	iv
List of Tables	v
1 Introduction	1
1.1 Introduction	1
1.2 Motivation	2
1.3 Dissertation Outline	3
2 Literature Review	4
2.1 MDP-based Optimization Frameworks	4
2.2 Incorporating network effects into system-level MR&R decision-making frame- works	6
2.3 Work Zone Scheduling Problems	6
2.4 Summary	7
3 A Simultaneous Network Optimization Approach for Pavement Man- agement Systems	8
3.1 Single Dimensional MDP-based Approach	8
3.2 Multidimensional MDP-based Approaches	10
3.3 A Simultaneous Network Optimization Approach	17
3.4 Parametric Study	19
3.5 Comparison of the facility-specific poicies	22
3.6 Discussion	23
4 Incorporating Network Considerations into the System-level Optimiza- tion Problem	25
4.1 Multidimensional MDP Problem	26
4.2 Approximate Dynamic Programming	27
4.3 Parametric Case Study	33

4.4	Comparison between ADP and other Network-based MR&R Decision-making Frameworks: A Qualitative Discussion	41
4.5	Discussion	43
5	Modeling Economic Interdependence Within System-level MR&R Decision-Making	44
5.1	Problem Formulation	44
5.2	Numerical Example	46
5.3	Results	48
5.4	Conclusions	48
6	Conclusion	49
6.1	Summary	49
6.2	Future Work	50
A	Cost Structure for Numerical Examples	52
B	Transition Probability Matrices	53
	Bibliography	56

List of Figures

1.1	A typical infrastructure management framework	1
1.2	Different types of network configurations	3
3.1	Comparison of the average system-level costs incurred by implementing SNO, Lagrangian relaxation, and TSBU	20
3.2	Comparison between the simulated costs and the expected system-level costs predicted at $t=1$ for SNO, Lagrangian relaxation, and TSBU	21
3.3	Comparison between facility-specific policies of SNO and TSBU for a facility from $n = 100$ scenario; For the action set, 1: do-nothing, 2: routine maintenance, 3: rehabilitation, and 4: reconstruction	22
4.1	Stylized 11-link pavement network for incorporating structural interdependence (individual link capacities expressed in brackets	34
4.2	Total costs comparison between SNO and ADP for different budget levels ($h=0.75$)	39
4.3	Box plots comparing ADP and SNO for varying budget levels and $h = 0.75$. . .	40
4.4	Convergence of the value function approximation for $B = 100$ scenario	41
4.5	Box plots comparing ADP and SNO in the absence of network constraints . . .	41
5.1	Modeling economies of scale for a network in series	45
5.2	Box plots comparing ADP and SNO across varying levels of EOS and budgets .	47

List of Tables

A.1	Cost structure for numerical example	52
B.1	Do nothing transition matrix	53
B.2	Routine maintenance transition matrix	54
B.3	Rehabilitation transition matrix	54
B.4	Reconstruction transition matrix	55

Acknowledgments

According to Wikipedia, the term “goat rodeo” refers to *a chaotic event where many things must go right for the situation to work*. In many ways, my time in Berkeley has been a goat rodeo, and for this, there are many people whom I wish to thank. Foremost, I would like to thank Samer for mentoring me through this journey. Through Samer’s diverse research interests, I have come to explore and appreciate a wide array of topics pertaining to transportation research. But it is as much through observing him and his interactions with fellow students including myself, that I have come to appreciate many other facets of an academic career. I have always been amazed at Samer’s ability to find time for his students, while maintaining extremely busy schedules as an administrator. I particularly cherish our frequent visits to coffee shops for meetings, and conferences have thus far proven to be a great opportunity to tap into Samer’s treasure trove of anecdotes. For instance, Samer has a keen sense of history about the academic community, and once I learnt from him at a conference that Prof. A, whom we had just met, was a student of Prof. B, who also happened to be the wife of late Prof. C, after whom the conference’s lifetime achievement award was named. It looks like I have happily caught the academic genealogy bug, for in the process of checking the facts for this anecdote, I discovered that one of Prof. C’s sons is also a Nobel laureate, whose talk I had the privilege of recently attending at Berkeley. While such trivia in and of itself may not carry much significance, but when put together, they lend an irresistible charm to the community one is seeking to be a part of.

Having been originally created as the Institute of Transportation and Traffic Engineering in 1948 through a California state legislature bill in response to the deferred maintenance of transportation facilities during World War II, Berkeley’s transportation research group has since made an imposing legacy for itself. Now rechristened as the Institute of Transportation Studies (ITS), a distinguishing feature of its transportation program has been the breadth of topics covered by its core faculty and affiliated faculty members, ranging from fundamental transportation areas, such as traffic flow theory and operations, transportation logistics, demand modeling, city and regional planning, to recent advancements, such as mobile sensing, intelligent transportation systems and life cycle assessment. At the same time, I have also felt inspired and humbled in equal measure by the camaraderie fostered by the members of this community. From the stories of late Prof. Gordon Newell playing ping-pong with his students and colleagues in the corridors of McLaughlin Hall to Profs. Daganzo and Cassidy teaming up for doubles in the annual ping-pong tournament to this day; from the baking experiments at the weekly cookie hours to the festive Christmas holiday parties involving Prof. Hansen’s reinterpretations of the Christmas carols - there is a potent mix of intellectual rigour and bonhomie that makes this program special. To this end, I am extremely grateful to its faculty members for being so generous in sharing their knowledge, ideas and experiences with us. In particular, Prof. Daganzo has been a constant source of inspiration for the vast body of work that he has accomplished. His child-like enthusiasm for research and instruction alike, and the simplicity of his nature shall long remain a cherished memory. I shall fondly remember the constant, reassuring presence of Joan when I first joined Berkeley,

and during the course of my stay, I have also come to greatly appreciate the influence she has had on shaping the behavioural component of transportation research at Berkeley. Alex (Skabardonis) perhaps best embodies the spirit of this department, marrying an exuberant personality with a very fine eye for detail. Finally, I would like to thank my dissertation committee members, Alex, Profs. Kaminsky and Max-Shen, for probing me when I formulated my research problems, encouraging me as I (eventually) made progress, and finally agreeing to sign on the completed manuscript.

Warm thanks must also go to my peers at varying stages of the program: Elliot Martin, Offer Grembek, Edward Huang, Josh Pilachowski, Athony Patire, Jing Xiong, Kwang Ho, Eric Gonzales, Vikash Gayah, Ilgin Guler, Eleni Christofa, GKB, Karthikgeyan Sivakumaran, Yiguang Xuan, Jeff Lidicker, Robert Campbell, Dylan, Ted Hunyh, Sowmya, Juan Argote, Michael Seelhorst, Yi Liu and Jean Doig. Their companionship, advice, encouragement, passion and work ethic have been a recurring driving force for me. As for my office mates at Club 116: Ipsita, DJ, Julia, Tierra, Vij, Haotian, Rui, Sebastian, Mo, Andre, Josh, Xiaofei, Darren and Jinwoo thank you for putting up with me, and sharing my successes, frustrations, and anxieties with each other, as if they were you own. Thank you!

Even as this acknowledgment threatens to reach epic proportions now, it is vital that I also convey my gratitude towards the Indian community at Berkeley. Studying and living away from India provided me with a unique opportunity to understand and explore my roots in isolation. In this regard, I found myself to be incredibly fortunate to have found a cohort of diverse individuals, with each person bringing to light a previously less understood aspect of my heritage. While the Center for South Asia Studies and SPICMACAY at Berkeley provided occasions on campus to discuss several India-based issues, our day-to-day interactions and conversations also threw up several opportunities to learn about topics, both Indian as well as global, academic and non-academic. To typecast their Indian-ness would be a disservice to their understanding of the ways of the world, and they were here, like everyone else, to pursue research as well. Hence, through them, I learnt that electrical engineering and computer science could imply natural language processing, human-computer interaction, information theory and semiconductor physics, and that research areas, such as control theory and computer aided design, are also common to the mechanical engineering department. They were my L^AT_EX and C++ instructors, and often my tech gurus. Then, to add to all this, we have shared our views on *chai*, cricket, composting, cooking, Ilaiyaraaja, Thyagaraja, among other topics. For that and more, I am indebted to them.

To my friends everywhere else, I (sheepishly) appreciate your patience in dealing with the extended periods of radio silence from my end. I look forward to restoring the communication channels shortly!

To my Tee, - . - . .. / -:- — — -.. :- / - .. -- .-... -.. .. / .-. .-. . - - ..-.-. -. -.-.

- . ..-.-.-.

Finally, to my family, thank you for your unconditional love and support. To my grandparents, thank you for grounding me in the manner in which you traverse through life making it seem so effortless. To my *Amma* and *Appa*, this degree is an outcome of the countless sacrifices made by you in raising me. I hope that I can pay it forward in times to come in a

manner that will best befit your courage, determination and selfless love.

My stay at Berkeley has been made financially feasible at different stages by the following sources: the ITS block grant, Northeast Gas Association, Samer's endowed professorship through Xenel Industries Ltd., Safe Transportation Research and Education Center (Safe-TREC) and the University of California Transportation Center (UCTC) dissertation grant. For their generosity, I shall remain eternally grateful to these institutions. I also wish to pay my gratitude to all the staff members at Berkeley whose work in the background makes the campus functional. Special thanks to Shelley Okimoto, Norine Shima, Kendra Levine, Rita Evans and Helen Bassham.

I hereby conclude my acknowledgments by thanking a motley group of entities whom/which I could not fit elsewhere: my undergraduate (IIT Delhi) and school teachers (in particular, Dr. Geetam Tiwari), The Topaz, friends in Berkeley who are neither TRANSOCians nor Indians (my sincerest apologies for so horribly miscasting you), Wikipedia, ESPNCRicinfo, Tanvi Chaubey, Netflix, Youtube, MusicindiaOnline, Saavn.

Chapter 1

Introduction

1.1 Introduction

Transportation infrastructure management is a decision-support framework used by agencies to manage and monitor their aging assets (roads, bridges, pipelines, etc.) under a resource constrained setting. It facilitates the identification of cost effective maintenance, rehabilitation and replacement (MR&R) policies by estimating the resulting improvements in the condition of the assets in the future years. In order to facilitate this decision-making, a typical infrastructure management framework involves three important components, as shown in figure 1.1.

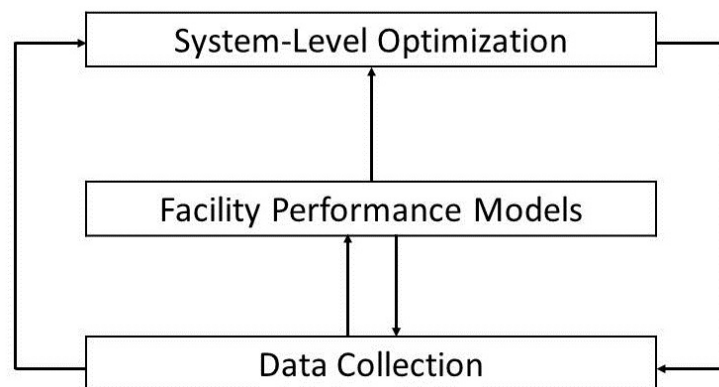


Figure 1.1: A typical infrastructure management framework

The first component of an infrastructure management system involves maintaining a repository of data pertaining to the assets, containing information such as the current and past condition states of the facilities, construction and maintenance history, weather indicators, traffic loading, etc. Using this information, facility performance models can be developed which provide the agency with some predictive power over how the facilities dete-

riorate over time. Finally, using the facility performance models and the information about the current state of the system, an optimization framework can be employed to select MR&R policies.

This dissertation focuses on the development of system-level MR&R optimization methodologies for infrastructure management systems, with a special emphasis on the management of road networks.

1.2 Motivation

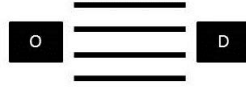
The goal of system-level infrastructure management decision-making is to identify optimal MR&R policies for a system of facilities. In the context of pavement management systems (PMS), the objective involves minimizing the agency expenditure as well as the costs incurred by road users (in the form of vehicle wear-and-tear, fuel usage, etc.) over a planning horizon. While most approaches in the literature have studied it as a problem of optimal allocation of limited financial resources, the interdependence between facilities, as introduced by a unifying network configuration, is often not accounted for. For instance, the implementation of MR&R activities on road networks can result in significant delays to travelers due to the loss in network capacity, detours, etc. According to one estimate, more than 60 million vehicles per hour per day of capacity were lost due to work zone activity on the National Highway System over a two week period in the United States in 2001 (Wunderlich and Hardesty 2003). Hence, given that the impact of scheduling work zones, especially in saturated traffic flow conditions, can be severe, it is important to systematically address and incorporate these user concerns within the decision-making process.

The recognition of an over-arching network configuration introduces several challenges, as well as opportunities, for system-level MR&R decision-making. Dekker et al. (1997) suggests that interactions between individual components of an infrastructure system can be classified into three different types: economic dependence (benefits/costs associated with joint maintenance), structural dependence (set of facilities collectively determining system performance such as connectivity or capacity) and stochastic dependence (presence of correlated deterioration factors like environment, loading). In the context of maintaining road networks, structural and economic interdependence are relevant issues which need to be incorporated within the MR&R decision-making process.

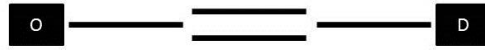
To further illustrate the concept of network-induced interdependence, consider a system comprising of four road segments. Without explicitly identifying the individual pavements within the network, it is impossible to gauge the impact of the proposed maintenance activities on the network performance. However, it can be inferred that if all the facilities in the system are arranged in series, as shown in figure 1.2(a), then each facility is critical for the functioning of the network. As a result, a partial/complete road closure during peak hours of traffic will adversely affect the traffic. In comparison, if all the road segments are in parallel (figure 1.2(b)), the network exhibits a very high level of redundancy. Consequently, potential road closures can be accommodated by rerouting traffic through parallel routes. A



(a) network in series



(b) network in parallel



(c) network with road segments in a combination of series and parallel

Figure 1.2: Different types of network configurations

more realistic network would perhaps comprise of links in both series and parallel, as shown in figure 1.2(c). Hence, in order to better mitigate the impact of construction activities on road users, the identification of optimal system-level MR&R policies should capture the relation between the road segments in a systematic manner.

1.3 Dissertation Outline

The outline of the dissertation is as follows; chapter 2 reviews the existing system-level MR&R decision-making frameworks and other relevant literature which helps formalize the research problems. In chapter 3, a system-level MR&R optimization methodology is proposed, which addresses some of the outstanding issues pertaining to budget allocation problems. Chapters 4 and 5 consider the structural and economic interdependence problems in the context of system-level MR&R decision-making. Finally, chapter 5 summarizes the findings from the dissertation and discusses the possibilities for extending this research further.

Chapter 2

Literature Review

The review of the existing literature pertaining to transportation infrastructure management is organized as follows. Section 2.1 discusses the different MR&R decision-making frameworks pertaining to transportation infrastructure management problems, with a special emphasis on Markov decision processes. Section 2.2 discusses the recent developments in the infrastructure management literature which take into account the network configuration within the decision-making process. Section 2.3 discusses work zone scheduling literature along with the strategies considered by state agencies to quantify the impact of construction activities on traffic operations. Finally, section 2.4 summarizes the literature to formalize the research problems.

2.1 MDP-based Optimization Frameworks

System-level MR&R decision-making paradigms in the transportation infrastructure management literature can be differentiated on the basis of their underlying assumptions of continuous/discrete condition state variables, continuous/discrete time horizons and/or deterministic/stochastic rates of deterioration. Continuous-time methods are useful for providing high-level strategic policies, but are not well suited for a detailed, tactical planing of MR&R activities. For such analysis, discrete-time methods are more extensively used in the literature. For a more detailed discussion of continuous-time frameworks, readers are encouraged to refer to Ouyang and Madanat (2006) and Sathaye and Madanat (2011), among others.

Discrete-state, discrete-time Markov decision process (MDP)-based frameworks have been widely used in infrastructure management, especially in the context of incorporating uncertainty in the underlying facility performance models. One of the first instances of using MDP frameworks for infrastructure management was the development of the Arizona pavement management system (Golabi et al. 1982). The LP-based approach utilized *randomized* policies to effectively accommodate budget constraints within the MR&R decision-making problem. Randomized policies are probabilistic in nature, wherein the optimal policy for a

facility in a given condition state is defined as a probability distribution function across two or more actions. The interpretation of randomized policies in a network setting relies on the assumption that all the facilities in the system deteriorate homogeneously. Consequently, such an approach can also be referred to as a *single dimensional MDP* problem. Other applications of this framework include the bridge management system, Pontis (Golabi and Shepard 1997), Smilowitz and Madanat (2000), Kuhn and Madanat (2005) and Madanat et al. (2006), among others.

Kallenberg (1994) shows that linear programming formulations for solving finite horizon MDP problems with constraints provide optimal solutions in the form of randomized policies. In the context of incorporating budget constraints, randomization of optimal policies can be interpreted as facilitating a complete utilization of the available budget, since a purely deterministic policy can either under-utilize or exceed the budget (Dimitrov and Morton 2009). However, the translation of the probabilistic policies into facility-specific decisions necessitates the presence of additional sub-routines, which is often cited as a limitation of randomized policy-based approaches.

In order to provide facility-specific decisions, other MDP-based optimization frameworks have focussed on obtaining facility-specific policies by using variants of the value iteration technique proposed by Richard Bellman (Bellman 1954). Since solving a system-level MDP problem using traditional dynamic programming techniques leads to computational challenges (referred to as the *curse of dimensionality*), researchers decompose the problem into two-stage (facility-level and system-level) problems. In the India Bridge Management System (Sinha et al. 1988), the Markovian deterioration assumption was relaxed to obtain deterministic performance models using regression techniques (Sinha et al. 1988). Ohlmann and Bean (2009) utilized a Lagrangian relaxation approach to relax the budget constraints, thereby decoupling the system-level MDP problem into facility-specific MDPs. Yeo et al. (2012) solved for a set of optimal and sub-optimal policies for each facility using a facility-level problem, and subsequently solved a multi-choice knapsack problem at the system-level which involves the budget constraint. An advantage of these techniques is that they do not assume the facilities to be homogeneous, and hence are classified as *multidimensional MDP* frameworks. However, since these approaches employ heuristics to simplify the system-level MDP problem, the optimality of the solution procedure is not guaranteed. In addition, most of these approaches are catered towards solving the resource allocation problem, and hence cannot always be extended to a general set of constraints.

In recent times, reinforcement learning/approximate dynamic programming (ADP) algorithms have also been applied to infrastructure management problems. Durango-Cohen (2004) applied some learning techniques to facility-level problems under imperfect deterioration information scenarios. Gao and Zhang (2009) and Kuhn (2010) utilized ADP frameworks to provide system-level MR&R policies in the context of budget allocation problems. Through the use of simulation techniques and value function approximations, ADP seeks to solve multidimensional MDP problems with complex constraints. However, one of the limitations of the approach is that theoretical optimality results are hard to guarantee. In addition, while Gao and Zhang (2009) and Kuhn (2010) demonstrate the applicability of

ADP-based approaches to infrastructure management problems, their performance vis-a-vis other MDP-based approaches was not evaluated.

2.2 Incorporating network effects into system-level MR&R decision-making frameworks

Network effects have primarily been incorporated into MR&R decision-making involving deterministic performance models. Ouyang (2007) used a deterministic infinite horizon dynamic program with multidimensional continuous state and control variables to account for travelers' route choices and the agency's resource allocation decisions simultaneously. Durango-Cohen and Sarutipand (2009) addressed the demand responsiveness associated with coordinating maintenance activities on a road network using a quadratic programming formulation. Ng et al. (2009) represented the traffic dynamics using a cell transmission model within a two-stage mixed integer programming formulation. Chu and Chen (2012) also proposed a bi-level programming formulation which provides threshold-based minimum cost MR&R policies using a deterministic user equilibrium problem to quantify the change in network traffic. In the context of MDP-based system-level MR&R decision-making, Furuya and Madanat (2012) extended the Yeo et al. (2012) framework to incorporate structural and economic interdependence in railway asset management.

In the management science literature, the economic interdependence problem has been studied as a *parallel machine replacement* problem, wherein the objective is to find a minimum cost policy for a set of economically interdependent machines across a planning horizon (Jones et al. 1991; Karabakal et al. 1994; Childress and Durango-Cohen 2005). Karabakal et al. (1994) proposed a mixed integer programming formulation involving a Lagrangian relaxation method to replace a set of machines within a budget constraint. Jones et al. (1991) and Childress and Durango-Cohen (2005) provide insights into the structure of optimal replacement strategies under non-resource constrained environments. For instance, the worse cluster replacement rule proposed by Childress and Durango-Cohen (2005) states that for problems with arbitrary cost functions, a machine is replaced only if all machines in worse states have been replaced.

2.3 Work Zone Scheduling Problems

On the operations side of infrastructure management, studies on work zone scheduling of maintenance activities have sought to incorporate the interaction between individual facilities at the system-level. Fwa et al. (1998) formulated an integer programming problem to provide an optimal schedule of activities which minimizes traffic delays caused by the resulting lane closures. Chang et al. (2001) employed a Tabu search-based methodology for an optimal work zone schedule. Hajdin and Lindenmann (2007) used the notion of corridors as a bundle of maintenance activities on various assets (roads, bridges, tunnels, etc.)

in order to determine optimal corridor lengths on a highway network that minimize user and agency costs. However, while these formulations highlight the interdependence between individual facilities at the system-level, the work zone scheduling problems do not concern themselves with the selection of MR&R activities, which also involves trade-offs associated with deferring maintenance.

Finally, at the policy level, state agencies utilize different construction options, which trade off duration, safety and loss of road capacity, based on the extent of closure undertaken. For example, urban freeway repair projects in California traditionally used seven or ten-hour nighttime closures because daytime closures were seen to cause unacceptable delays to weekday peak travel. However, nighttime closures are also associated with adverse impacts, such as poor safety for road users and construction crews, and longer closure times (Lee and Ibbs 2005). In recognition of these drawbacks, decision-support tools such as Construction Analysis for Pavement Rehabilitation Strategies (CA4PRS) have been developed in order to assess multiple highway rehabilitation strategies such as continuous (round-the-clock) operations during 55 hour weekend closures, or 72 hour weekday along with night-time closures, with the help of traffic simulation models.

2.4 Summary

Based on the overview of existing literature, there exist two major gaps in the literature, both in the area of MDP-based optimization frameworks, wherein this dissertation can contribute to:

1. The literature indicates that multiple MDP frameworks have been proposed to solve system-level MR&R decision-making problems involving resource constraints. Given their common underlying modeling assumptions, it is important to compare these approaches and benchmark them with each other so as to ascertain the relative merits and demerits of the frameworks in the context of budget allocation problems.
2. The literature also reveals that there exist two aspects to the management of infrastructure facilities: the financial allocation of resources for MR&R activities, which can be referred to as a planning problem, and the operational-level implementation, which pertains to the scheduling of the MR&R activities. Given the widespread usage of MDP-based optimization frameworks in transportation infrastructure management, it is important to extend the state-of-the-art MDP-based methodologies to integrate these two aspects of the decision-making process.

In the next chapter, a new optimization framework is proposed, which helps bridge the gap between single dimensional and multidimensional MDP decision-making frameworks for budget allocation problems.

Chapter 3

A Simultaneous Network Optimization Approach for Pavement Management Systems

Based on the literature review in chapter 2, it is observed that there exist multiple MDP-based frameworks which are aimed at providing MR&R policies in a resource constrained setting. These frameworks have been broadly classified into single dimensional and multidimensional MDP problems. In the case of single dimensional MDP formulations, the identification issues associated with randomized policies have thus far acted as a barrier for comparing these methods with the multidimensional MDP frameworks which provide facility-specific MR&R policies. In order to alleviate this issue, a simultaneous network optimization framework (SNO) is proposed in this chapter, which incorporates the favourable modeling techniques of both methodologies, while providing optimal facility-specific MR&R policies for budget allocation problems.

Prior to discussing the problem formulation of the SNO framework, it would be instructive to compare and contrast the formulations of the single dimensional and multidimensional MDP approaches. The frameworks presented below are believed to be the state-of-the-art approaches in their respective categories, and provide the necessary insights to develop the SNO framework. The discussion is restricted finite planning horizon problems, since the infinite planning horizon problem provides steady state policies, which are not always readily implementable.

3.1 Single Dimensional MDP-based Approach

For the implementation of the Arizona pavement management system, Golabi et al. (1982) proposed a linear programming formulation involving randomized policies, which has seen been widely used in the area of infrastructure management. In order to solve the decision-making problem in a given year t of a finite planning horizon of T years, the optimization

problem can be formulated as follows:

$$\min_w N \left(\sum_{s \in \mathbf{S}} \sum_{a \in \mathbf{A}} \left(\sum_{\tau=t}^T \alpha^{\tau-t} (c(s, a) + u(s)) w_{sa\tau} + \alpha^{T+1-t} \tilde{V}(s) w_{sa(T+1)} \right) \right) \quad (3.1)$$

subject to

$$\sum_{a \in \mathbf{A}} w_{sa\tau} = f_s^t \quad \forall s \in \mathbf{S} \quad (3.2)$$

$$\sum_{s \in \mathbf{S}} \sum_{a \in \mathbf{A}} w_{sa\tau} = 1 \quad \forall \tau = t+1, \dots, T, \quad (3.3)$$

$$N \left(\sum_{s \in \mathbf{S}} \sum_{a \in \mathbf{A}} c(s, a) w_{sa\tau} \right) \leq B_\tau \quad \forall \tau = t, \dots, T, \quad (3.4)$$

$$\sum_{r \in \mathbf{S}} \sum_{a \in \mathbf{A}} p_a(r, s) w_{ra\tau} = \sum_{a \in \mathbf{A}} w_{sa(\tau+1)} \quad \forall s \in \mathbf{S}, \forall \tau = t, \dots, T, \quad (3.5)$$

$$w_{sa\tau} \in [0, 1] \quad \forall s \in \mathbf{S}, \forall a \in \mathbf{A}, \forall \tau = t, \dots, T+1,$$

where,

$w_{sa\tau}$: fraction of facilities in state s to which action a is applied in year τ
(*randomized policies*),

f_s^t : fraction of facilities in state s in year t (the first year of optimization),

$c(s, a)$: cost incurred by the agency to implement action a , when a facility is in state s ,

$u(s)$: cost incurred by users due to vehicle wear-and-tear, when a facility is in state s ,

B_τ : agency's annual budget in year τ ,

$p_a(r, s)$: probability of a facility transitioning from state r to s , when action a is selected,

α : discount amount factor,

$\tilde{V}(s)$: salvage value associated with state s at the end of the planning horizon,

N : number of facilities in the network,

\mathbf{A} : action space associated with a facility (including do-nothing),

\mathbf{S} : state space associated with a facility.

Herein, equation 3.1 refers to the objective, which is to minimize the expected system-level user-plus-agency costs, incurred from year t to the end of the planning horizon; equation 3.2 represents the state of the system at the start of the optimization; equation 3.3 ensures that the randomized policies sum up to one for each year; equation 3.4 forces the expected agency expenditure to be within the annual budget constraint; and equation 3.5 represents the Chapman-Kolmogorov equations, which relate the policies of a given year with the policies of the subsequent year.

In order to implement the recommendations from the top-down approach, the randomized policies need to be associated with individual pavement sections, either using engineering

judgment, or with the help of additional sub-programs within the PMS. If the size of the network is sufficiently large, the policies implemented in the future time periods should also be consistent with the predictions of the randomized policies, due to the law of large numbers.

The LP formulation provides an optimal as well as a computationally attractive framework for solving the constrained MDP problem. The use of randomized policies allows for budget constraints to be imposed on all future actions, while maintaining the Markovian evolution of the state of the system. As a result, it provides agencies with a defensible procedure for preparing multi-year budget plans for MR&R decision-making. However, a limitation of the top-down approach is that the use of randomized policies precludes the identification of facility-specific actions from the optimization results.

3.2 Multidimensional MDP-based Approaches

3.2.1 Two Stage Bottom-Up Approach

In order to determine facility-specific policies for a pavement network, Yeo et al. (2012) formulated a two stage bottom-up (TSBU) approach, which consists of a facility-level and a system-level problem. In the first stage, the facility-level problem is solved to obtain optimal and sub-optimal policies for each facility, which act as inputs for the second stage. The system-level problem is then represented as a multi-choice knapsack problem, which incorporates the budget constraint for the current year. The decoupled nature of the formulation is motivated by the curse of dimensionality associated with solving a system-level dynamic programming problem involving a multidimensional state space.

3.2.1.1 Facility-level Problem

The objective of the facility-level optimization problem is to identify the optimal and sub-optimal policies for each facility, along with their associated to-go costs for each time period of a finite planning horizon. Herein, the optimal policy is defined by the action which minimizes the expected cost-to-go from the current year to the end of the planning horizon, for a given state s and year t . The motivation behind identifying the alternate policies is to provide greater flexibility with budget allocation at the system-level, since the sum of all the optimal policies might exceed the available budget. The optimization problem, represented as a discrete-state discrete-time MDP, can be solved using a backward-recursive dynamic programming approach with the following formulation:

$$a_{k\tau}(s) = \arg \min_{a \in \mathbf{A} - \{a_{j\tau}, j \leq k-1\}} \left[c(s, a) + u(s) + \alpha \sum_{r \in S} (p_a(s, r) V_{1(\tau+1)}(r)) \right], \quad \forall k = 1, \dots, |\mathbf{A}|, \quad \forall s \in \mathbf{S}, \forall \tau = t, \dots, T, \quad (3.6)$$

$$V_{k\tau}(s) = \min_{a \in \mathbf{A} - \{a_{j\tau}, j \leq k-1\}} \left[c(s, a) + u(s) + \alpha \sum_{r \in S} (p_a(s, r) V_{1(\tau+1)}(r)) \right], \quad \forall k = 1, \dots, |\mathbf{A}|, \quad \forall s \in \mathbf{S}, \forall \tau = t, \dots, T, \quad (3.7)$$

$$V_{1(T+1)}(s) = \tilde{V}(s), \quad \forall s \in \mathbf{S}, \quad (3.8)$$

where,

$a_{k\tau}(s)$: k^{th} optimal action when a facility is in state s in year τ ($k = 1$ is optimal),

$V_{k\tau}(s)$: expected cost-to-go associated with the k^{th} optimal action, from year τ to the end of the planning horizon, when a facility is in state s ,

$V_{1(T+1)}(s)$: salvage value associated with state s at the end of the planning horizon, $\tilde{V}(s)$.

In the absence of system-level constraints in the facility-level formulation, an assumption is being made that the future costs correspond to an optimal policy implementation, as denoted by $V_{1(\tau+1)}(r)$ (1=optimal) in equations 3.6 and 3.7. In effect, the formulation implies that optimality/sub-optimality is only restricted to the current year, and in the subsequent years, the budget would be sufficient for selecting the optimal actions for each facility. Finally, equation 3.8 specifies a state-dependent salvage value to the cost to-go function associated with the end of the planning horizon.

3.2.1.2 System-level Problem

The objective of the system-level problem is to allocate the annual budget for MR&R activities, so as to minimize the expected cost-to-go for the entire network. Using the ranked set of actions from the facility-level problem as an input, the problem is formulated as a multiple-choice knapsack problem:

$$\min_x \sum_{k=1}^{|\mathbf{A}|} \sum_{i=1}^N V_{kt}^{(i)}(\mathbf{s}_t(i)) x_{a_{kt}^{(i)}} \quad (3.9)$$

$$\text{subject to } \sum_{i=1}^N \sum_{k=1}^{|\mathbf{A}|} c(\mathbf{s}_t(i), a) x_{ia} \leq B_t, \quad (3.10)$$

$$\sum_{a \in \mathbf{A}} x_{ia} = 1 \quad \forall i = 1, \dots, N, \quad (3.11)$$

$$x_{ia} \in \{0, 1\} \quad \forall i = 1, \dots, N, \forall a \in \mathbf{A},$$

where,

- $a_{kt}^{(i)}, V_{kt}^{(i)}$: k^{th} optimal action and the corresponding expected cost-to-go for facility i , obtained from the facility-level problem for the year of decision-making, t ,
- $x_{a_{kt}^{(i)}}^{(i)}$: 1 if the action corresponding to $a_{kt}^{(i)}$ is selected for facility i ; 0 otherwise,
- $\mathbf{s}_t(i)$: condition state associated with facility i in year t .

Equation 3.9 represents the objective function, defined as the expected cost-to-go for the network, based on the actions selected for each facility; equation 3.10 indicates that the total amount spent on MR&R activities should be within the annual budget, and equation 3.11 ensures that exactly one action (including do-nothing) is selected for each facility.

The system-level problem requires information about the condition state of each facility at the beginning of each year through annual inspections. Hence, in order to implement the two stage bottom-up approach, the system-level optimization needs to be re-solved in each year of the planning horizon. On the other hand, since the facility-level problem is solved for the entire planning horizon, the optimal and alternate policies for each facility need not be calculated again.

The use of an integer programming formulation has the benefit of providing unique policies for individual facilities. However, the disjointed nature of this approach suffers from the limitation that the facility-level policies are developed without acknowledging the limited budget availability in the future years. Consequently, it is difficult to justify the optimistic assumption in the facility-level formulation of considering optimal policies in the future years.

3.2.2 Modified Two Stage Bottom-up Approach: Incorporating Lagrangian Relaxation Methods

In order to obtain non-randomized/deterministic policies for a collection of heterogeneous assets, Ohlmann and Bean (2009) employed a Lagrangian relaxation technique to account for the presence of budget constraints within the system-level optimization problem. While the objective of the MR&R decision-making problem is not to obtain deterministic policies, the Lagrangian relaxation approach can also be suitably adapted to re-order the optimal/alternate policies obtained from the TSBU approach so as to better reflect the resource constrained setting of the future years.

A deterministic policy implies that every condition state-time period pair has a unique action associated with it. Let the optimal deterministic policy matrix be defined as $\mathbf{A}^* = \{a_{s\tau}^*\}$, which acknowledges the presence of budget constraints in the current as well as future time periods. In order to identify these optimal policies, Ohlmann and Bean (2009) propose the following mathematical programming formulation:

$$(\mathbf{P}) \quad \min_{\mathbf{A}^*} \sum_{i=1}^N \sum_{s \in \text{mathbb{S}}} \left(\sum_{\tau=t}^T \alpha^{\tau-t} (c(s, a_{s\tau}^*) + u(s)) q_{i,s,a_{s\tau}^*,\tau} + \alpha^{T+1-t} \tilde{V}(s) q_{i,s,a_{s(T+1)}^*,T+1} \right), \quad (3.12)$$

subject to

$$\sum_{s \in \mathbf{S}} q_{i,s,a_{s\tau}^*,\tau} = 1 \quad \forall i = 1, \dots, N, \forall \tau = t, \dots, T, \quad (3.13)$$

$$\sum_{i=1}^N \sum_{s \in \mathbf{S}} c(s, a_{s\tau}^*) q_{i,s,a_{s\tau}^*,\tau} \leq B_\tau \quad \forall \tau = t, \dots, T, \quad (3.14)$$

$$\sum_{s' \in \mathbf{S}} P(s|s', a_{s'\tau}^*) q_{i,s',a_{s'\tau}^*,\tau} = q_{i,s,a_{s(\tau+1)}^*,\tau+1} \quad \forall i = 1, \dots, N, \forall s \in \mathbf{S}, \forall \tau = t, \dots, T, \quad (3.15)$$

$$\begin{aligned} a_{s\tau}^* &\in \mathbf{A}, \quad q_{i,s,a_{s\tau}^*,\tau} \in [0, 1] & \forall i = 1, \dots, N, \forall s \in \mathbf{S}, \\ & & \forall \tau = t, \dots, T+1, \end{aligned}$$

where,

$q_{i,s,a_{s\tau}^*,\tau}$: probability that facility i is in state s in year τ , when action $a_{s\tau}^*$ is implemented,

B_τ : agency's annual budget in year τ .

The problem formulation outlined above looks similar to the top-down approach described in section 3.1. However, the identification of a deterministic policy, \mathbf{A}^* , unlike a randomized policy, cannot be made using a linear programming approach, since the actions, $a_{s\tau}^*$, are only implicitly modeled using $q_{i,s,a_{s\tau}^*,\tau}$.

In order to solve (\mathbf{P}) , a Lagrangian relaxation approach is motivated, especially since the individual facilities are linked to each other only through the budget constraints (equation 3.14). Hence, relaxing the budget constraints using a Lagrangian multiplier, $\boldsymbol{\theta} \geq \mathbf{0}$, yields the following objective function:

$$\begin{aligned} (\mathbf{L}_\theta) \quad \min_{\mathbf{A}^*} \quad & \sum_{i=1}^N \sum_{s \in \mathbf{S}} \left(\sum_{\tau=t}^T \alpha^{\tau-t} (c(s, a_{s\tau}^*) + u(s)) q_{i,s,a_{s\tau}^*,\tau} + \alpha^{T+1-t} \tilde{V}(s) q_{i,s,a_{s(T+1)}^*,T+1} \right) \\ & + \sum_{\tau=t}^T \theta_\tau \left(\sum_{i=1}^N \sum_{s \in \mathbf{S}} c(s, a_{s\tau}^*) q_{i,s,a_{s\tau}^*,\tau} - B_\tau \right), \\ = \quad \min_{\mathbf{A}^*} \quad & \sum_{i=1}^N \sum_{s \in \mathbf{S}} \sum_{\tau=t}^T \alpha^{\tau-t} \left(c(s, a_{s\tau}^*) + u(s) + \frac{\theta_\tau}{\alpha^{\tau-t}} c(s, a_{s\tau}^*) \right) q_{i,s,a_{s\tau}^*,\tau} \\ & + \sum_{i=1}^N \sum_{s \in \mathbf{S}} \alpha^{T+1-t} \tilde{V}(s) q_{i,s,a_{s(T+1)}^*,T+1} - \sum_{\tau=t}^T \theta_\tau B_\tau, \end{aligned} \quad (3.16)$$

Since (\mathbf{L}_θ) relaxes some of the constraints within (\mathbf{P}) , its optimal solution represents a lower bound to the optimal solution of the primal problem. In other words, let Z^* be the optimal solution of (\mathbf{P}) , and let Z_θ^* be the optimal solution of (\mathbf{L}_θ) . For $\boldsymbol{\theta} \geq \mathbf{0}$, $Z_\theta^* \leq Z^*$.

In order to obtain the solution to the primal problem, the dual problem can be solved in the form of $\max_{\boldsymbol{\theta} \geq \mathbf{0}} \mathbf{L}_\theta$. The solution procedure proposed by Ohlmann and Bean (2009)

utilizes a subgradient optimization method, wherein the Lagrangian multipliers, $\boldsymbol{\theta}$, are iteratively adjusted by obtaining tight upper and lower bounds to (\mathbf{P}) . Herein, the lower bound is obtained by solving (\mathbf{L}_θ) , whereas an upper bound is obtained finding a feasible solution to the primal problem, (\mathbf{P}) .

3.2.2.1 Generating lower bounds: solving (\mathbf{L}_θ)

The benefit of relaxing the budget constraints in equation 3.16 is that the resulting problem formulation can be solved as N independent MDPs with a penalized cost function, $c_\theta(s, a, \tau)$:

$$c_\theta(s, a, \tau) = c(s, a) + \frac{\theta_\tau}{\alpha^{\tau-t}} c(s, a). \quad (3.17)$$

The penalized cost function is equivalent to the original cost function at $\boldsymbol{\theta} = \mathbf{0}$, which corresponds to the scenario when the budget constraint is not significant. However, $\boldsymbol{\theta} \neq \mathbf{0}$ indicates that the budget is a limiting factor, and it suitably penalizes the cost function to reflect the resource constraint. It is also important to note that the expensive MR&R actions become relatively more expensive in comparison to the cheaper MR&R actions. Consequently, the ordering of the optimal and the sub-optimal set of policies may also change from the constrain-free scenario.

Using $c_\theta(s, a, \tau)$, (\mathbf{L}_θ) can be solved for a given value of $\boldsymbol{\theta}$ by implementing the backward recursive dynamic programming approach previously described in the context of the facility-level problem of the TSBU approach:

$$a_{k\tau\theta}(s) = \arg \min_{a \in \mathbf{A} - \{a_{j\tau}, j \leq k-1\}} \left[c_\theta(s, a, \tau) + u(s) + \alpha \sum_{r \in S} (p_a(s, r) V_{1(\tau+1)\theta}(r)) \right], \quad \forall k = 1, \dots, |\mathbf{A}|, \\ \forall s \in \mathbf{S}, \forall \tau = t, \dots, T, \quad (3.18)$$

$$V_{k\tau\theta}(s) = \min_{a \in \mathbf{A} - \{a_{j\tau}, j \leq k-1\}} \left[c_\theta(s, a, \tau) + u(s) + \alpha \sum_{r \in S} (p_a(s, r) V_{1(\tau+1)\theta}(r)) \right], \quad \forall k = 1, \dots, |\mathbf{A}|, \\ \forall s \in \mathbf{S}, \forall \tau = t, \dots, T, \quad (3.19)$$

The sum of the penalized costs-to-go, $\sum_{i=1}^N V_{1t\theta}(s_{it})$, represents the resulting objective of (\mathbf{L}_θ) , wherein \mathbf{s}_t represents the condition state of the system in year t . However, since the budget constraints have been relaxed, it is possible that the optimal policies, (\mathbf{L}_θ) , represented by $A_\theta = \{a_{1\tau\theta}(s); \tau = t, \dots, T, s \in S\}$, may not always satisfy the budget constraints. Consequently, the lower bound provides an infeasible set of deterministic policies.

3.2.2.2 Generating upper bounds: obtaining a feasible solution

In order to update the Lagrangian multipliers, $\boldsymbol{\theta}$, it is important to identify a feasible upper bound to the primal problem. In Ohlmann and Bean (2009), the authors suggest a greedy repair heuristic wherein MR&R actions are greedily chosen from a set of sub-optimal actions $\{a_{k\tau}^\theta(s); k = 2, 3, 4, \dots\}$, in a manner that the budget constraint is satisfied in every year. As part of this work, an alternate heuristic is proposed which utilizes the system-level optimization problem of the TSBU approach.

The objective of the problem formulation is to select MR&R actions for a given year t , so as to minimize the expected penalized system cost-to-go subject to the budget constraint. Herein, an additional state variable is also used: f_{ist} is defined as the probability that facility i is in state s in year t .

$$\min_x \sum_{s \in \mathbf{S}} \sum_{k=1}^{|\mathbf{A}|} \sum_{i=1}^N f_{ist} V_{kt\theta}^{(i)}(s) x_{s, a_{kt\theta}^{(i)}(s)} \quad (3.20)$$

subject to

$$\sum_{i=1}^N \sum_{s \in \mathbf{S}} \sum_{a \in \mathbf{A}} f_{ist} c(s, a) x_{sa}^{(i)} \leq B_t, \quad (3.21)$$

$$\sum_{a \in \mathbf{A}} x_{sa}^{(i)} = 1 \quad \forall i = 1, \dots, N, \forall s \in \mathbf{S} \quad (3.22)$$

$$x_{sa}^{(i)} \in \{0, 1\} \quad \forall i = 1, \dots, N, \forall a \in \mathbf{A}, \forall s \in \mathbf{S},$$

where,

$a_{kt\theta}^{(i)}(s), V_{kt\theta}^{(i)}(s)$: k^{th} optimal action and the corresponding penalized expected cost-to-go for facility i in state s in year t (obtained from equations 3.18-3.19),

$x_{s, a_{kt\theta}^{(i)}(s)}^{(i)}$: 1 if the action corresponding to $a_{kt\theta}^{(i)}(s)$ is selected; 0 otherwise.

If the current state of the system is identified, f_{ist} is a binary variable, and the problem formulation is equivalent to the system-level problem of the TSBU approach presented in equations 3.20-3.22. However, the advantage of using f_{ist} is that once the optimal actions are chosen for the given time period, the optimization for the next time period can also be implemented by identifying $f_{is(t+1)}$ as follows:

$$f_{is(t+1)} = \sum_{a \in \mathbf{A}} \sum_{s' \in \mathbf{S}} P(s|s', a) f_{is't} x_{s'a}^{(i)} \quad \forall i = 1, \dots, N, \forall s \in \mathbf{S}, \quad (3.23)$$

Once $f_{is(t+1)}$ is constructed, the solution procedure can be repeated until the end of the planning horizon. Finally, an upper bound, Z_f^* , to the primal problem (\mathbf{P}) can be constructed as follows:

$$Z_f^* = \sum_{i=1}^N \sum_{s \in \mathbf{S}} \sum_{a \in \mathbf{A}} \left(\sum_{\tau=t}^T \alpha^{\tau-t} (c(s, a) + u(s)) \tilde{q}_{i,s,a,\tau} + \alpha^{T+1-t} \tilde{V}(s) \tilde{q}_{i,s,a,T+1} \right) \quad (3.24)$$

where,

$$\tilde{q}_{i,s,a,\tau} = f_{is'\tau} x_{s'a}^{(i)} \quad \forall i = 1, \dots, N, \forall s \in \mathbf{S}, \forall a \in \mathbf{A} \quad (3.25)$$

Hence, by using f_{ist} , the system-level problem presented here, unlike the Monte Carlo simulation-based approach of the TSBU approach, can be solved for the future years to obtain a more realistic estimate of the expected future costs.

It is important to note here that while the original problem proposed by Ohlmann and Bean (2009) was intended to obtain a set of deterministic policies for a resource constrained MDP setting, TSBU and the Lagrangian relaxation approach do not always provide deterministic policies. TSBU provides a feasible solution to the MR&R decision-making problem, and the objective of the Lagrangian relaxation approach is to modify the structure of the optimal/sub-optimal policies so as to improve the quality of the feasible solution.

3.2.2.2.1 Adjusting the Lagrangian multiplier

In order to update the Lagrangian parameters, θ , Ohlmann and Bean (2009) employ a sub-gradient optimization method. The motivation behind the technique is to increase/decrease the value of θ as a function of the gap between the best upper (U) and lower (L) bounds, as well as the extent of budget constraint violation. The pseudocode of the algorithm as presented by Ohlmann and Bean (2009) is as follows:

Initialization:

Let $\delta = 2, \epsilon = 0.001, c = 0, noImprovCount = 0$, and $countLimit = 250$.

Let $bestUpperIter = 1$ and $bestLowerIter = 1$.

Let $\theta(c) = 0, U = \infty, L = -\infty$.

while $\frac{U-L}{U} > \epsilon$

Let $\theta = \theta(c)$.

Obtain Z_θ^* by solving (\mathbf{L}_θ) .

if $Z_\theta^* > L$ then

Let $L = Z_\theta^*$.

Set $noImproveCount = 0$ and $bestUpperIter = c$.

else

Set $noImproveCount = noImproveCount + 1$.

end if

Obtain Z_f^* by solving the modified two stage bottom-up approach.

if $Z_f^* < U$

Let $U = Z_f^*$
 Set $bestLowerIter = c$.
end if
 Set $c = c + 1$.
 $\theta_\tau(c) = \max \left\{ 0, \theta_\tau(c-1) + \eta \left(\sum_{i=1}^N \sum_{s \in \mathbf{S}} c(s, a_{s\tau}^*) q_{i,s,a_{s\tau}^*,\tau} - B_\tau \right) \right\}$, where

$$\eta = \frac{\delta[U - Z_\theta^*]}{\sum_{\tau=t}^T \left[\left(\sum_{i=1}^N \sum_{s \in \mathbf{S}} c(s, a_{s\tau}^*) q_{i,s,a_{s\tau}^*,\tau} - B_\tau \right) \right]^2}$$

if $noImproveCount > countLimit$
 Set $\delta = \delta/2$.
 Set $noImproveCount = 0$.
end if
end while

In conclusion, the Lagrangian relaxation technique helps incorporate the budget constraints within the optimal/sub-optimal policies. It is an improvement over the TSBU approach, since the policies determined by the TSBU approach correspond to the $\boldsymbol{\theta} = \mathbf{0}$ scenario. A limitation of the technique is that the estimation of the expected penalised future cost-to-go also assumes that the optimal policy will be implemented in the future years, which may not be possible for all facilities. However, the choice of the optimal policy itself might be better than the one chosen by the TSBU approach.

The reader is encouraged to refer to Ohlmann and Bean (2009) for more details with regards to the implementation of the Lagrangian relaxation approach.

3.3 A Simultaneous Network Optimization Approach

Based on the discussion of the single and multidimensional MDP approaches, it can be inferred that facility-specific policies need to be developed in accordance with the resource constraints imposed on the current as well as future years. Herein, the LP-based randomized policy approach satisfies all requirements, except for providing facility-specific policies. At the same time, determining facility-specific policies alludes to an integer programming formulation. Keeping this in mind, an approach can be developed by modifying the LP formulation into a mixed-integer linear programming formulation, as shown below:

$$\begin{aligned}
 \min_{x,w} \quad & \sum_{i=1}^N \sum_{a \in \mathbf{A}} (c(\mathbf{s}_t(i), a) + u(\mathbf{s}_t(i))) x_{iat} \\
 & + N \left(\sum_{s \in \mathbf{S}} \sum_{a \in \mathbf{A}} \left(\sum_{\tau=t+1}^T \alpha^{\tau-t} (c(s, a) + u(s)) w_{sa\tau} + \alpha^{T+1-t} \tilde{V}(s) w_{sa(T+1)} \right) \right) \quad (3.26)
 \end{aligned}$$

subject to

$$\sum_{a \in \mathbf{A}} x_{iat} = 1 \quad \forall i = 1, \dots, N, \quad (3.27)$$

$$\frac{1}{N} \left(\sum_{i=1, \dots, N | \mathbf{s}_t(i)=r} x_{iat} \right) = w_{rat} \quad \forall r \in \mathbf{S}, \forall a \in \mathbf{A}, \quad (3.28)$$

$$\sum_{s \in \mathbf{S}} \sum_{a \in \mathbf{A}} w_{sa\tau} = 1 \quad \forall \tau = t+1, \dots, T, \quad (3.29)$$

$$N \left(\sum_{s \in \mathbf{S}} \sum_{a \in \mathbf{A}} c(s, a) w_{sa\tau} \right) \leq B_\tau \quad \forall \tau = t, \dots, T, \quad (3.30)$$

$$\sum_{r \in \mathbf{S}} \sum_{a \in \mathbf{A}} p_a(r, s) w_{ra\tau} = \sum_{a \in \mathbf{A}} w_{sa(\tau+1)} \quad \forall s \in \mathbf{S}, \forall \tau = t, \dots, T, \quad (3.31)$$

$$w_{sa\tau} \in [0, 1] \quad \forall s \in \mathbf{S}, \forall a \in \mathbf{A}, \forall \tau = t, \dots, T+1,$$

where,

x_{iat} : 1 if action a is selected for facility i ; 0 otherwise (t refers to the current year),

$w_{sa\tau}$: fraction of the network in state s to which action a is applied in year τ , where τ is representative of all the future years.

In terms of the objective function and the resulting optimal solution, SNO is identical to the approach provided by Golabi et al. (1982). The only modification in the problem formulation is with regards to the use of binary integer variables for the current year, t , as is evident from the objective function (equation 3.26). The constraint of interest is equation 3.28, which defines the randomized policies for the current year in terms of the integer variables. Once the relationship between the two sets of variables is established, it is then possible to determine the expected future costs in terms of the randomized policies.

The salient feature of SNO is that it provides facility-specific policies for the current year using a single optimization routine, while utilizing the randomized policies to calculate the expected future costs. This allows for budget constraints to be imposed on the future years, as well as retaining the optimal nature of the LP formulation. In comparison, TSBU is internally inconsistent, as it does not account for the system-level interdependencies at the facility-level problem. In the case of the Lagrangian relaxation-based approach, while the identification of the optimal/sub-optimal policies incorporates the budget constraints using the Lagrangian multipliers, there is scope for the penalized cost-to-go estimates to be incorrectly estimated.

The SNO framework needs to be implemented for every year of the planning horizon, since the condition state associated with each facility, $\mathbf{s}_t(i)$, needs to be identified at the beginning of each year.

3.4 Parametric Study

For evaluating the proposed methodology (SNO), a parametric study was conducted to compare its performance with TSBU and the Lagrangian relaxation-based approach. The condition state of the facilities was defined using an eight point ordinal index, where 1 is the best state and 8 is deemed to be an unacceptable state by the agency. For the purpose of illustration, four types of activities were considered: do-nothing, routine maintenance, rehabilitation and reconstruction. The agency and user cost structure, shown in table A.1, was taken from Madanat (1993). Herein, maintenance and rehabilitation activities become prohibitively more expensive as the state worsens, whereas reconstruction incurs a constant cost. The user cost also increases as the facility deteriorates, and a high penalty cost is imposed when the facility is in the non-permissible condition state ($s = 8$). The transition probability matrices for the different MR&R alternatives, as shown in tables B.1-B.4, were also adapted from Madanat (1993), but were suitably modified to reflect the increasing levels of maintenance effectiveness. The planning horizon consisted of 15 years and the discount rate was 5%. The salvage value at the end of the planning horizon was set equal to the user costs, which can be interpreted as a proxy for the quality of the terminating state of the facility.

In order to generate the different scenarios, the annual budget was fixed at $B = 250$ units, while the number of facilities, n , was varied to be 10, 100, and 500. Hence, for a network of 10 facilities, an annual budget of 250 units would be sufficiently high, whereas, for $n=500$, the same budget would be severely constraining. The initial condition of the facilities was uniformly distributed between states 1 and 7 (the non-permissible condition state 8 was excluded), so as to represent a wide range of condition states in the system.

Given the stochastic nature of deterioration, the results were generated using a Monte Carlo simulation method. Monte Carlo simulation is a popular sampling technique, wherein random information is generated using an artificial process (typically a uniform distribution), so as to pick a random observation from a population (Powell 2007). For the parametric study, the condition states for a new decision-making year were simulated based on the facility-specific actions recommended by the optimization routine for the previous year of decision-making using a uniform random number generator. The process was then repeated for each year, until the end of the planning horizon. For each scenario, 1000 simulations were carried out to determine the average system-level user-plus-agency costs incurred by the agency (in net present value).

3.4.1 Results

Figure 3.1 represents the average system-level costs incurred by the agency using SNO, the Lagrangian relaxation approach and TSBU. As the budget is kept constant, the costs for all three approaches, represented on a log-scale, increase with an increase in the number of facilities in the system. For $n=10$, all three optimization approaches perform equally well.

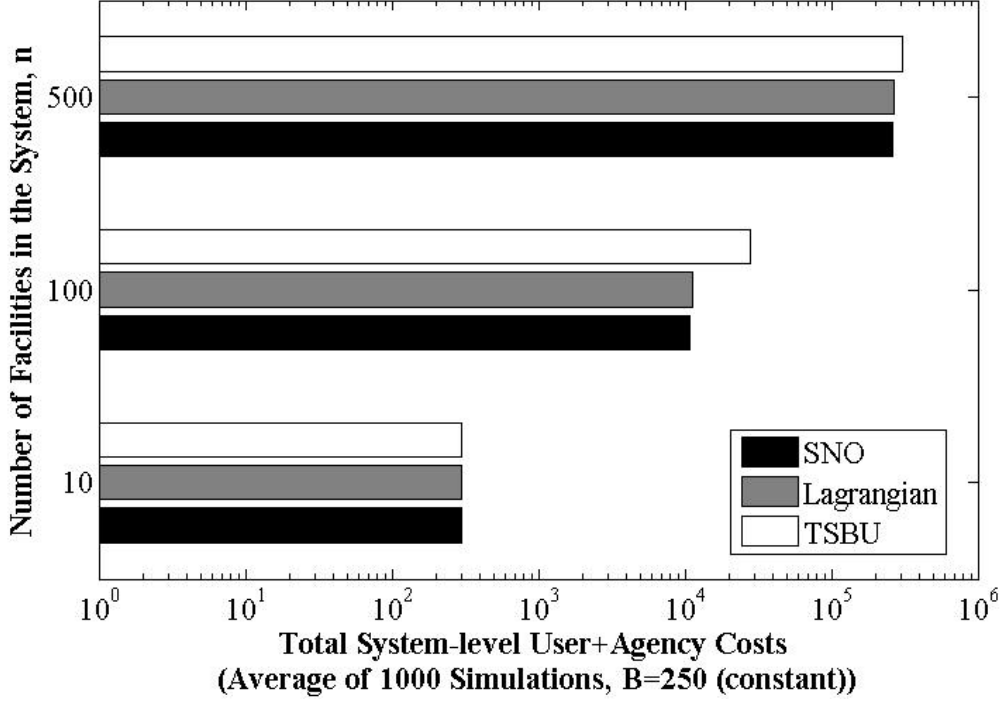


Figure 3.1: Comparison of the average system-level costs incurred by implementing SNO, Lagrangian relaxation, and TSBU

However, as the budget constraint becomes more severe, SNO and the Lagrangian relaxation approach start providing lower costs than TSBU.

A more informative assessment of the three approaches can be made by comparing the distributions of the simulation results, as illustrated in figure 3.2. Herein, a box plot representation shows the median (the horizontal line inside the box), the lower and upper quartiles (the edges of the box), and the overall spread of the simulation results (the whiskers extending above and below the box). In addition, a dot, signifying the expected system-level costs, predicted by each optimization routine at $t = 1$, has also been marked on the plot. Ideally, as the number of simulations tends to infinity, it would be anticipated that the average of the costs realized through simulation and the expected cost predicted by the optimization should become identical. Hence, the performance of the three approaches can also be evaluated on how closely the realized costs match with the a-priori expected minimum costs.

Figure 3.2(a) shows the box plot corresponding to $n = 10$ (the unconstrained budget scenario). In this case, the distribution of the simulated costs is identical for SNO, the Lagrangian relaxation approach, and TSBU. Also, the medians of the box plots coincide with the costs expected at $t = 1$, indicating that both approaches predict the future costs accurately. For $n = 100$ (figure 3.2(b)), it can be seen that while Lagrangian relaxation

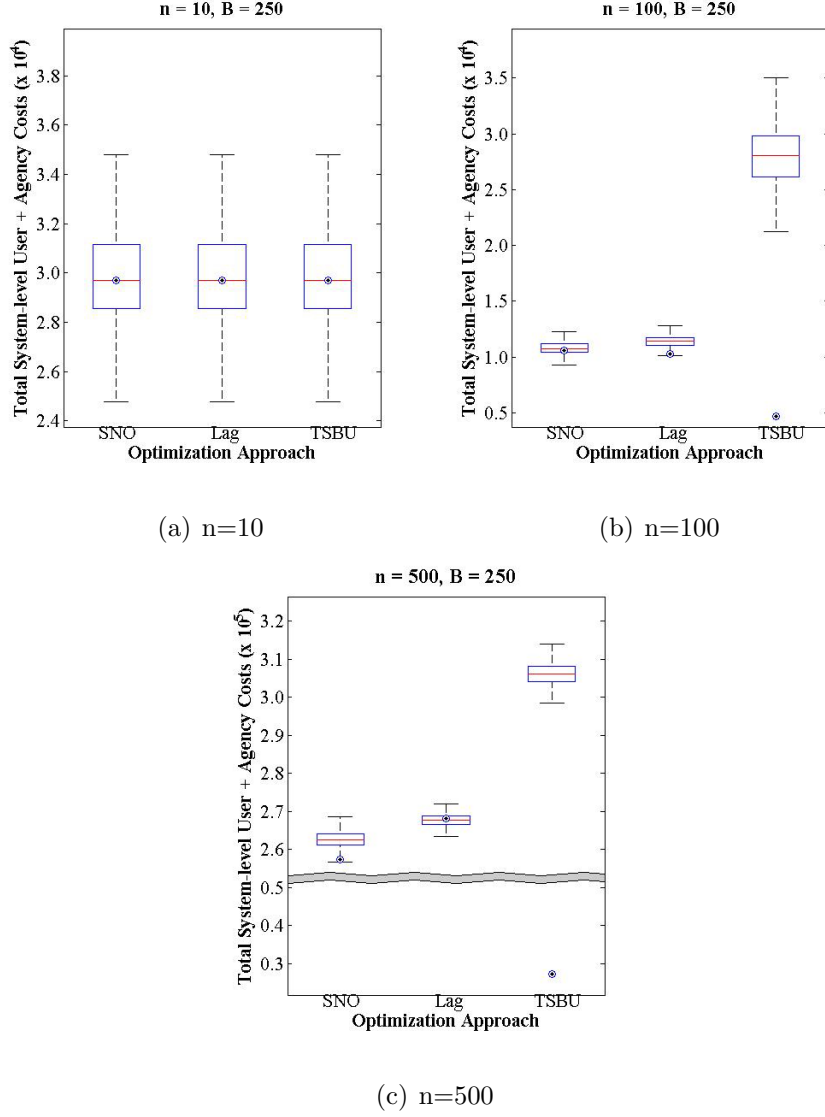


Figure 3.2: Comparison between the simulated costs and the expected system-level costs predicted at $t=1$ for SNO, Lagrangian relaxation, and TSBU

and SNO provide a-priori expected cost-to-go estimates within the range of the simulation results, differences between the predicted and the realized costs begin to emerge for the TSBU approach. In fact, it can be observed that TSBU becomes increasingly inconsistent with its predictions as the budget constraint becomes tighter (i.e., as the number of facilities increase). Consequently, it can be understood that TSBU's choice of policies becomes sub-optimal as the budget constraint becomes more restrictive. In comparison, the Lagrangian relaxation method is able to re-order the optimal/alternate policies to reflect the severity of the budget constraints. The simulation results also indicate that SNO performs marginally better than

the Lagrangian relaxation method in the presence of restrictive budget constraints. In the case of deviations of the predicted cost-to-go from the median of the box plot, the differences can be attributed to insufficient Monte Carlo simulations.

3.5 Comparison of the facility-specific policies

In order to highlight the shortcomings of the TSBU approach, a comparison between the facility-specific policies obtained from SNO and TSBU is illustrated in figure 3.3. Herein, the two plots show the evolution of the condition state of a facility starting in state 7, and the corresponding actions recommended by SNO and TSBU, respectively. The results correspond to the $n = 100$ scenario. The optimal policy recommended by TSBU for state 7 in year $t = 1$, is reconstruction (4), followed by rehabilitation (3), routine maintenance (2), and do-nothing (1). In comparison, the system-level analysis of SNO recommends the optimal actions to be distributed between do-nothing (1) and reconstruction (4). Owing to the budget constraints, TSBU implements the rehabilitation action (second optimal), while SNO implements do-nothing for the facility under consideration. As the simulation evolves over time, SNO implements a reconstruction action in year 3, whereas the TSBU policies result in the facility staying in state 8 for prolonged periods of time. The resulting user+agency cost accrued through SNO and TSBU are 178.57 and 609.42 units respectively.

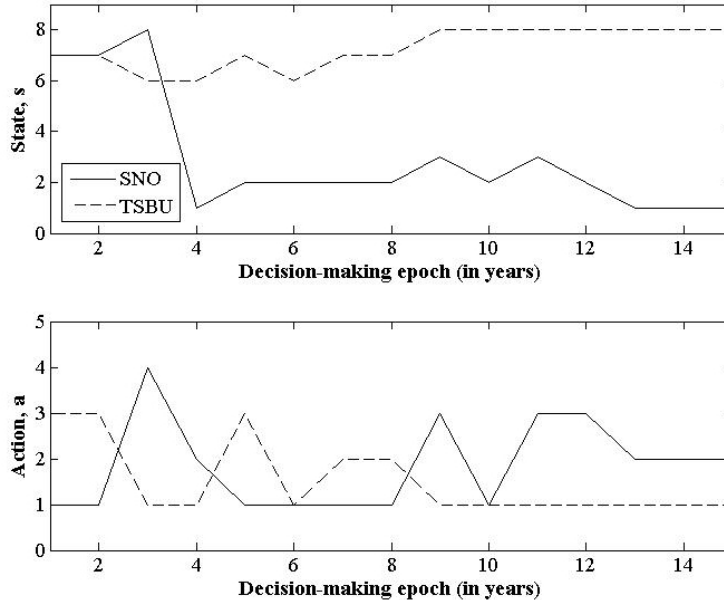


Figure 3.3: Comparison between facility-specific policies of SNO and TSBU for a facility from $n = 100$ scenario; For the action set, 1: do-nothing, 2: routine maintenance, 3: rehabilitation, and 4: reconstruction

While the above discussion corresponds to a single simulation run for a facility representing a worst-case scenario, some insights can still be obtained on the underlying decision-making process of both approaches. The policy matrix generated by TSBUS assumes that an optimal policy will be implemented in the future, which in the case of a significantly deteriorated condition state is to reconstruct. However, the available budget prevents the selection of a reconstruction action, leading to multiple implementations of the rehabilitation action. In comparison, SNO recognizes the budget constraints in the current, as well as the future years, and concludes that neither routine maintenance nor rehabilitation activities will lead to a significant improvement in the condition state of the facility in the long run. Consequently, SNO recommends either doing nothing, or reconstructing, for facilities in state 7, which ensures that over a period of time, most facilities will be reconstructed.

3.5.1 Implementation Issues

In the parametric study undertaken above, the optimization was carried out over a finite planning horizon. However, from an agency’s perspective, infrastructure assets, like pavements and bridges, may not have predefined useful lives. In such cases, a more realistic accounting practice would be to use a rolling planning horizon, wherein at every decision epoch, a new T -year planning horizon is solved for. A long enough planning horizon ensures that issues pertaining to salvage value selection become insignificant due to the discounting of future costs. In addition, steady-state policies and costs can also be incorporated into the SNO framework as a proxy for salvage values, as demonstrated in Golabi et al. (1982).

3.6 Discussion

In this chapter, the top-down approach proposed by Golabi et al. (1982) was extended using SNO which allows for facility-specific decision-making using randomized policies. The results of the parametric study indicate that SNO is as effective or better than the other MDP approaches in providing system-level MR&R policies for resource allocation problems. Finally, a Lagrangian relaxation-based approach was employed to overcome the limitations of the two-stage bottom-up approach, which in turn was shown to be sub-optimal and inconsistent for scenarios with constrained financial resources.

The contribution of SNO lies in facilitating a comparison between the single-dimensional and multi-dimensional methodologies in MDP-based MR&R decision-making frameworks. Also, by addressing the identification issues associated with randomized policies, the mixed integer programming formulation of SNO provides a potential framework for incorporating network considerations in a MDP setting.

From a methodological perspective, the results of the parametric case study also indicate that benchmarking and internal consistency checks are useful tools while evaluating a stochastic optimization approach. In the absence of any theoretical guarantee on optimality, comparing the performance of an approach with other state-of-the-art methods provides a

good measure of its efficacy. In contrast, internal consistency checks ensure that the observed costs are always consistent with the a-priori expectations, even if it is known to be sub-optimal.

In the next chapter, the focus shifts towards incorporating network considerations into system-level MR&R decision-making. In particular, the structural interdependence problem is formally introduced, wherein the scheduling MR&R activities on road segments adversely impacts the network capacity.

Chapter 4

Incorporating Network Considerations into the System-level Optimization Problem

In order to model the network dynamics within the system-level infrastructure management problem, examples of potential network-based constraint/formulation include calculating the maximum flow through the network, or solving a deterministic user equilibrium problem to determine the network travel time. Such computations require the knowledge of the effective capacities on each link on the network, which in turn are a function of the MR&R activities being implemented on the links. Consequently, it is important to identify MR&R decision-making frameworks which provides facility-specific policies.

In the previous chapter, the SNO framework was proposed, which provides optimal facility-specific policies for a given decision-making year of a resource allocation problem. However, one of its shortcomings is that the network constraints cannot be simultaneously imposed on the cost estimates associated with the future years, as the randomized policies are represented as a probabilistic distribution over the action space. Consequently, while SNO is optimal for budget allocation problems, it may provide sub-optimal, but feasible, policies upon the inclusion of network-based constraints.

In order to benchmark the solutions obtained from SNO, the use of an approximate dynamic programming (ADP) framework is motivated. The advantage of ADP is that it facilitates the modeling of complex problem formulations within an MDP setting. Also, in comparison to the other multidimensional MDP frameworks discussed in chapter 3, ADP is general modeling framework. However, a limitation of ADP is that an optimal solution cannot always be guaranteed.

The rest of the chapter is structured as follows: the need for ADP is first motivated by presenting the Bellman equation and the curse of dimensionality associated with it. Subsequently, an ADP framework is presented in a general fashion, so as to simplify the problem formulation and facilitate a holistic understanding of the technique. Finally, a parametric study is conducted on a stylized network to infer the impact of network-based constraints

on the decision-making methodologies, wherein the results obtained from ADP and SNO are compared.

4.1 Multidimensional MDP Problem

An exact solution to a finite horizon multidimensional MDP problem can be obtained by employing a backward recursive dynamic programming approach, which is based on Richard Bellman's *principle of optimality* (Bellman 1954). The optimality rule states the following: *an optimal policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decisions*. In other words, the principle alludes to the Markovian assumption, wherein the optimal policy in a given year of the planning horizon is only dependent on the current state of the system.

In the context of MR&R decision-making, the problem is formulated as follows: if the state of the pavement network at the beginning of year t is represented by \mathbf{s}_t , then the optimal cost-to-go from year t to the end of the planning horizon, T , can be recursively obtained using the following set of equations:

$$V_\tau(\mathbf{s}_\tau) = \min_{\mathbf{a}_\tau \in \mathbf{A}_c^N} \left[c(\mathbf{s}_\tau, \mathbf{a}_\tau) + \alpha \sum_{\mathbf{s}' \in \mathbf{S}^N} P(\mathbf{s}' | \mathbf{s}_\tau, \mathbf{a}_\tau) V_{\tau+1}(\mathbf{s}') \right] \quad \forall \mathbf{s}_\tau \in \mathbf{S}^N, \quad \tau = T, \dots, t, \quad (4.1)$$

$$a_\tau(\mathbf{s}_\tau) = \arg \min_{\mathbf{a}_\tau \in \mathbf{A}_c^N} \left[c(\mathbf{s}_\tau, \mathbf{a}_\tau) + \alpha \sum_{\mathbf{s}' \in \mathbf{S}^N} P(\mathbf{s}' | \mathbf{s}_\tau, \mathbf{a}_\tau) V_{\tau+1}(\mathbf{s}') \right] \quad \forall \mathbf{s}_\tau \in \mathbf{S}^N, \quad \tau = T, \dots, t, \quad (4.2)$$

$$V_{T+1}(\mathbf{s}_{T+1}) = \bar{V}(\mathbf{s}_{T+1}) \quad \forall \mathbf{s}_{T+1} \in \mathbf{S}^N, \quad (4.3)$$

where,

$V_\tau(\mathbf{s}_\tau)$: value function associated with state, \mathbf{s}_τ , in year τ , representing the minimum expected cost-to-go from year τ to the end of the planning horizon,

$a_\tau(\mathbf{s}_\tau)$: the optimal policy associated with state, \mathbf{s}_τ , in year τ ,

$\bar{V}(\mathbf{s}_{T+1})$: salvage value associated with state, \mathbf{s}_{T+1} , at the end of the planning horizon,

$c(\mathbf{s}_\tau, \mathbf{a}_\tau)$: unit costs associated with implementing the action set, \mathbf{a}_τ , when the network is in state, \mathbf{s}_τ ,

$P(\mathbf{s}'|\mathbf{s}_\tau, \mathbf{a}_\tau)$: one-year transition probability of being in state, \mathbf{s}' , in year $\tau + 1$, given that the network is in state, \mathbf{s}_τ , in year τ , and action set, \mathbf{a}_τ , is implemented,
 α : discount amount factor,
 N : number of facilities in the network,
 \mathbf{S}^N : feasible state space of the network,
 \mathbf{A}_c^N : constrained feasible action space of the network (constraints can include finite budget, scheduling, network flow, etc.).

The problem formulation are almost identical to the facility-level problem formulation in the TSBU approach, except that the state space in this formulation is defined over the network, and instead of finding facility-level policies, the intention is to find an optimal policy set which satisfies all the system-level constraints.

The advantage of a backward recursive dynamic programming approach is that it guarantees an optimal policy. However, it suffers from the *curse of dimensionality*, wherein the computational costs associated with solving the Bellman equation increase exponentially with the size of the network. A major consideration in this regard is that the optimization problem is required to be solved for all possible state space combinations, which increase exponentially with the number of facilities in the network. For instance, if $|\mathbf{S}| = 8$, the number of condition states possible for $N = 5$ are $8^5 = 32,768$; whereas for $N = 10$, the combinations rise up to $8^{10} = 1,073,741,824$. Other factors contributing towards an increase in computational complexity include, looping over the entire state space to calculate the expectation of the future cost-to-go ($\sum_{\mathbf{s}' \in \mathbf{S}^N} P(\mathbf{s}'|\mathbf{s}_\tau, \mathbf{a}_\tau) V_{\tau+1}(\mathbf{s}')$), as well as an exponential increase of the action space, \mathbf{A}_c^N . Consequently, such an approach is not suitable for solving a system-level MR&R decision-making problem.

4.2 Approximate Dynamic Programming

Approximate dynamic programming is a MDP-based modeling framework which seeks to overcome the dimensionality issues associated with traditional dynamic programming methods. It employs an algorithmic strategy of stepping forward through time, which obviates the need to loop through the entire state space in future time periods. Researchers have applied ADP techniques to diverse problems, ranging from learning how to play backgammon (Tesauro 1995) to large-scale vehicle fleet management (Simão et al. 2009). The most notable references pertaining to ADP include Bertsekas and Tsitsiklis (1996), Sutton and Barto (1998) and Powell (2007).

The term "approximate" in ADP refers to the fact that the value functions need to be estimated using simulations, and may not always converge to the optimal values. In comparison, a backward recursive dynamic programming approach guarantees optimal policies, as it exactly computes the value functions for all years. However, the advantage of ADP is that it provides a framework to solve high-dimensional MDP problems which cannot be attempted otherwise using analytical techniques. It also allows for complex network considerations to

be represented within the decision-making problem. Hence, while attempting to solve MDP problems using ADP, the inherent trade-off lies in sacrificing optimality for computational tractability.

There are many strategies available for estimating value functions and the resulting policies in the ADP literature. The details of the particular methodology undertaken in this work are outlined in the following subsections: firstly, the need for approximating the value function with lower dimensional parametric functions is motivated. Subsequently, the procedures for stepping forward through time are detailed, and finally, the learning algorithm for updating the value function approximation is discussed.

4.2.1 Value Function Approximation

In the Bellman equation, the cost-to-go function, $V_t(\mathbf{s}_t)$, is stored using a standard look-up table representation, wherein each element of the state space ($|\mathbf{S}|^N$) is uniquely identified. A look-up table representation in the context of an ADP framework has several shortcomings. Firstly, the memory requirements can be unsurmountable for high-dimensional problems. Secondly, calculating the future cost-to-go requires looping over the entire state space, which leads to computational inefficiencies. Thirdly, the learning rate associated with a look-up table representation can be very slow, since each element of the state space must be visited multiple times in order to develop a good estimate of its cost-to-go. Consequently, there is a need to approximate the value function so as to make its representation scalable from the point of view of computational efficiency as well as memory allocation.

The ADP literature recommends different techniques to work with large state spaces. One such approach relies on the use of aggregation methods to scale down the state and action spaces. An alternate approach is to develop analytical functions using parameters which exploit the structure of the problem in order to capture important attributes of the value function. In this paper, a combination of both approaches is adopted for approximating the value function. In particular, a set of linear, separable basis functions are chosen, which can be summed up to provide an estimate of the future cost-to-go for a given state-action pair, $(\mathbf{s}_t, \mathbf{a}_t)$, in year t :

$$\tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t) = \sum_{b=1}^{|\mathbf{B}_t|} \theta_b(\mathbf{s}_t, \mathbf{a}_t) \phi_b(\mathbf{s}_t, \mathbf{a}_t), \quad (4.4)$$

where,

$\tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t)$: expected cost-to-go from year $t + 1$ to the end of the planning horizon, when action set, \mathbf{a}_t , is chosen for state, \mathbf{s}_t ; also referred to as the Q -function,

$\phi_b(\mathbf{s}_t, \mathbf{a}_t)$: user-specified basis function which captures specific attributes of the cost-to-go function; also referred to as the Q -factor,

$\theta_b(\mathbf{s}_t, \mathbf{a}_t)$: weight associated with a given basis function, which is iteratively updated within the ADP framework,

$|\mathbf{B}_t|$: the total number of basis functions for time period, t .

Using $\tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t)$, system-level MR&R policies can be obtained as follows:

$$\min_{\mathbf{a}_t \in \mathbf{A}_t^N} c(\mathbf{s}_t, \mathbf{a}_t) + \alpha \tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t). \quad (4.5)$$

Herein, it is noteworthy to distinguish $\tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t)$ from $V_t(\mathbf{s}_t)$. $V_t(\mathbf{s}_t)$ represents the optimal expected cost-to-go associated with state, \mathbf{s}_t , from year t to the end of the planning horizon. A limitation of approximating the optimal cost-to-go is that it still necessitates a calculation of the expected future cost-to-go for a given action set, \mathbf{a}_t . In comparison, the Q -function can be viewed as approximating the expected future cost-to-go using an augmented state, $(\mathbf{s}_t, \mathbf{a}_t)$, which leads to computational savings. The equivalence between $\tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t)$ and $V_t(\mathbf{s}_t)$ can be established by comparing equations 4.5 and 4.1:

$$\tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t) \approx \sum_{\mathbf{s}' \in \mathbf{S}} P(\mathbf{s}' | \mathbf{s}_t, \mathbf{a}_t) V_{t+1}(\mathbf{s}'). \quad (4.6)$$

It is important to note that $\tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t)$ excludes the costs incurred in year t , since the corresponding costs can be easily calculated using $c(\mathbf{s}_t, \mathbf{a}_t)$.

4.2.2 Stepping forward through time

For a given ADP iteration, n , let the current state of the network be \mathbf{s}_t^n , and the optimal action set selected on the basis of equation 4.5 be \mathbf{a}_t^n . Once \mathbf{a}_t^n has been identified, the process of stepping forward in time is undertaken by generating a sample realization of the future state of the system, \mathbf{s}_{t+1}^n , using a Monte Carlo simulation procedure. As discussed in the parametric case study of chapter 3, the procedure involves using a uniform random number generator to yield a feasible state transition from the transition probability distribution, $P(\mathbf{s}_{t+1}^n | \mathbf{s}_t^n, \mathbf{a}_t^n)$. Let the transition procedure be symbolically represented as:

$$\mathbf{s}_{t+1}^n = \Psi(\mathbf{s}_t^n, \mathbf{a}_t^n, \omega^n), \quad (4.7)$$

wherein, ω^n represents a vector of uniform random numbers generated from the distribution, $U(0,1)$.

Once a sample state of the system for year $t + 1$ is ascertained, the optimal action set for year $t + 1$, \mathbf{a}_{t+1}^n , can be solved for by using equation 4.5. In order to use equation 4.5, it is assumed that the current estimate of the Q -function, $\tilde{Q}_{t+1}(\mathbf{s}_{t+1}^n, \mathbf{a}_{t+1}^n)$ is the best available estimate of the future cost-to-go.

The procedure is subsequently repeated until the end of the planning horizon, and a sequence of state-action pairs, $[(\mathbf{s}_1^n, \mathbf{a}_1^n), (\mathbf{s}_2^n, \mathbf{a}_2^n), \dots, (\mathbf{s}_T^n, \mathbf{a}_T^n), \mathbf{s}_{T+1}^n]$, is obtained. The sequence is referred to as a *sample path*, since it is a simulated realization of a stochastic process.

4.2.3 Updating the value function approximation

The concept of a sample path is fundamental to the idea of learning in approximate dynamic programming (Powell 2007). Its importance lies in the fact that the costs accrued over a sample path can be compared with the costs predicted by the latest estimate of the future cost-to-go, $\tilde{Q}_t^{n-1}(\mathbf{s}_t^n, \mathbf{a}_t^n)$. If the realized costs are very different from the costs predicted by the Q -function, it implies that the current parameter estimates are poor, thus allowing for the weights associated with the Q -factors to be updated. Conversely, if the predicted and realized costs are similar, it implies that the Q -function provides an accurate representation of the future costs.

In this paper, a *temporal difference* (TD) learning algorithm is employed for updating the value function approximation. TD learning is a two-stage procedure which relies on a forward pass to generate a sample path, and a backward pass to update the parameters. More importantly, it generalizes the different types of estimation errors which can be obtained from a given sample path.

For instance, a one-period look-ahead temporal difference error compares the Q -function estimates of the states visited in two successive years of a sample path, as shown below:

$$\delta_t^n = c(\mathbf{s}_{t+1}^n, \mathbf{a}_{t+1}^n) + \alpha \tilde{Q}_{t+1}^{n-1}(\mathbf{s}_{t+1}^n, \mathbf{a}_{t+1}^n) - \tilde{Q}_t^{n-1}(\mathbf{s}_t^n, \mathbf{a}_t^n), \quad (4.8)$$

The one-period look-ahead update can also be classified as a *bootstrapping* approach, wherein a value estimate is updated based on an existing value estimate (Sutton and Barto 1998). Dynamic programming is another example of bootstrapping, since the cost-to-go in year t is estimated using the cost-to-go estimate corresponding to year $t + 1$. However, a limitation of the one-period look-ahead update is that, unlike dynamic programming, if the initial estimates of the value function approximation in year $t + 1$ are poor, the update may lead to a bias in the future cost estimation.

An alternate approach for updating the Q -factors is to compare $\tilde{Q}_t^n(\mathbf{s}_t^n, \mathbf{a}_t^n)$ with the costs realized through the sample path, which can be represented as follows:

$$\tilde{z}_t^n = c(\mathbf{s}_{t+1}^n, \mathbf{a}_{t+1}^n) + \alpha c(\mathbf{s}_{t+2}^n, \mathbf{a}_{t+2}^n) + \dots + \alpha^{T-t+1} c(\mathbf{s}_{T+1}^n, \mathbf{a}_{T+1}^n). \quad (4.9)$$

Unlike bootstrapping, \tilde{z}_t^n provides a cost-to-go estimate that is purely based on the Monte Carlo simulation of the sample path. Consequently, \tilde{z}_t^n as an estimate is representative of the costs associated with the current state of the value function approximations. However, a cost-to-go estimate based on the entire sample path does not account for the possibility that the difference between the predicted and realized cost-to-go could have been generated by the sub-optimal nature of the policies belonging to the intermediate years. In such a case, the Q -factor update of the current policy, \mathbf{a}_t^n , may suffer from an over-fitting problem.

The TD learning algorithm generalizes the bootstrapping and Monte Carlo-based learning approaches, shown in equations 4.8 and 4.9, as follows:

$$\Delta_t^n = \sum_{\tau=t}^T (\alpha\lambda)^{\tau-t} \left(c(\mathbf{s}_{\tau+1}^n, \mathbf{a}_{\tau+1}^n) + \alpha \tilde{Q}_{\tau+1}^{n-1}(\mathbf{s}_{\tau+1}^n, \mathbf{a}_{\tau+1}^n) - \tilde{Q}_{\tau}^{n-1}(\mathbf{s}_{\tau}^n, \mathbf{a}_{\tau}^n) \right), \quad (4.10)$$

$$= \sum_{\tau=t}^T (\alpha\lambda)^{\tau-t} (\delta_{\tau}^n), \quad (4.11)$$

where,

Δ_t^n : temporal difference error associated with the current estimate of the future cost-to-go, $\tilde{Q}_t^{n-1}(\mathbf{s}_t^n, \mathbf{a}_t^n)$,

$\lambda \in [0, 1]$: heuristic discount amount factor which determines the contribution of the errors associated with policies implemented farther down the sample path.

The learning framework presented in equations 4.10 and 4.11 is referred to as TD(λ) learning, wherein the parameter, λ , generalizes the learning process. It can be shown that TD(0) is equivalent to equation 4.8, whereas TD(1) reduces the problem to estimating the temporal difference error using the costs accrued over the sample path (equation 4.9). More generally, the artificial discounting introduced by the parameter, λ , allows the algorithm to look into the future, while accounting for the possibility that the policies implemented further along the sample path may be sub-optimal.

Once the temporal difference error is calculated for a given year t , the weights associated with the Q -factors can be updated using a stochastic gradient algorithm:

$$\theta_{tb}^n \longleftarrow \theta_{tb}^{n-1} + \gamma_n \nabla_{\theta_{tb}} (\hat{v}_t^n) \Delta_t. \quad (4.12)$$

where,

$$\hat{v}_t^n = c(\mathbf{s}_t^n, \mathbf{a}_t^n) + \alpha \tilde{Q}_t^{n-1}(\mathbf{s}_t^n, \mathbf{a}_t^n), \quad (4.13)$$

and γ_n represents the step-size of the stochastic gradient algorithm. The step-size determines the magnitude of the update made along the direction of the error-minimizing gradient.

There exist various step-size rules in the stochastic gradient algorithm literature which can be used in order to achieve convergence of the parameters being updated. However, the three basic conditions are as follows:

$$\gamma_n \geq 0, \quad n = 1, 2, \dots, \quad (4.14)$$

$$\sum_{n=1}^{\infty} \gamma_n = \infty, \quad (4.15)$$

$$\sum_{n=1}^{infy} \gamma_n^2 < \infty. \quad (4.16)$$

In this paper, a *search-then-converge* step-size rule is adopted, which produces an extended period of learning. The step-size formula is represented as follows:

$$\gamma_n = \gamma_0 \frac{\left(\frac{b}{n} + a\right)}{\left(\frac{b}{n} + a + n^\beta\right)}, \quad (4.17)$$

where, γ_0, a, b , and β are parameters to be determined.

This class of step-size rules is termed "search-then-converge" because they provide for a period of high step-sizes (while the search for optimal policies is taking place) after which the step-size declines (to achieve convergence). The degree of delayed learning is controlled by the parameters, b and a (Powell 2007).

In conclusion, the proposed ADP framework for solving a system-level MR&R decision-making problem can be summarized using the following pseudo-code:

Step 0. Initialization:

Step 0a. Initialize $\tilde{Q}_t^0(\mathbf{s}_t, \mathbf{a}_t)$, $\forall \mathbf{s}_t \in \mathbf{S}^N, \forall \mathbf{a}_t \in \mathbf{A}^N, \forall t = 1, \dots, T + 1$.

Step 0b. Choose an initial state \mathbf{s}_1^0 .

Step 1. Do for $n = 1, 2, \dots, \mathcal{N}$:

Step 2. Do for $t = 1, 2, \dots, T$:

Step 2a. Solve:

$$\hat{v}_t^n = \min_{\mathbf{a}_t \in \mathbf{A}_c^N} c(\mathbf{s}_t^n, \mathbf{a}_t) + \alpha \tilde{Q}_t^{n-1}(\mathbf{s}_t^n, \mathbf{a}_t),$$

and let \mathbf{a}_t^n be the optimal solution corresponding to \hat{v}_t^n .

Step 2b. Compute $\mathbf{s}_{t+1}^n \leftarrow \Psi(\mathbf{s}_t^n, \mathbf{a}_t^n, \omega^n)$

Step 3. Initialize $\Delta_T \leftarrow 0$, $\hat{v}_{T+1}^n \leftarrow \bar{V}(\mathbf{s}_{T+1}^n)$, wherein, $\bar{V}(\mathbf{s}_{T+1}^n)$ represents the salvage value associated with the end of the planning horizon.

Step 4. For $t = T, T - 1, \dots, 1$, do:

Step 4a. $\delta_t \leftarrow \hat{v}_{t+1}^n - \tilde{Q}_t^{n-1}(\mathbf{s}_t^n, \mathbf{a}_t^n)$.

Step 4b. $\Delta_t \leftarrow \Delta_t + \delta_t$.

Step 4c. $\theta_{tb}^n \leftarrow \theta_{tb}^{n-1} + \gamma_n \nabla_{\theta_{tb}}(\hat{v}_t^n) \Delta_t$.

Step 4d. $\Delta_{t-1} \leftarrow \alpha \lambda \Delta_t$.

Herein, \mathcal{N} represents the duration of the training period in which the Q -factor weights are updated. Once the Q -factors have converged to their final values, the facility-specific MR&R policies can be obtained using equation 4.5.

4.2.4 Exploration vs. Exploitation

A fundamental challenge within approximate dynamic programming is that in order to estimate the cost-to-go for a given state-action pair, the state-action combination must be visited as part of the sample path. As a result, a trade-off must be made between visiting a state which leads to the lowest costs ("exploitation") versus visiting a state in order to obtain information about the value of being in that state ("exploration") (Powell 2007). Prior research indicates that allowing for intermittent exploration during the training period provides better results than a pure exploitation strategy. In this paper, a ϵ -greedy policy is employed, wherein a sub-optimal policy is randomly selected with a probability, ϵ .

4.2.5 Comparing SNO with ADP

The advantage of ADP is that the optimal actions selected at every time step of the sample path always satisfy the underlying constraints of the decision-making process. Consequently, the value function approximation updated using the sample path is also consistent with the constraints of the system-level MDP problem. Also, since ADP is modeled as a multi-dimensional MDP problem, it is capable of handling the heterogeneity among individual facilities in the network. However, owing to its simulation-based approach, ADP does not provide any optimality guarantees. In fact, its performance depends on several factors: choice of basis functions, learning strategy, convergence rate, among others.

In comparison, SNO is an optimal approach for budget allocation problems. Its limitation arises from the fact in spite of providing facility-specific policies in the first year, it essentially solves a single-dimensional MDP problem. Consequently, the use of randomized policies, while suitable for representing the average resource consumption in future time periods, is not effective in accounting for the network considerations.

4.3 Parametric Case Study

In this section, the network considerations are formally introduced into the MR&R decision-making framework as mathematical constraints. The structural interdependence between the individual facilities is incorporated by imposing a network capacity constraint on the MR&R activity selection. In particular, an agency-chosen network capacity threshold limits the loss in network capacity due to the implementation of the chosen MR&R policies by an upper bound. The framework also allows for different work zone options to be considered (partial vs. complete closure), wherein time, money and loss in link capacity can be traded-off.

The use of capacity as a performance measure can be interpreted as a supply-based criterion, wherein the agency seeks to provide enough capacity during MR&R activity implementation so that the associated origin-destination demand can be met. While it is recognized that a capacity-based approach does not take into account the demand on individual links, it can be argued that the traffic can be re-routed to maximize capacity utilization, using adequate signages and real-time information dissemination systems.

4.3.1 Network Representation

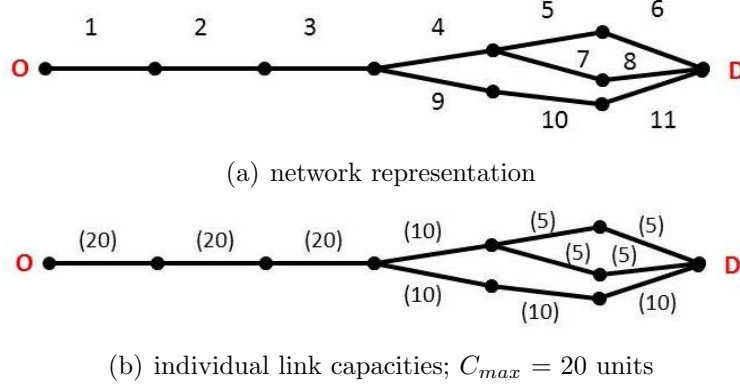


Figure 4.1: Stylized 11-link pavement network for incorporating structural interdependence (individual link capacities expressed in brackets)

Let us consider a stylized network configuration consisting of 11 road segments and 10 nodes connecting an origin and a destination, as shown in figures 4.1(a) and 4.1(b). Figure 4.1(b) represents the individual link capacities, using which the resulting network capacity, C_{max} , can be obtained as 20 units. It is to be noted that C_{max} represents the maximum capacity available to the users in the absence of MR&R activities.

In terms of notation, the pavement network is defined as a graph, $G = (N, E)$, wherein the edges, E , represent the road segments, and the nodes, N , representing the points of intersection between any two facilities. An individual road segment is defined as a link, (q, r) .

4.3.2 Value Function Approximation: ADP

In order to implement an ADP methodology, the foremost step is to model the value function approximation for the network under investigation. In this regard, the choice of basis functions plays an important role in determining the quality of the solutions obtained using ADP. The objective is to exploit the properties of the problem such that the learning process is reduced to estimating a few key parameters of interest.

Based on the network configuration depicted in figures 4.1(a) and 4.1(b), it is possible to identify groups of road segments which respond similarly to the network capacity constraints. For instance, it can be argued that the groups defined as, $\{1, 2, 3\}$, $\{4\}$, $\{5, 6, 7, 8\}$ and $\{9, 10, 11\}$ have a near-identical policy response towards the network-based constraints. In the case of links, 1-3 and 9-11, since all the road segments are in series, a MR&R activity/construction type chosen for any segment within the group leads to the same loss in capacity of the network. For sections 5-8, the symmetry in the network configuration is also taken into consideration while aggregating the corresponding state-action space. The

road segment, $\{4\}$, is considered as a singleton for the purpose of approximating the value function.

Finally, given the Markovian evolution of the system, it can also be assumed that for any given year, t , the optimal policy for a facility, (q, r) is only a function of its current state, and the action-construction type pair under consideration. Consequently, the total number of basis functions chosen for the value function approximation, $|\mathbf{B}|$, can be calculated as follows:

$$|\mathbf{B}| = T|G_N||\mathbf{S}||\mathbf{A}||L|, \quad (4.18)$$

where,

T : length of the planning horizon,

$|G_N|$: number of groups having a homogeneous response to the capacity constraints,

$|L|$: number of construction options available (partial and complete closures).

The savings from using a value function approximation, in terms of memory requirement, can be gauged from the fact that for a look-up table representation (equation 4.1), the parameter state space comprises of $T|\mathbf{S}|^N|\mathbf{A}|^N|L|^N$ combinations.

Using this Q -function representation, the future cost-to-go associated with implementing the action-construction type combinations, $(\mathbf{a}_t, \mathbf{l}_t)$, when the network is in state, \mathbf{s}_t , can be obtained as follows:

$$\tilde{Q}_t(\mathbf{s}_t, \mathbf{a}_t, \mathbf{l}_t) = \sum_{(q,r) \in E} \theta_{t, G_N(q,r), \mathbf{s}_t(q,r), \mathbf{a}_t(q,r), \mathbf{l}_t(q,r)}, \quad (4.19)$$

where, $G_N(q, r)$ refers to the group number associated with facility, (q, r) .

It is important to acknowledge here that approximating the value function is not a precise science. For instance, a large state space adversely affects the convergence of ADP since it requires more parameters to be learnt. In addition, there is also a greater need for exploration to adequately cover the parameter space. Conversely, having too few parameters may lead to a poor approximation of the future cost-to-go, leading to inefficient solution. The approach involves having some a-priori expectations about the decision-making process, followed by modeling those beliefs using the approximation function, and evaluating its performance using simulations. Based on the quality of the simulation results, further modifications can be made to the chosen basis functions, before arriving at a suitable value function approximation.

4.3.3 Problem Formulation: ADP

The objective of the MR&R optimization is to provide optimal policies for each facility, (q, r) , in the network, while satisfying the budget restrictions and the network capacity constraints:

$$\min_{x, \tilde{x}, \tilde{z}, y} \sum_{(q,r) \in E} \sum_{a \in \mathbf{A}} \sum_{l \in \mathbf{L}} (c(\mathbf{s}_t(q, r), a, l) + u(\mathbf{s}_t(q, r))) + \alpha \theta_{t, G_N(q, r), \mathbf{s}_t(q, r), a, l} x_{(q, r), a, l, t}, \quad (4.20)$$

subject to

$$\sum_{(q,r) \in E} \sum_{a \in \mathbf{A}} \sum_{l \in \mathbf{L}} c(\mathbf{s}_t(q, r), a, l) x_{(q, r), a, l, t} \leq B \quad (4.21)$$

$$\sum_{a \in \mathbf{A}} \sum_{l \in \mathbf{L}} x_{(q, r), a, l, t} = 1 \quad \forall (q, r) \in E, \quad (4.22)$$

$$\sum_{\Delta=1}^{52} \tilde{x}_{(q, r), a, l, \Delta} \geq x_{(q, r), a, l, t} \quad \forall a \in \mathbf{A}, \forall l \in \mathbf{L}, \quad (4.23)$$

$$\sum_{\Delta - d_{a, l} < \Delta' \leq \Delta} \tilde{x}_{(q, r), a, l, \Delta'} \leq \tilde{z}_{(q, r), a, l, \Delta} \quad \forall a \in \mathbf{A}, \forall l \in \mathbf{L}, \quad (4.24)$$

$$y_{(D, O), \Delta} \geq hC_{max} \quad \forall \Delta = 1, \dots, 52, \quad (4.25)$$

$$C_{(q, r)} \left[1 - \left(\sum_{a \in \mathbf{A}} \sum_{l \in \mathbf{L}} \kappa_{(q, r), a, l} \tilde{z}_{(q, r), a, l, \Delta} \right) \right] \geq y_{(q, r), \Delta} \quad \forall (q, r) \in E, \quad (4.26)$$

$$\sum_{r: (r, q) \in E} y_{(r, q), \Delta} - \sum_{r: (q, r) \in E} y_{(q, r), \Delta} = 0 \quad \forall (q, r) \in E, \quad (4.27)$$

$$x_{(q, r), a, l, t}, \tilde{x}_{(q, r), a, l, \Delta}, \tilde{z}_{(q, r), a, l, \Delta} \in \{0, 1\}, y_{(q, r), \Delta} \in \mathbf{R}^+ \quad \forall a \in \mathbf{A}, \forall l \in \mathbf{L}, \quad (4.27)$$

where,

- $x_{(q, r), a, l, t}$: 1 if the MR&R action-construction type pair (a, l) , is selected for facility (q, r) in year t ; 0 otherwise,
- $\tilde{x}_{(q, r), a, l, \Delta}$: 1 if the action-construction type pair, (a, l) , is started in week Δ for facility (q, r) ; 0 otherwise,
- $\tilde{z}_{(q, r), a, l, \Delta}$: 1 if the action-construction type pair, (a, l) , is under implementation in week Δ for facility (q, r) ; 0 otherwise,

- $y_{(D,O),\Delta}$: available network capacity in the presence of MR&R activities in week Δ , represented as a virtual edge connecting destination D and origin O,
- $y_{(q,r),\Delta}$: available link capacity for facility (q, r) in week Δ ,
- $c(\mathbf{s}_t(q, r), a, l)$: agency cost of undertaking the action-construction type pair, (a, l) ,
- $u(\mathbf{s}_t(q, r))$: user costs linked with increase in vehicle wear-and-tear, fuel usage, etc.,
- B : annual budget,
- $d_{(q,r),a,l}$: duration of implementing the action-construction type pair, (a, l) ,
- h : fraction of the maximum capacity of the network, C_{max} , representing a minimum network capacity threshold,
- $C_{(q,r)}$: maximum link capacity associated with facility (q, r) , as defined when no MR&R activity is scheduled,
- $\kappa_{(q,r),a,l}$: loss in capacity associated with implementing the action-construction type pair, (a, l) , for facility (q, r) .

Herein, equation 4.20 represents the objective function, defined as minimizing the expected system-level cost-to-go as based on the estimates of the Q -function; equation 4.21 indicates that the total amount spent on MR&R activities should be within the annual budget; equation 4.22 assigns exactly one action-construction type pair (including do-nothing) to each facility. The scheduling constraints, as represented by equations 4.23 - 4.24, are modeled as a non-preemptive scheduling problem (Sousa and Wolsey 1992), wherein the assigned activity is completed in one sequence; equation 4.23 ensures that only the chosen action-construction type pair is considered for assessing the feasibility of the MR&R scheduling, and equation 4.24 ensures that the chosen action-construction type pair undergoes a continuous construction period of $d_{(q,r),a,l}$ weeks. Equation 4.25 guarantees that the network capacity in the presence of MR&R activities does not violate the minimum network capacity threshold; equation 4.26 represents the loss in link capacity associated with implementing an action-construction type pair on facility (q, r) ; and equation 4.27 represents the flow conservation equation for every node in the network.

It is important to note that the optimization routine only imposes the constraints on the policies associated with the current year of decision-making. However, since the sample path is generated by solving the optimization problem in each time period, the underlying budget and network considerations are always satisfied along the sample path. Consequently, the resulting Q -function updates, which are also based on the sample path traversed, are consistent with the constraints of the problem.

4.3.4 Problem Formulation: SNO

In section 3.3, the SNO framework was described in the context of budget allocation problems. In order to add the network-based constraints into the formulation, the budget allocation problem can be reformulated similar to the ADP formulation using a facility-specific decision variable identical to the ADP formulation, $x_{(g,r),a,l,t}$. Hence, the network constraints (equations 4.22 - 4.27) can be modeled in an identical fashion.

The policies obtained using SNO also satisfy the network constraints for the current year of the decision-making, t . However, the disadvantage of SNO is that its future cost estimation relies on the use of randomized policies which cannot accommodate the network constraints. Consequently, while SNO still provides feasible policies, the optimality of the solutions is no longer guaranteed.

4.3.5 Scenario Generation: Supplementary Information

The condition state of the facilities is evaluated using an eight point ordinal index similar to the case study in chapter 3, where 1 is the best state and 8 is deemed to be an unacceptable state by the agency. The four types of activities under consideration include to do nothing, routine maintenance, rehabilitation and reconstruction. It is assumed that maintenance activities can be implemented overnight and hence lead to no loss in capacity of a road segment. There exist two reconstruction options: a partial road closure which requires 10 weeks of construction time and causes 30% loss in capacity, while a complete road closure can be completed in 2 weeks, but leads to a 100% loss in capacity, i.e., the link is rendered inaccessible. These values are not based on empirical data, but are representative of the kind of trade-offs which can be expected in real-life scenarios.

The agency and user cost structure are identical to the values shown in table A.1. The planning horizon consists of 15 years and the discount rate is chosen to be 5%. The salvage value at the end of the planning horizon is set equal to the user costs, wherein the user costs can be interpreted as a proxy for the quality of the terminating state of the facility.

The parameters corresponding to ADP's step-size rule are $a = 300$; $b = 0.5$; $\beta = 0.2$; $\gamma_0 = 0.7$; . The artificial discounting factor, λ , is taken to be 0.4, and the Q -function is trained for 1500 iterations. The algorithms are programmed in C++, and the optimization problems are solved using CPLEX® on a Windows-based OS with a 3.10 GHz processor and 4GB RAM.

In order to compare ADP and SNO, three budget levels are considered, $B = 50; 100; 150$ units. The network capacity threshold, h , is chosen to be 0.75, such that the network constraints are active for the purpose of the investigation. The initial condition of the facilities is uniformly distributed between the states, 1(good), 4(moderate), and 7(poor), within each group, g , so as to capture a range of deterioration levels. The performance of the two methodologies is compared on the basis of implementing each scenario 1000 times using Monte Carlo simulations.

4.3.6 Results

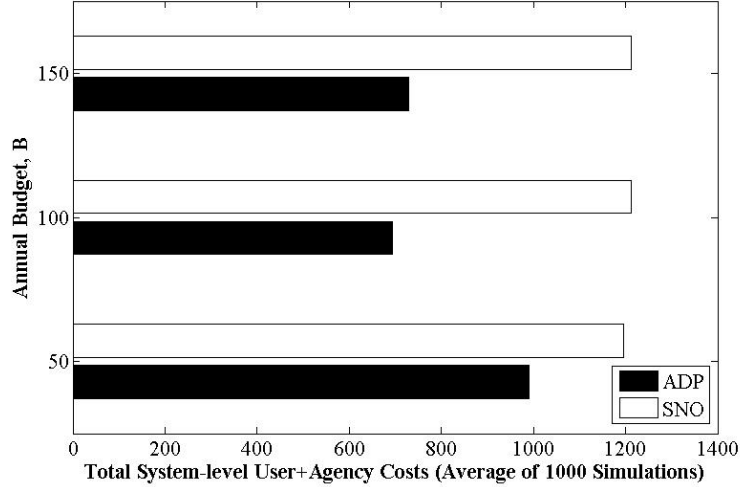


Figure 4.2: Total costs comparison between SNO and ADP for different budget levels ($h=0.75$)

Figure 4.2 represents the simulation-average of the system-level user+ agency costs incurred from using SNO and ADP. The results indicate that in the presence of network constraints, ADP performs better than SNO (on an average) across all budget levels.

Using the box plot representation, figures 4.3(a) to 4.3(c) indicate that ADP's predicted costs align well with the median of the costs realized by the simulation. On the other hand, SNO provides an inaccurate prediction of future costs as the budget increases. The similarity of the cost distributions in figures 4.3(b) and 4.3(c) indicates that the network constraints prevents any significant improvements that can be achieved by an increase in the budget levels.

The primary reason for the disparity in SNO's simulated and predicted costs is that the randomized policies corresponding to the future years do not account for the network constraints. In the context of the numerical example, reconstruction activities are required to be excluded for certain road segments due to capacity considerations. However, since the network constraints can only be imposed on the facility-specific MR&R policy selection, SNO defers the activities to the future years, in the form of randomized policies. As the sample path is simulated, the policies predicted for the future years are not realized, and the gap between the predicted and the realized costs widens. With an increase in the available budget, a greater share of the randomized policies are allocated for reconstruction activities, thus leading to more inaccuracies in the estimation of the future costs. In comparison, ADP updates its value function approximation based on states/actions that are in agreement with the network and budget constraints. Consequently, the costs predicted by ADP at $t = 1$ are consistent with the realized cost distribution.

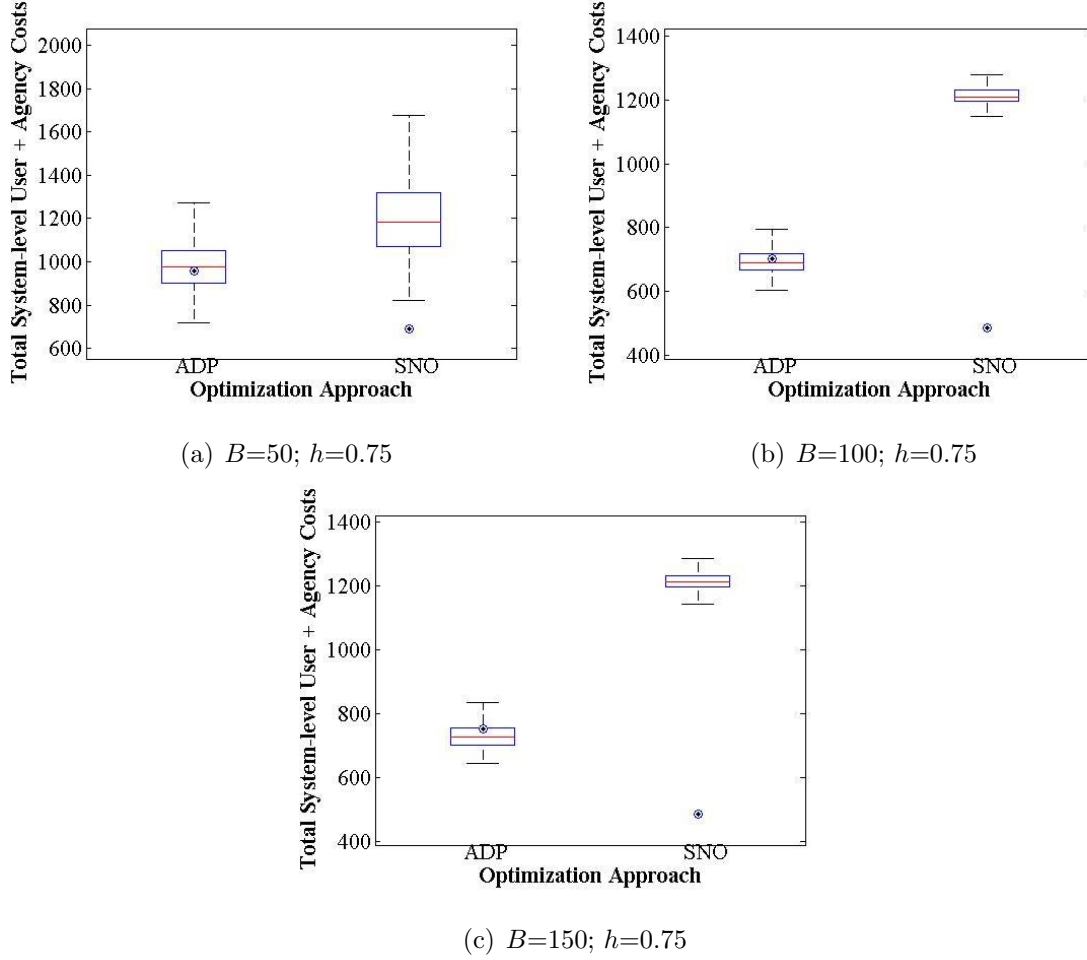


Figure 4.3: Box plots comparing ADP and SNO for varying budget levels and $h = 0.75$

Figure 4.4 shows the convergence of the Q -function for $B = 100$ scenario. As can be observed, the realized and the predicted costs converge to their final values after 500-600 iterations, which implies that the $TD(\lambda)$ learning algorithm is suitable for updating the Q -function.

Finally, in order to assess the efficacy of ADP in a pure budget allocation setting, the network capacity threshold can be fixed to $h = 0$, thereby implying that the network capacity does not constrain the decision-making. The results corresponding to the $B = 50$ and $B = 100$ scenarios are shown in figures 4.5(a) and 4.5(b). These results illustrate that, in the absence of the network constraints, SNO remains the preferred approach for pure budget allocation problems. On the other hand, while ADP does not perform as well as SNO, it still provides internally consistent results, in that the predicted and realized costs are similar in nature.

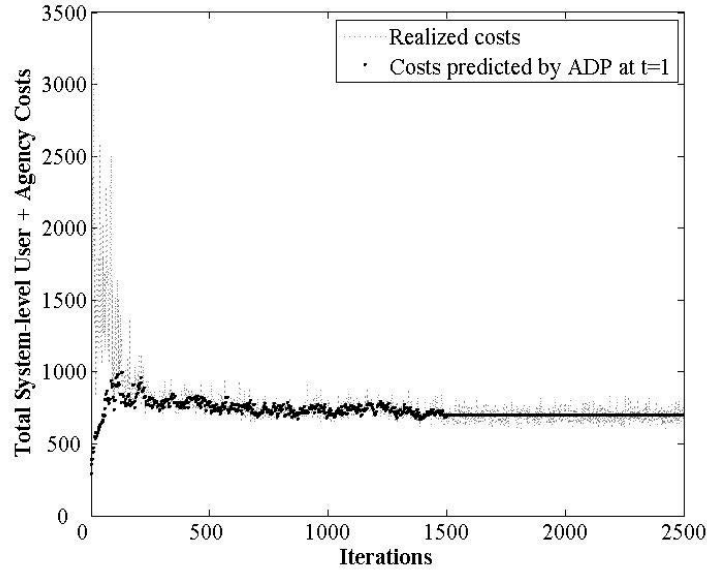


Figure 4.4: Convergence of the value function approximation for $B = 100$ scenario

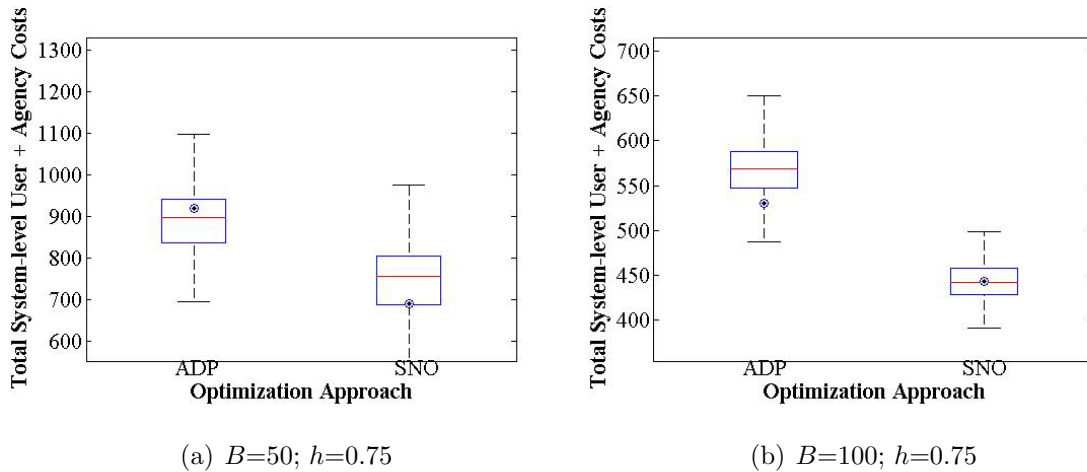


Figure 4.5: Box plots comparing ADP and SNO in the absence of network constraints

4.4 Comparison between ADP and other Network-based MR&R Decision-making Frameworks: A Qualitative Discussion

As illustrated by the results of the parametric study, ADP provides the necessary computational flexibility to solve multidimensional MDPs involving network considerations. However,

given the emerging interest in incorporating network-induced facility interdependence into MR&R decision-making, it is beneficial to qualitatively discuss the differences between the proposed approach and the other network-based decision-making frameworks in the literature.

The use of a minimum network capacity threshold implies that the agency seeks to maintain a minimum network performance guarantee in the presence of MR&R activities. It inherently assumes that the traffic can be rerouted to maximize the use of the available capacity. However, such an approach has certain shortcomings. A capacity-based performance metric does not account for the demand on the individual links, which in turn may depend on the effective link capacity, route travel time, existing condition of the road segment, etc. Other researchers have addressed the issue of demand responsiveness in different ways. Ouyang (2007) modeled the traffic distribution using a route choice model based on the users' travel time and the current state of the facility. Durango-Cohen and Sarutipand (2009) adjusted the demand response on an individual link as a linear function of its historical demand, its current condition state and effective capacity, as well as the condition state and capacity of the other links in the network. Ng et al. (2009) and Chu and Chen (2012) formulated bi-level optimization approaches, wherein the MR&R activity selection takes place in the first level while changes in the network travel time are computed in the second level.

The advantage of using ADP is that a variety of modeling assumptions can be incorporated within the decision-making process. For instance, a MDP-based framework can be developed involving a traffic assignment/simulation-based approach, wherein the impact of the MR&R activities on the network travel time will be learnt as part of the future cost estimation process. However, adequate consideration needs to be given to the ease of solving the underlying single-period optimization.

In a related issue, the concept of network capacity does not scale very well to multiple origin-destination (OD) pairs, since it is required to be solved as a multi-commodity network flow problem. However, in the special case of a multiple origin-single destination or single origin-multiple destination network, the concept of network capacity thresholds can be easily extended by using OD-specific thresholds. Alternatively, a travel time-based approach can be used, which is not affected by the network configuration.

Finally, a prominent phenomenon observed in deterministic deterioration-based frameworks is that facilities in series are grouped as *complements* and maintained together, whereas *substitutes* representing facilities in parallel are maintained at different times (Durango-Cohen and Sarutipand 2009). A similar structure is not observed using a network capacity threshold since it only maintains a lower limit on the network flow. In comparison, incorporating link travel time/flow parameters within the objective function, as proposed by other network-based MR&R approaches, can incentivize better co-ordination among construction activities. Another mitigating factor diffusing the substitution/complementarity effects is that the duration of construction activities is measured in weeks. Consequently, the scheduling sub-problem allows the flexibility of scheduling MR&R activities within the same year without having an overlap.

Nevertheless, the presence of a hard constraint such as the network capacity threshold does demonstrate some indirect substitution and complementarity effects within the decision-making process. For instance, given any two links in series, construction activity on either or both links impacts the network capacity identically. In this regard, the two links can be assumed to be complements. Similarly, there exists a scenario wherein two parallel links requiring reconstruction cannot be simultaneously scheduled owing to the network capacity threshold. In this case, one of the two links might be forced to defer its reconstruction to a future time period. Also, as indicated in section 4.3.2, the presence of these inter-facility relationships was also exploited in the selection of the basis functions.

4.5 Discussion

In this chapter, the structural interdependence problem was motivated to account for the adverse impact of construction activities on the road network. With the help of a parametric case study, it is shown that ADP provides a cohesive MR&R decision-making framework which incorporates both budget allocation and work zone scheduling problems. The results also provide important insights into the applicability of ADP-based approaches in the context of infrastructure management. In particular, by benchmarking its performance with SNO, it is shown that ADP is most beneficial for modeling problems wherein finding optimal policies using analytical frameworks is not feasible. The results also indicate that randomized policies, while ideal for modeling budget allocation problems, do not adequately capture the network constraints effectively.

In the next chapter, the economic interdependence problem is explored, wherein cost savings can be achieved by combining MR&R activities across contiguous road sections. Hence, in this case, incorporating the network considerations facilitates a better utilization of the available budget.

Chapter 5

Modeling Economic Interdependence Within System-level MR&R Decision-Making

The costs associated with implementing MR&R activities can be broken down into a fixed set-up component (e.g., for machine rental and operation, labour), and a variable component that is proportional to the intensity of the activity (e.g., material expenditures) (Ouyang 2007). Assuming that the fixed set-up component can be shared across multiple projects, incorporating economies of scale (EOS) within the budget allocation problem can lead to significant to a better utilization of the available budget.

From a methodological standpoint, the economic interdependence problem requires for the optimization framework to recognize the contiguous sections, i.e., pavements in series, in order to achieve the resulting cost savings. Once again, it can be inferred that such a cost structure cannot be adequately represented using randomized policies. As a result, there is a case for investigating an ADP framework to account for the network dynamics.

In this chapter, the performance of ADP and SNO is assessed under different EOS scenarios involving increasing levels of fixed cost representation. Also, the economic interdependence problem is studied using sets of pavements in series, since only contiguous groups of pavement sections can exploit a common fixed cost set-up. However, even though the study does not include the network capacity constraints discussed previously, both aspects of network interdependence can be readily combined.

5.1 Problem Formulation

Let the cost of implementing an action, a , as represented by $c(s, a)$, be subdivided into its fixed and variable cost components as follows:

$$c(s, a) = c_f(a) + c_v(s, a). \quad (5.1)$$

Using this representation, the cost associated with implementing action, a , across n pavement sections in series, can be represented in the following way:

$$\equiv \sum_{i=1}^n c(\mathbf{s}(i, i+1), a) - (n-1)c_f(a), \quad (5.2)$$

where, $\mathbf{s}(i, i+1)$ indicates the condition states of the link $(i, i+1)$. Also, $n-1$ is indicative of the cost savings associated with grouping n pavement sections into a single project.

In order to incorporate these fixed cost savings into the optimization routine, the number of contiguous sections associated with a given MR&R activity needs to be identified with the problem formulation. Alternatively, it is more convenient to model a group of n contiguous sections as a collection of $n-1$ overlapping pairs of pavement sections undergoing a common MR&R activity. For instance, figure 5.1 shows that a group of three and four contiguous pavement sections can be equivalently represented using two and three overlapping pairs of pavement sections respectively. Hence, the economies of scale can be calculated by identifying the number of pairs associated with an MR&R activity.

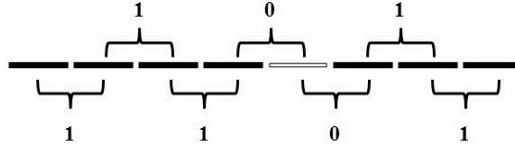


Figure 5.1: Modeling economies of scale for a network in series

With regards to the choice of basis functions, it is assumed that the value function approximation depends only on the state-action pair of the facility. The a-priori beliefs in this case are that all facilities in series are indistinguishable from each other. The number of basis functions to be estimated are $|\mathbf{B}| = T|\mathbf{S}||\mathbf{A}||L|$. Using this information, the optimization problem can be formulated as follows:

$$\begin{aligned} \min_{x,f} & \sum_{i=1}^{|E|} \sum_{a \in \mathbf{A}} \left(c(\mathbf{s}_t(i, i+1), a) + u(\mathbf{s}_t(i, i+1)) + \alpha \theta_{t, \mathbf{s}_t(i, i+1), a} \right) x_{(i, i+1), a, t} \\ & - \sum_{i=1}^{|E|-1} \sum_{a \in \mathbf{A}} c_f(\mathbf{s}_t(i, i+1), a) f_{(i, i+2), a}, \end{aligned} \quad (5.3)$$

subject to

$$\sum_{a \in \mathbf{A}} x_{(i,i+1),a,t} = 1 \quad \forall i = 1, \dots, |E|, \quad (5.4)$$

$$f_{(i,i+2),a,t} \leq x_{(i,i+1),a,t} \quad \forall i = 1, \dots, |E| - 1, \forall a \in \mathbf{A}, \quad (5.5)$$

$$f_{(i,i+2),a,t} \leq x_{(i+1,i+2),a,t} \quad \forall i = 1, \dots, |E| - 1, \forall a \in \mathbf{A}, \quad (5.6)$$

$$\sum_{i=1}^{|E|} \sum_{a \in \mathbf{A}} c(\mathbf{s}_t(i, i+1), a) x_{(i,i+1),a,t} - \sum_{i=1}^{|E|-1} \sum_{a \in \mathbf{A}} c_f(\mathbf{s}_t(i, i+1), a) f_{(i,i+2),a,t} \leq B, \quad (5.7)$$

$$x_{(i,i+1),a,t} \in \{0, 1\}, f_{(j,j+2),a,t} \in [0, 1] \quad \forall a \in \mathbf{A}, \forall i = 1, \dots, |E|, \forall j = 1, \dots, |E| - 1,$$

where,

$f_{(i,i+2),a,t}$: 1, if two adjacent pavements sections $(i, i+1)$ and $(i+1, i+2)$ have a common MR&R activity, a ; 0 otherwise,

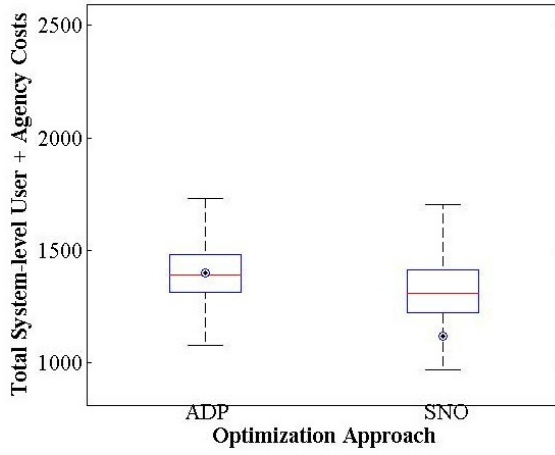
$|E|$: number of facilities in the network; $|E| - 1$ represents the number of nodes in the network (excluding origin and destination), since all facilities are in series.

The equations of interest in the problem formulation outlined above are equations 5.5 and 5.6, which ensure that $f_{(i,i+2),a,t}$ is 1 only when both, $x_{(i,i+1),a,t}$ and $x_{(i+1,i+2),a,t}$ are 1. In other words, the two equations perform the logical AND operation. An interesting feature of the problem formulation is that since $f_{(i,i+2),a,t}$ are bounded by integer variables and its impact of the objective function is to reduce the value of the objective function, its integrality assumption can be relaxed, while ensuring that $f_{(i,i+2),a,t}$ remains a binary variable.

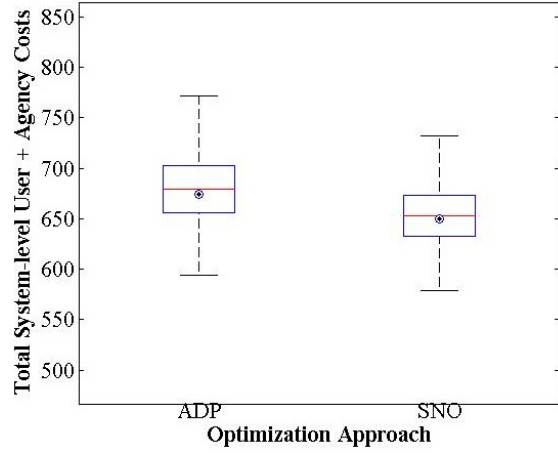
For SNO, the economies of scale can be identically modeled using the facility-specific policies. However, the future cost estimation cannot accommodate the fixed cost component due to the use of randomized policies. Hence, as was the case with structural interdependence, SNO is no longer optimal. In the context of network capacity, the randomized policies violated the capacity constraints. However, in this case, the randomized policies are more conservative in nature, as the problem assumes that economies of scale do not exist in the future. Hence, it is possible that the chosen randomized policies may eventually lead to certain cost savings when the facility-level policies are re-computed. The cause for concern is that these savings are not represented in the estimated future costs, and thus lead to the selection of sub-optimal facility-level policies.

5.2 Numerical Example

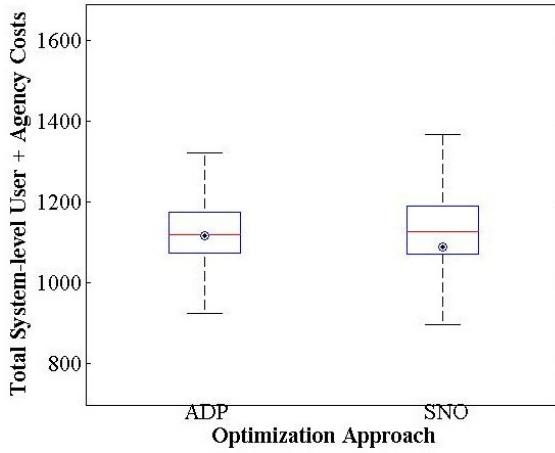
In order to assess the performance of ADP and SNO in incorporating economies of scale in MR&R decision-making, a network of 15 pavement sections in series is considered. It is also assumed that there are no EOS associated with preventive maintenance activities. In comparison, the rehabilitation and reconstruction activities have a fixed cost associated with them. Keeping the total cost of implementing the MR&R activities fixed and identical



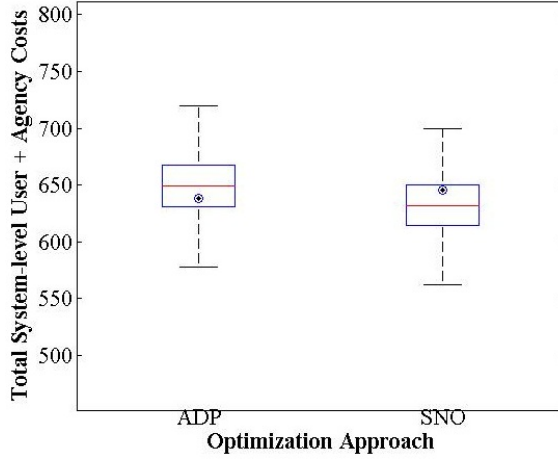
(a) No EOS; $B = 50$



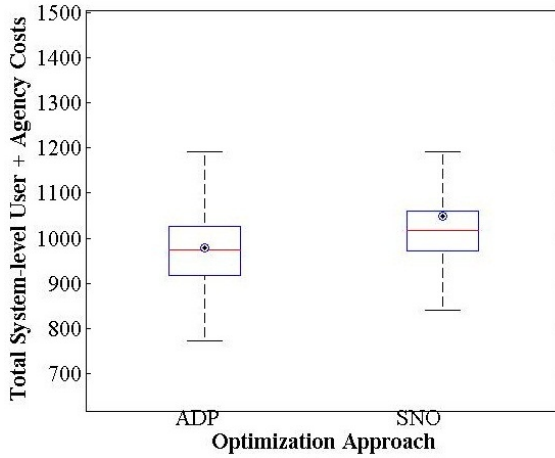
(b) No EOS; $B = 100$



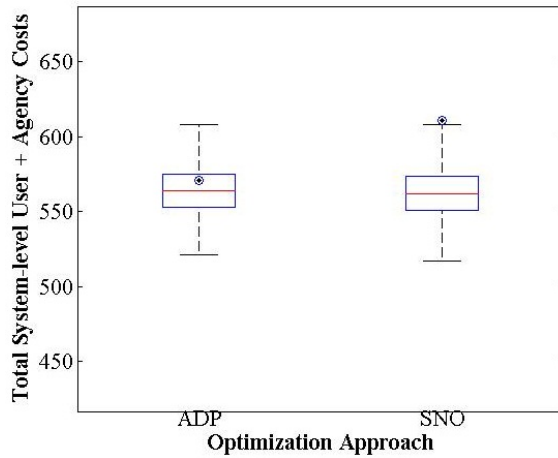
(c) Moderate EOS; $B = 50$



(d) Moderate EOS; $B = 100$



(e) High EOS; $B = 50$



(f) High EOS; $B = 100$

Figure 5.2: Box plots comparing ADP and SNO across varying levels of EOS and budgets

to table A.1, scenarios corresponding to zero, moderate and high EOS are generated for rehabilitation and reconstruction activities.

For moderate EOS, the fixed costs corresponding to rehabilitation and reconstruction are 2 and 10 units respectively. In terms of the percentage of the total rehabilitation cost (fixed + variable), the fixed costs contribute 9-52 %, depending on the condition state of the facility. For the replacement costs, the fixed costs represent around 39% of the total costs, and it does not vary with the condition state of the road section.

For the high EOS scenario, the fixed cost component of the rehabilitation activities is increased to 3 units, which contributes towards 14-79% of the total costs. The fixed cost contribution for a replacement activity is 15 units (57% of the total costs).

Finally, two annual budget levels of 50 and 100 units are considered, and the decision-making process is simulated for both ADP and SNO for 1000 runs using Monte Carlo simulations.

5.3 Results

Figure 5.2 represents the distribution of costs incurred through ADP and SNO for the different EOS and budget level combinations. The simulation results indicate that in the absence of economies of scale (figures 5.2(a)-5.2(b)), SNO performs slightly better than ADP, as is expected. For moderate and high EOS ((figures 5.2(c)-5.2(f)), both approaches perform equally well, with ADP improving upon SNO in the lower budget scenarios. In terms of the costs predicted at $t = 1$, both approaches fare reasonably well, except for one scenario, wherein SNO over-predicts the costs (figure 5.2(f)).

5.4 Conclusions

A significant difference between the economic interdependence problem and the network interdependence problem is that the randomized policies obtained in the presence of EOS provides a conservative estimate of the future costs. If the MR&R actions most benefited by the economies of scale are implemented intermittently, the randomized policies can still provide a reasonably good lower bound for the future costs. However, it is only in scenarios wherein the available budget is very limited, or the incentive to repeatedly combine MR&R actions is very high, that a randomized policy can fail to capture the future costs effectively. Consequently, it is under those scenarios that ADP performs better than SNO.

It is acknowledged that the fixed cost representation used for the case study may not be representative of the true trade-offs observed in the real world. However, here the motivation of the study was to compare the two optimization frameworks when one gradually moves away from a pure budget allocation problem, for which optimal solutions are known, to an EOS setting where SNO may no longer be optimal.

Chapter 6

Conclusion

This chapter provides some concluding remarks for the work presented as part of this dissertation. Section 6.1 provides a summary of the major findings in the dissertation. Section 6.2 discusses ways in which this research can be extended in the future.

6.1 Summary

This dissertation focuses on furthering the development of MDP-based system-level MR&R decision-making frameworks in transportation infrastructure management. Based on the survey of the infrastructure management literature, two problems of interest were identified. The first problem concerns itself with the difficulty in comparing the various MDP frameworks developed for solving budget allocation problems. The second problem of interest involves moving beyond traditional budget allocation problems to incorporate network considerations into system-level decision-making.

With regards to the first research problem, the simultaneous network optimization (SNO) framework, discussed in chapter 3, makes an important contribution towards gaining a better understanding of single dimensional and multidimensional MDP-based approaches for resource constrained MDP problems. The implicit assumption of a homogeneous system in single dimensional MDPs is less restrictive for budget allocation problems, as it allows for the randomize policies to efficiently estimate the expected budget consumption in the future years. While the Golabi et al. (1982) approach has been successfully implemented in practice, the absence of facility-specific policies is often cited as a limitation of the approach. By providing a suitable framework to overcome the identification issues associated with randomized policies, SNO allows for the various MDP-based optimization frameworks to be compared with each other. The results of the comparative study indicate that SNO performs better than the state-of-the-art MDP methodologies for budget allocation problems.

On incorporating network considerations into the MR&R decision-making problem, the presence of an underlying network configuration introduces heterogeneity among the individual road segments. As a result, the optimal nature of randomized policy frameworks is no

longer guaranteed. In addition, most multidimensional MDP frameworks developed for solving budget allocation problems are not easily extendible to account for network constraints. Consequently, the use of approximate dynamic programming is motivated, as it is a general modeling framework used for solving MDP problems.

In the context of infrastructure management, the use of approximate dynamic programming is a recent phenomenon. While previous research establishes that MR&R decision-making problems can be modeled using ADP for budget allocation problems, as part of this dissertation, it is shown that ADP is more relevant for modeling problems involving complex inter-facility dynamics. In particular, the parametric case study presented in chapter 4 illustrates a cohesive MR&R decision-making methodology for policy makers, which integrates both budget allocation and work zone scheduling problems in a unified framework.

From a research perspective, this dissertation provides important insights into the suitability of ADP-based approaches for infrastructure management problems. It also highlights the benefit of benchmarking ADP against other MDP frameworks, especially since theoretical guarantees of optimality cannot be obtained. For instance, while the structural interdependence problem showed that ADP performs significantly better than SNO, the economic interdependence problem (chapter 5) indicates that the performance of ADP can be better/worse than SNO depending on the parameters of the problem. Some of the shortcomings of ADP have also been pointed out, such as the subjectivity in modeling value function approximations, and the issue of exploration versus exploitation.

In conclusion, it is evident from this research that ADP is a robust modeling framework. With rapid improvements in computation power taking place, ADP provides an exciting opportunity for researchers as well as practitioners to look beyond traditional budget allocation problems and model the intricacies of system-level MR&R decision-making.

6.2 Future Work

The research presented as part of this dissertation can be improved along several directions. The discussion on future work is classified here into two parts: Section 6.2.1 outlines the avenues which can be explored to make the ADP framework perform better. Section 6.2.2 briefly describes some problems of interest, which are either direct extensions of the work presented in the dissertation, or other problems pertaining to infrastructure management.

6.2.1 Methodological Improvements

A major concern with any ADP implementation is the issue of exploration versus exploitation. Since the state-action space for infrastructure management problems is invariably large, it is important to identify effective learning strategies which balance the need for exploration and exploitation. In this regard, a promising technique called the knowledge gradient approach is discussed in Powell (2007). The objective of this approach is to maintain estimates of the mean and the variance of the basis function parameters being learnt, and explore for

new solutions with an objective of minimizing the variance associated with the state-action pairs.

The dissertation describes a preliminary implementation of an approximate dynamic programming framework for infrastructure management systems. In texts such as Bertsekas and Tsitsiklis (1996), Sutton and Barto (1998), and Powell (2007), several different ADP techniques, such as approximate policy iteration, actor-critic algorithms, approximate linear programming, etc., have been described. The suitability of these techniques for infrastructure management problems needs to be investigated.

6.2.2 Related Problems of Interest

- In section 4.4, some of the limitations associated with a capacity-based framework were discussed, and potential alternatives, such as the use of a travel time-based network performance measure, were suggested. Some of these alternative modeling techniques would be useful to implement ADP on larger networks.
- The work presented for the economic interdependence problem should be further explored using insights from the parallel machine replacement literature.
- In recent times, state augmentation techniques have been used by researchers to model richer Markovian transition matrices which consider the impact of the prior condition states and maintenance history on the facility deterioration (Robelin and Madanat 2007). An expansion of the state space can lead to computational challenges with regards to efficiently solving the SNO framework. The suitability of an ADP approach is worth exploring in this context.

Appendix A

Cost Structure for Numerical Examples

Table A.1: Cost structure for numerical example

Maintenance Activity	Pavement State							8
	1	2	3	4	5	6	7	
	Acceptable							Unacceptable
Do-Nothing	0	0	0	0	0	0	0	0
Maintenance	0.04	0.15	0.31	0.65	0.83	1.4	2	6.9
Rehabilitation	3.81	3.91	4.11	6.64	9.06	10.69	12.31	21.81
Replacement	25.97	25.97	25.97	25.97	25.97	25.97	25.97	25.97
User Costs	0	2	4	8	14	22	25	100

Appendix B

Transition Probability Matrices

Table B.1: Do nothing transition matrix

s_τ	$s_{\tau+1}$							
	1	2	3	4	5	6	7	8
1	0.6	0.4	0	0	0	0	0	0
2	0	0.5	0.5	0	0	0	0	0
3	0	0	0.4	0.6	0	0	0	0
4	0	0	0	0.35	0.65	0	0	0
5	0	0	0	0	0.3	0.7	0	0
6	0	0	0	0	0	0.2	0.8	0
7	0	0	0	0	0	0	0.1	0.9
8	0	0	0	0	0	0	0	1

Table B.2: Routine maintenance transition matrix

s_τ	$s_{\tau+1}$							
	1	2	3	4	5	6	7	8
1	0.85	0.15	0	0	0	0	0	0
2	0	0.73	0.37	0	0	0	0	0
3	0	0	0.62	0.38	0	0	0	0
4	0	0	0	0.52	0.48	0	0	0
5	0	0	0	0	0.43	0.57	0	0
6	0	0	0	0	0	0.35	0.65	0
7	0	0	0	0	0	0	0.29	0.71
8	0	0	0	0	0	0	0	1

Table B.3: Rehabilitation transition matrix

s_τ	$s_{\tau+1}$							
	1	2	3	4	5	6	7	8
1	1	0	0	0	0	0	0	0
2	0.85	0.15	0	0	0	0	0	0
3	0	0.85	0.15	0	0	0	0	0
4	0	0	0.85	0.15	0	0	0	0
5	0	0	0	0.85	0.15	0	0	0
6	0	0	0	0	0.85	0.15	0	0
7	0	0	0	0	0	0.85	0.15	0
8	0	0	0	0	0	0	0.85	0.15

Table B.4: Reconstruction transition matrix

s_τ	$s_{\tau+1}$							
	1	2	3	4	5	6	7	8
1	1	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0
5	1	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0
7	1	0	0	0	0	0	0	0
8	1	0	0	0	0	0	0	0

Bibliography

- [1] R. Bellman. “The theory of dynamic programming”. In: *Bulletin of American Mathematical Society* 60.6 (1954), pp. 503–515.
- [2] Dimitri P. Bertsekas and John N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996. ISBN: 1886529108.
- [3] Y. Y. Chang et al. “A tabu search based approach for work zone scheduling”. In: *Proceeding of the Transportation Research Board 80th Annual Meeting, Washington, DC*. 2001.
- [4] Suzanne Childress and Pablo Durango-Cohen. “On parallel machine replacement problems with general replacement cost functions and stochastic deterioration”. In: *Naval Research Logistics (NRL)* 52.5 (2005), pp. 409–419.
- [5] James C Chu and Yin-Jay Chen. “Optimal threshold-based network-level transportation infrastructure life-cycle management with heterogeneous maintenance actions”. In: *Transportation Research Part B: Methodological* 46.9 (2012), pp. 1123–1143.
- [6] R. Dekker et al. “A review of multi-component maintenance models with economic dependence”. In: *Mathematical Methods of Operations Research* 45.3 (1997), pp. 411–435.
- [7] N. B. Dimitrov and D. P. Morton. “Combinatorial design of a stochastic markov decision process”. In: *Operations Research and Cyber-Infrastructure* (2009), pp. 167–193.
- [8] P. L. Durango-Cohen. “Maintenance and repair decision making for infrastructure facilities without a deterioration model”. In: *Journal of Infrastructure Systems* 10.1 (2004), pp. 1–8.
- [9] Pablo L Durango-Cohen and Pattharin Sarutipand. “Maintenance optimization for transportation systems with demand responsiveness”. In: *Transportation Research Part C: Emerging Technologies* 17.4 (2009), pp. 337–348.
- [10] Akitoshi Furuya and Samer Madanat. “Accounting for Network Effects in Railway Asset Management”. In: *Journal of Transportation Engineering* (2012).
- [11] T. F. Fwa et al. “Scheduling of pavement maintenance to minimize traffic delays”. In: *Transportation Research Record: Journal of the Transportation Research Board* 1650.1 (1998), pp. 28–35.

- [12] L. Gao and Z. Zhang. “Approximate Dynamic Programming Approach to Network-Level Budget Planning and Allocation for Pavement Infrastructure”. In: *Transportation Research Board 88th Annual Meeting*. Number 09-2344, 2009.
- [13] K. Golabi and R. Shepard. “Pontis: A system for maintenance optimization and improvement of US bridge networks”. In: *Interfaces* 27.1 (1997), pp. 71–88.
- [14] Kamal Golabi et al. “A statewide pavement management system”. In: *Interfaces* 12.6 (1982), pp. 5–21.
- [15] R. Hajdin and H. P. Lindenmann. “Algorithm for the planning of optimum highway work zones”. In: *Journal of Infrastructure Systems* 13.3 (2007), pp. 202–214.
- [16] Philip C Jones et al. “Parallel machine replacement”. In: *Naval Research Logistics (NRL)* 38.3 (1991), pp. 351–365.
- [17] L. C. M. Kallenberg. “Survey of linear programming for standard and nonstandard Markovian control problems. Part I: Theory”. In: *Mathematical Methods of Operations Research* 40.1 (1994), pp. 1–42.
- [18] Nejat Karabakal et al. “Parallel replacement under capital rationing constraints”. In: *Management Science* 40.3 (1994), pp. 305–319.
- [19] K. Kuhn and S. Madanat. “Model uncertainty and the management of a system of infrastructure facilities”. In: *Transportation Research Part C: Emerging Technologies* 13.5-6 (2005), pp. 391–404.
- [20] K. D. Kuhn. “Network-Level Infrastructure Management Using Approximate Dynamic Programming”. In: *Journal of Infrastructure Systems* 16.2 (2010), pp. 103–111.
- [21] E. B. Lee and C. W. Ibbs. “Computer simulation model: Construction analysis for pavement rehabilitation strategies”. In: *Journal of Construction Engineering and Management* 131.4 (2005), pp. 449–458.
- [22] S. Madanat. “Optimal infrastructure management decisions under uncertainty”. In: *Transportation Research Part C: Emerging Technologies* 1.1 (1993), pp. 77–88.
- [23] Samer Madanat et al. “Adaptive optimization and systematic probing of infrastructure system maintenance policies under model uncertainty”. In: *Journal of Infrastructure Systems* 12.3 (2006), pp. 192–198.
- [24] ManWo Ng et al. “Optimal Long-Term Infrastructure Maintenance Planning Accounting for Traffic Dynamics”. In: *Journal of Computer-Aided Civil and Infrastructure Engineering* 24.7 (2009), pp. 459–469. ISSN: 10939687. DOI: 10.1111/j.1467-8667.2009.00606.x.
- [25] J. W. Ohlmann and J. C. Bean. “Resource-constrained management of heterogeneous assets with stochastic deterioration”. In: *European Journal of Operational Research* 199.1 (2009), pp. 198–208.
- [26] Y. Ouyang. “Pavement resurfacing planning for highway networks: parametric policy iteration approach”. In: *Journal of Infrastructure Systems* 13.1 (2007), pp. 65–71.

- [27] Y. Ouyang and S. Madanat. “An analytical solution for the finite-horizon pavement resurfacing planning problem”. In: *Transportation Research Part B: Methodological* 40.9 (2006), pp. 767–778.
- [28] W. B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*. John Wiley & Sons, Inc., 2007.
- [29] Charles-Antoine Robelin and Samer M Madanat. “History-dependent bridge deck maintenance and replacement optimization with Markov decision processes”. In: *Journal of Infrastructure Systems* 13.3 (2007), pp. 195–201.
- [30] Nakul Sathaye and Samer Madanat. “A bottom-up solution for the multi-facility optimal pavement resurfacing problem”. In: *Transportation Research Part B: Methodological* 45.7 (2011), pp. 1004–1017. ISSN: 01912615. DOI: 10.1016/j.trb.2011.03.002.
- [31] H. P. Simão et al. “An approximate dynamic programming algorithm for large-scale fleet management: A case application”. In: *Transportation Science* 43.2 (2009), pp. 178–197.
- [32] K.C. Sinha et al. *The Development of Optimal Strategies for Maintenance, Rehabilitation and Replacement of Highway Bridges, Vol. 1: The Elements of the Indiana Bridge Management System (IBMS)*. Tech. rep. FHWA/IN/JHRP-88/15-1. Joint Highway Research Project. Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 1988.
- [33] K. Smilowitz and S. Madanat. “Optimal inspection and maintenance policies for infrastructure networks”. In: *Journal of Computer-Aided Civil and Infrastructure Engineering* 15.1 (2000), pp. 5–13.
- [34] J.P. Sousa and L.A. Wolsey. “A time indexed formulation of non-preemptive single machine scheduling problems”. In: *Mathematical Programming* 54.1 (1992), pp. 353–367.
- [35] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. Cambridge University Press, 1998.
- [36] G. Tesauro. “Temporal difference learning and TD-Gammon”. In: *Communications of the ACM* 38.3 (1995), pp. 58–68.
- [37] K. Wunderlich and D. Hardesty. *A snapshot of summer 2001 work zone activity*. Tech. rep. EDL: 13793. National Transportation Library, FHWA, U.S. Department of Transportation, 2003.
- [38] H. Yeo et al. “Algorithms for bottom-up maintenance optimisation for heterogeneous infrastructure systems”. In: *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance* (2012). , DOI: 10.1080/15732479.2012.657649.