## UNDERSTANDING UNCERTAINTY: A REINFORCEMENT LEARNING APPROACH FOR PROJECT-LEVEL PAVEMENT MANAGEMENT SYSTEMS

by

Ayatollah Yehia

# A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

### MASTER OF APPLIED SCIENCE

in

### THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Civil Engineering)

### THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2020

© Ayatollah Yehia, 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

#### UNDERSTANDING UNCERTAINTY: A REINFORCEMENT LEARNING APPROACH FOR PROJECT-LEVEL PAVEMENT MANAGEMENT SYSTEMS

submitted by	Ayatollah Yehia	in partial fulfilment of the requirements for			
the degree of	Master of Applied Science				
in	Civil Engineering				
Examining Committee:					

Dr. Omar Swei, Assistant Professor, Department of Civil Engineering, UBC Supervisor

Dr. Terje Haukaas, Professor, Department of Civil Engineering, UBC Supervisory Committee Member

## Abstract

Transportation agencies have limited fiscal resources to manage their pavement infrastructure. Planning for the future includes uncertainty, such as the uncertainty of future traffic levels, cost of rehabilitation actions, price indices, among others. Deterioration modeling also includes uncertainty, such as random and measurement uncertainty. Failing to consider these uncertainties may lead to sub-optimal management policies that are unable to adapt to the future. Thus, the objective of this thesis is to develop a reinforcement learning algorithm to manage pavement systems at the project-level that minimizes the life-cycle cost.

The deterioration model developed uses an iterative-methods approach to estimate infrastructure performance models based on sampling theory. The model addresses the issue around measurement uncertainty underlying infrastructure condition assessments for continuous distress indicators and its effect on the parametric models underlying decision-support tools. Through a case study of pavement roughness data collected as part of Federal Highway Administration's long-term pavement performance program, the new approach reduces the unexplained variance that would typically enter decision-support tools by 14%. It also addresses concerns around heteroscedasticity surrounding conventional methods, allowing modelers to recover efficiency in their statistical estimates.

Finally, the Q-learning algorithm with an  $\varepsilon$ -greedy policy efficiently learns an optimal management policy for infrastructure assets while simultaneously incorporating several sources of uncertainty. An important advantage of this approach is that it is model-free and non-parametric, imposing no restrictions on the structure of the uncertain inputs. This study subsequently implements the Q-learning approach across three separate case studies. The

proposed algorithm leads to the selection of a management policy that, on average, reduces expected life-cycle costs between 3% and 15% compared to traditional infrastructure management approaches.

This research contributes to the pavement management literature by creating improved performance models and providing a holistic view of uncertainties in the management process. There are several opportunities to expand upon this research which are discussed.

## Lay Summary

This research introduces an innovative tool to manage pavement systems. A deterioration model was developed to consider the uncertainty caused by random variability and deterioration measurements. By doing so, the deterioration accuracy increased in comparison to current models. The advantage of improving deterioration accuracy is more accurate maintenance policies for pavements. This helps agencies manage their limited budgets more efficiently.

A reinforcement learning tool that minimizes the cost of managing a pavement throughout its life was also designed. This tool also considers the uncertainty of future traffic volumes and price indices for rehabilitation activities. When compared to how pavements are currently being managed, the reinforcement learning tool reduced the agency's management cost. The advantage of this tool is that it can be applied to any pavement management system, allowing for "flexibility between locations and input variables". These results help transportation agencies manage their limited budgets by making more optimal decisions.

## Preface

The research presented in this thesis was conducted by me under the supervision of Dr. Omar Swei (Assistant Professor at the University of British Columbia). I was responsible for conducting the bulk of the literature review, writing the scripts involved in this project as well as analyzing the results. Emily Wong provided additional support for the literature reviews in Chapter 2 and 4. Dr. Omar Swei continuously provided feedback and advice throughout the thesis.

Chapter 2, 3 and 4 are the heart of the thesis. Chapter 3 was published in *Transportation Research Part C: Emerging Technologies* (Yehia, A. and Swei, O. (2020). 'Probabilistic Infrastructure Performance Models: An Iterative-Methods Approach', *Transportation Research Part C: Emerging Technologies*, 111, pp. 245–254.) while Chapter 2 and 4 were presented at the Annual Transportation Research Board conference located in Washington D.C, USA in January 2020.

# **Table of Contents**

Abstractiii
Lay Summaryv
Prefacevi
Table of Contents vii
List of Tablesix
List of Figuresx
Lists of Symbols, Abbreviations, or Other xi
Acknowledgements xii
Dedication xiii
Chapter 1. Introduction1
Chapter 2. Life-cycle planning of pavements 6
2.1 Life-cycle phases of a pavement
2.2 LCCA components
2.2.1 Uncertainty in LCCA
2.2.2 LCCA Inputs 10
2.2.3 Existing Studies 12
2.3 Limitations and Contribution
Chapter 3. Deterioration Modeling18
3.1 Existing Pavement Degradation Models

3.2 Methodology
3.3 Case Study Analysis
3.4 Results
3.5 Contributions
Chapter 4. Reinforcement Learning 41
4.1 RL Approaches in Relevant Applications
4.2 Methodology 45
4.2.1 <i>Q</i> -learning Algorithm for the Optimization of Pavement LCCA
4.3 Description of Case Studies
4.4 Case Study Results
4.5 Contributions
Chapter 5. Conclusion 60
5.1 Future work
References

# List of Tables

Table 1. Summary of existing LCCA studies and uncertainties	. 15
Table 2. Estimate of Equation 12 using (a) the proposed IRLS approach and (b) the typical	1
OLS regression approach	. 37
Table 3. Key inputs for the probabilistic LCCA model	. 51
Table 4. Probabilistic cost and effect of different available actions	. 51
Table 5. Design life and analysis period for case studies as well as initial action per the	
traditional LCCA approach	. 52
Table 6. Design life, analysis period for case studies and initial action per the traditional	
LCCA approach	. 57

# **List of Figures**

Figure 1. Phases of a pavement's life cycle7
Figure 2. Histogram of standard error of mean IRI condition (meters per kilometer) for 99%
of pavement facilities
Figure 3. Average standard error of mean IRI condition (meters per kilometer) of paving
facilities across time
Figure 4. Q-Q plot comparing the distribution of the weighted residuals (y-axis) to a standard
normal distribution (x-axis)
Figure 5. Minimum <i>Q</i> -value at starting period across all actions in learning environment (i.e.,
Algorithm 2)
Figure 6. Cost-to-go at starting period for each action in learning environment (i.e.,
Algorithm 1) for Case Number 1 and 2
Figure 7. Cost-to-go at starting period for each action in learning environment (i.e.,
Algorithm 1) for Case Number 1 and 2
Figure 8. Probabilistic LCC of Q-learning and traditional LCC approaches per Algorithm 2
for Case Number 2

# Lists of Symbols, Abbreviations, or Other

#### Pavement related abbreviations

AC	Asphalt Concrete
FHWA	Federal Highway Administration
IRI	International Roughness Index
LCA	Life-Cycle Assessment
LCC	Life-Cycle Cost
LCCA	Life-Cycle Cost Application
LTPP	Long-Term Pavement Performance
MEPDG	Mechanistic-Empirical Pavement Design Guide
MDP	Markov Decision Process
PMS	Pavement Management System
RL	Reinforcement Learning
SN	Structural Number

### Statistics abbreviations

GLS	Generalized Least Squares	
IRLS	Iterative Reweighted Least Squares	
OLS	Ordinary Least Squares	
Q-Q	Quantile-Quantile	
RMSE	Root Mean Square Error	
WLS	Weighted Least Squares	
2SLS	2-Stage Least Squares	

## Acknowledgements

I would like to thank my supervisor, Dr. Omar Swei, for his guidance and support during the last two years. Dr. Omar made me a better programmer and researcher and I will always be grateful for his advice and patience.

The professors I had the pleasure of taking courses from, made this research stronger. Dr. Tarek Sayed, Dr. Alex Bigazzi and Dr. Alan Russel in the civil engineering department and Dr. Mehrdad Oveisi-Fordoei and Dr. Giuseppe Carenini in the computer science department, thank you all for providing me with the pieces I needed to put this puzzle together.

My family, although we were thousands of miles away for months on end, never stopped being right by my side. Thank you, Dr. Sherif, my father, for being the first to motivate me on this unconventional path. My mother, who never stopped sending me love and food from home. My siblings who kept me on my metaphorical (and literal) toes.

My friends made sure I stayed balanced and motivated. Thank you for taking me out of the office occasionally to grab some coffee and have a chat. I don't think I would've enjoyed this process without our dosa nights and gatherings.

This research has been partially supported by the National Sciences and Engineering Research Council (NSERC) of Canada Discovery Grant Program

# Dedication

I dedicate this research to my family, both by genetics and by choice. Thank you for supporting me along the way.

## **Chapter 1. Introduction**

Pavement systems are one of the most important facilities of transportation infrastructure, with nine trillion tonne-kilometers of transport and freight being supported on over fifteen trillion kilometers every year worldwide (Santero et al. 2011a). An efficient system provides economic opportunities and benefits (i.e., the creation of jobs, an increase in private investment, trade stimulation) while a deficient system can have negative economic effects (i.e., a constraint on the migration of labor, slower market expansion, poor material handling) (Rodrigue 2017; Deng 2013). Thus, it is important for transportation agencies to maintain pavement systems at a certain performance level to provide societal benefits, (e.g., lower travel costs for goods and services) (Boarnet 1997). However, agencies have limited fiscal resources to manage their existing infrastructure (Arif and Bayraktar 2018). In 2016, the American Society of Civil Engineers (ASCE) conducted a survey that illustrates the gap between the funding available for networks and the funding required to maintain them; road networks are about 50% underfunded (ASCE 2016). With the condition of American roads being graded a 'D' or a large facility of the system requiring maintenance by the ASCE, it is evident more funding is required (ASCE 2017).

To manage this infrastructure with limited budgets, pavement management systems (PMS) can be used to make informed decisions (Torres-Machi et al. 2018). A PMS is a tool used to determine scheduling of pavement activities, allocation of resources and budgeting (FAA 2006). The Department of Transportation in Arizona (ADOT) was one of the first to adopt the use of a PMS in 1980; there were issues on how to distribute money to Arizona districts and the effects of road conditions with budget cuts (Golabi et al. 1982). The system had to: (a) contribute to the planning process by contributing to budget allocation, (b) be relatively simple to use and, (c) be flexible enough for sensitivity analysis to occur. This used Markov models and linear programming to model pavement deterioration and minimize agency cost respectively (Li et al. 2006). This influenced the development of similar international systems that utilize optimization models; some of these models are discussed in Chapter 2.

A PMS can be applied at the network level, which focuses on a cost-efficient way to maintain the pavement network, or the project level, which focuses on a certain facility of the network (Ismail et al. 2009). While a PMS may consist of different components, the four main elements are: (1) network inventory, (2) condition evaluation, (3) prediction models on performance and, (4) planning (Shahin 2005). A network inventory is a record of the pavement facilities in a network, where a facility is the smallest unit that a construction and maintenance action can be applied (Ismail et al. 2009). Pavement condition evaluation considers the functional evaluation of a pavement, such as its roughness and surface distresses while the structural evaluation of a pavement considers its structural capacity. Performance prediction models at both the project and network-level focus on evaluating the condition of the system and determining the rehabilitation actions required. These models mainly include the predicted physical deterioration of the pavement system and may be updated when the true pavement behavior is available (Chootinan et al. 2006). Tools such as regression, straight line extrapolation and Markovian models can be used to develop these models (Ismail et al. 2009). At the network level, these prediction models are used for budget planning and condition forecasting, while at the project level, prediction models are used to decide maintenance actions with the aid of life-cycle cost analysis. These prediction models can evaluate potential design and maintenance alternatives and estimate the required thickness for this structure given the parameters (Abaza 2004). Planning methods, the final component of PMS, is the module where future maintenance and rehabilitation actions are decided (Ismail et al. 2009). These action

plans are determined based on the: (1) current pavement condition, (2) predicted pavement performance and, (3) available fiscal resources.

Within performance prediction models at the project level, life-cycle cost analysis (LCCA) and life-cycle assessment (LCA) are available for use. LCCA is used to reduce the costs over the lifetime of a facility by evaluating design, construction and maintenance actions to meet that objective of minimizing the agency's costs (Santos et al. 2019). LCCA considers the cost of materials (initial construction) as well as the cost of rehabilitation actions in order to use limited budgets more efficiently. LCA is used to estimate the direct and indirect environmental impact of a pavement system, such as the impact of alternative construction and maintenance activities (Araújo et al. 2014; Yu and Lu 2012). The impact category of this assessment is typically in Global Warming Potential (GWP), where CO<sub>2</sub>, CH<sub>4</sub> and N<sub>2</sub>O or greenhouse gases are converted into CO<sub>2</sub>-equivalent emissions (Araújo et al. 2014; Huang et al. 2009).

However, there are challenges with using deterioration models, LCCA and LCA in PMS. In LCCA and LCA, this challenge is the uncertainty in estimating future parameters, such as deterioration, costs and traffic, with regards to short term and long-term decisions (Swei et al. 2015). If inputs are treated as deterministic, this may cause decision-makers to choose actions that are no longer suitable in the future. With regards to deterioration models, a critical component of a PMS, a low-fidelity deterioration model will lead to the selection of a sub-optimal construction design and maintenance schedule (Swei et al. 2018).

With limited fiscal resources, it is important to have a pavement management tool that: (a) considers the uncertainty in pavement measurements in deterioration modeling and, (b) considers the uncertainties present in life-cycle cost (LCC) for maintenance scheduling over a selected period

3

of time. While the literature provides countless studies on pavement management tools that consider the aforementioned challenges, these studies do not focus on every aspect of these points. For example, a PMS may focus on many rehabilitation alternatives but only considers one possible source of uncertainty due to computational limits. Thus, the objective of this thesis is to bridge the gap between theory and practice by designing a tool that incorporates uncertainty in the deterioration and management processes while optimizing decisions. Improving deterioration modeling by incorporating uncertainty will lead to better degradation estimates, which, in turn, lead to a more optimal method to allocate limited budgets. With uncertainties being pervasive and the need to be more proactive in pavement management decisions, it is important now, more than ever, to create this tool. By using reinforcement learning to make smarter decisions that are riddled with uncertainty, this developed tool will support a more sustainable future.

The specific goals of this thesis are to:

- 1. Develop a pavement degradation model that is able to calculate the International Roughness Index (IRI) after accounting for both random and measurement uncertainties.
- 2. Develop a script that calculates the fiscal cost of a transportation agency to maintain a pavement section for a specified time period. The script should also account for various present and future uncertainties such as pavement degradation, price indices, the cost of construction and maintenance actions and traffic growth. The output of this script should be a policy that minimizes the LCC of a pavement facility at the project level.

Chapter 2 of this thesis reviews the existing LCA/LCCA models at the project level. Previously designed models and their considered sources of uncertainty are discussed prior to identifying the existing gaps. Chapter 3 discusses deterioration model, which includes reviewing the existing

literature, describing the designed method and model as well as the discussing the results of the deterioration model. Chapter 4 includes reinforcement learning; a literature review, the chosen algorithm and the results of using reinforcement learning on the listed objectives. Finally, Chapter 5 discusses the results, reviews the objectives and identifies areas of opportunity.

## **Chapter 2. Life-cycle planning of pavements**

Life Cycle Cost Analysis (LCCA) was first developed by the U.S Department of Defense in the 1960s to increase the cost-effectiveness of government spending (Shields and Young 1991). In the 1990s, pavement LCCA expanded into the federal literature, including vehicle-operating-cost models (Liu et al. 2015). In 1995, the Federal Highway Administration (FHWA) made it compulsory to utilize LCCA for National Highway System projects that cost more than \$25 million in order to "reduce long-term costs and improve quality and performance" (FHWA 1996). FHWA and AASHTO continue to provide guidance to states developing their own LCCA procedures, as more than 80% of states in the United States use LCCA and consider agency cost to manage their limited fiscal resources (Ozbay et al. 2004). Both researchers and practitioners incorporate the uncertainty of relevant input parameters in LCCA.

LCCA in pavement systems can be used to account for the economic impact materials and rehabilitation alternatives have (Liu et al. 2015). For example, applications of LCCA may compare materials used in rigid and flexible pavements or pavement types at the project and network level (Batouli et al. 2017). In short, LCCA provides agencies with a comprehensive framework to make economically driven decisions between alternative investments (Santero et al. 2011a, Guo et al. 2019)

Life Cycle Assessment (LCA) has been used since the 1990s to evaluate the environmental impact of pavement infrastructure (Santero et al. 2011a). Like LCCA, it is a decision-support tool that is able to assess the environmental impact and burden of alternatives throughout the pavement's lifecycle (Zhang et al. 2008). Figure 1, taken from Santero et al. (2011a), illustrates the relevant lifecycle phases for pavements. However, as social and environmental effects do not affect agency costs, they are often discarded; only 40% of states in the United States consider social and environmental impacts (Heidari et al. 2020).



Figure 1. Phases of a pavement's life cycle (Santero et al. 2011a)

For infrastructure, the goal of LCA/LCCA is to quantify the cost effectiveness and environmental impacts for design, maintenance and rehabilitation alternatives that meet certain service requirements (Reigle and Zaniewski 2002; Walls and Smith 1998). By quantifying and comparing the behavior of pavements when alternative materials and rehabilitation techniques are applied, sustainability goals can be met (Liu et al. 2015). However, LCCA and LCA do not make a decision; rather these tools are used to support a management decision.

This thesis specifically focuses on LCCA as most transportation agencies have implemented a form of this application in their systems. Future work includes an integrated LCCA-LCA tool.

#### 2.1 Life-cycle phases of a pavement

Figure 1 illustrates the five main stages of a pavement's lifecycle along with the activities associated with each stage. The first stage, the material production phase, includes the extraction of raw materials, such as limestone and the production of cement and other pavement materials (Liu et al. 2015). This phase also includes any necessary transportation of materials. The construction phase includes onsite construction equipment and traffic delay caused by construction activities. The longest part of a pavement's lifecycle is the use phase (Santero et al. 2011b). An engine will exert effort to keep a vehicle's tires rolling over a pavement; this is known as rolling resistance (Trupia et al. 2017). In the use phase, rolling resistance, albedo (i.e., radiative forcing), lighting that is used on pavements, among other factors are how pavements interact with the environment phase (Santero et al. 2011b). With rolling resistance, roads with high volumes and heavy truck traffic will have a higher environmental impact than low volume roads with few trucks. However, it's been found that many pavement LCA studies do not fully expand on the use phase in their analysis (Santero et al. 2011a). These studies may use absolute values for total traffic emissions instead of those that are specifically the pavement's contribution.

The maintenance phase includes rehabilitation actions that occur during the life of the pavement. This phase is connected to every phase of the life-cycle as this phase has impacts similar to those of the materials and construction phase. Thus, it can occur during the use phase and after the end-of-life phase (Santero et al. 2011a). At the end-of-life phase, pavements can be either (a) demolished and landfilled (i.e. "cradle-to-grave"); (b) demolished and recycled (i.e. "cradle-to-cradle"); or (c) remain and support the following pavement structure (i.e. "cradle-to-cradle") (Santero et al. 2011b; Liu et al. 2012). Pavements are typically not chosen to be demolished and landfilled due to the decline of natural resources and increasing construction prices. However,

there may be specific materials that cannot be recycled (Santero et al. 2011b). If a pavement is recycled or supports the following pavement structure, the environmental impacts become complicated to determine, as assumptions, forecasting and allocation uncertainty affect this decision. Environmental impacts, such as demolition and transportation impacts, are included in this scenario (Liu et al. 2015).

As the tool is focusing on reducing LCC for agencies, the construction and maintenance phases, which also include material production phase, are the ones that are most important. These actions directly affect agencies. The end-of-life phase for this tool assumes that pavement sections can be salvaged depending on their last rehabilitation action and deterioration rate. Chapter 4 expands on the salvage rate in more detail.

#### **2.2 LCCA components**

#### 2.2.1 Uncertainty in LCCA

There are several inputs that LCCA uses to make an informed decision. These measures are associated with two types of uncertainty. The first type of uncertainty is aleatory (i.e., random) uncertainty. This uncertainty is associated with randomness in samples and parameters, as measured values could be different from their true values (Babashamshi et al. 2016; Ilg et al. 2017). The second type of uncertainty is epistemic uncertainty. This is associated with system performance as well as variability (Reza et al. 2013). Epistemic uncertainty can stem from a lack of information, ambiguity and incomplete data (Zhang et al. 2006; Reza et al. 2013). While both types of uncertainty involve variability in values, aleatory uncertainty is statistical, resulting from stochastic methods involving experimental data such as Monte Carlo simulations (Reza et al. 2013).

To account for these uncertainties that underly the planning of transportation infrastructure, agencies rely on probabilistic LCCA to evaluate alternative pavement design and maintenance schedules (Abdelaty et al. 2016; Swei et al. 2015). While these models incorporate several sources of uncertainty, pavement engineers typically use these models to evaluate only a few available design and construction alternatives (Swei et al. 2015; Pittenger et al. 2012). This reality has motivated researchers to develop multiple optimization-based approaches to minimize life-cycle costs for a pavement facility. The important advantage of an optimization-based approach is that it facilitates a greater exploration of available design and maintenance choices. Due to computational limitations, however, these methods tend to only consider uncertainty around pavement degradation, as further discussed in the upcoming section. Alternatively, these approaches may be deterministic (Wu et al. 2017), thereby failing to account for uncertainty in relevant input factors (e.g., construction costs, future traffic). By treating these inputs as deterministic, current optimization-based approaches leave decision-makers susceptible to suboptimal investment strategies that are unable to adapt to unknown future conditions.

#### 2.2.2 LCCA Inputs

Rehabilitation actions, which include routine maintenance and reconstruction, are the costs that directly affect agencies. The uncertainty with rehabilitation actions is due to the varying cost of these actions and limited information of the cost of maintenance actions. Several studies have implemented uncertainty of cost in LCCA, which will be discussed in the following section. Maintenance timing and deterioration also affect LCC, as there are two possible strategies for maintaining pavement segments. The first strategy is conventional maintenance, where an agency will only maintain when a pavement segment has reached its maximum deterioration. However, preventative maintenance (i.e., maintenance that occurs when a threshold, before maximum

deterioration, is reached) is more cost-effective than conventional maintenance (Babashamsi et al. 2016; Wu et al. 2017). Determining when to maintain depends on how fast a pavement deteriorates, which is typically measured using the International Roughness Index (IRI), which is expressed in units of slope (m/km, in/mi) (Park et al. 2007; Liu et al. 2015). A lower IRI value, such as 0.0 m/km, means a pavement is perfectly smooth, but this is not feasible due to pavements having bumps and dips. There is no maximum IRI, but an IRI above 8 m/km means a vehicle cannot access that pavement except at reduced speeds. In the designed system mentioned in Chapter 4, the minimum IRI when a pavement segment is constructed and maintained is 1.0 m/km (63.6 in/mi) while the maximum IRI is 2.36 m/km (150 in/mi). Deterioration is affected by several factors, such as average annual daily truck traffic (AADTT), age of a pavement segment and pavement segment's thickness or structural number (SN). Deterioration also has associated uncertainties such as aleatory and measurement uncertainty, which affect how fast a segment degrades. Thus, future planning of maintenance actions is associates with several uncertainties that need to be accounted for.

Traffic, as mentioned previously, affects deterioration and ultimately, is an impact factor in LCCA. Future traffic projections are usually based on traffic growth rates, which are calculated from historical traffic data. However, these are subject to uncertainty as growth rates may fluctuate and thus, not follow the projected demand (Li and Madanu 2009; Wu et al. 2017). Rates can vary depending on the geographic location of a segment and their proximity to growth areas.

Maintenance treatment intensity refers to the type of rehabilitation treatment chosen for a segment depending on its deterioration (Swei et al. 2016). Failing to apply the right maintenance treatment at the scheduled time could result in a shorter service life (Li and Madanu 2009). Uncertainty with maintenance treatment intensity stems from: (1) when to apply the treatment, which is associated

11

with scheduling, dependent on deterioration and, (2) the type of treatment to be applied, dependent on deterioration.

In LCCA, discount rates are one of the most significant factors, which determines the present value of costs in the future by reflecting historical trends over a long time period (Santero et al. 2011c; Wu et al. 2015). FHWA recommends using discount rates that are published in the latest version of the White House Office of Management and Budget (OMB). However, there is no universal agreement on the discount rate that should be applied to pavement projects. State agencies utilize discrete real rates, which factor out inflation rates, that range from 3% to 5.3% while a small number of states use a probabilistic analysis to address the underlying uncertainty of the discount rate due to fluctuations throughout a pavement's life-cycle (Li and Madanu 2009; Wu et al. 2017).

#### 2.2.3 Existing Studies

A large number of probabilistic LCCAs evaluate a few alternatives and use Monte Carlo simulations to propagate their life-cycle cost uncertainty. Probabilistic LCCAs have focused on the uncertainty in future traffic volumes (Guo et al. 2012; Harvey et al. 2012; Jawad and Ozbay 2006; Li and Madanu 2009; Reigle and Zaniewski 2002; Salem et al. 2003; Zhang et al. 2010), construction costs, such as material prices (Harvey et al. 2012; Huang et al. 2004; Swei et al. 2015; Tighe 2001), maintenance treatment intensities (Li and Madanu 2009), and future maintenance schedules, which is done through pavement deterioation models (Harvey et al. 2012; Huang et al. 2012; Huang et al. 2004; Salem et al. 2003; Swei et al. 2015). There are also several studies on optimization-based methods that consider more rehabilitation actions and generate an optimal policy. There are deterministic approaches, such as Santos et al. (2019) and Mamlouk et al. (2000). While deterministic models are simple and computationally efficient, they fail to capture the uncertainty in an infrastructure's stochastic deterioration (Morcous and Lounis 2005). Morcous and Lounis

(2005) developed genetic algorithms that utilize Markov-chain models to optimize the maintenance of infrastructure networks. These captured the uncertainties that are present in infrastructure deterioration, measurement errors and condition assessment.

There are several other optimization-based approaches that only consider the underlying uncertainty of pavement deterioration. Kuhn and Madanat (2005) developed a robust linear programming algorithm to minimize agency and user cost. Kuhn (2009) used approximate dynamic programming via value function approximation to generate the optimal policies for individual facilities. These were then used to generate network-level maintenance schedules with fiscal constraints. To minimize the life-cycle cost of a specific system, Guo et al. (2019) used a heuristic enumeration approach that incorporated uncertainty in both pavement deterioration and cost, while other studies focused on a single source of uncertainty. Finally, while Jawad and Ozaby (2006) and Zhang et al. (2010) do not account for the uncertainty in pavement deterioration, both account for the uncertainty in future traffic and/or costs via genetic algorithms and dynamic programming. Table 1 summarizes these findings and categorizes them by the considered sources of uncertainty and LCCA method.

While the aforementioned optimization-based methods should theoretically lead to near optimal solutions that achieve the objective of minimizing life-cycle costs, there are two major limitations. First, by excluding several relevant sources of uncertainty, these approaches are susceptible to selecting management policies that are unable to react to future conditions that deviate from the predicted expectations. The computational power required for these methods limit the uncertainties that can be considered. This is due to the "curse of dimensionality", in which the state space grows as the problem becomes more complex (Tack and Chou 2002). Optimization-based LCCA that utilize Markov models (i.e., model-based approaches) also require transitional probabilities in

order to move from one state to the other. A state in a pavement management system model could be the pavement condition in a system in a point of time, before or after a rehabilitation action has been selected (Guillaumot et al. 2003). Typically, these are derived from empirical data that may differ between agencies; an agency would be unable to benefit from an LCCA that does not accurately use its system's data, which is another limitation (Ravirala and Grivas 1995; Guillaumot et al. 2003). These enforce restrictive assumptions and/or constraints, such as the specification of factors (i.e., pavement deterioration) (Powell 2009). Durango-Cohen (2004) and Medury and Madanat (2013) have explored the benefits of adapting a model-free reinforcement learning algorithm, which will be further discussed in future chapters.

## Table 1. Summary of existing LCCA studies and uncertainties

Study	<b>Considered Sources of Uncertainty</b>			Method		
	Traffic	Maintenance Intensity	Maintenance Timing and Deterioration	Cost	Non-Optimization	Optimization
Mamlouk et al. (2000)						$\checkmark$
Tighe (2001)				$\checkmark$	✓	
Reigle and Zaniewski (2002)	✓				√	
Salem et al. (2003)	$\checkmark$		$\checkmark$		$\checkmark$	
Durango-Cohen (2004)			$\checkmark$			$\checkmark$
Huang et al. (2004)			$\checkmark$	$\checkmark$	$\checkmark$	
Kuhn and Madanat (2005)			$\checkmark$			$\checkmark$
Jawad and Ozbay (2006)	$\checkmark$			$\checkmark$		$\checkmark$
Li and Madanu (2009)	$\checkmark$	✓		$\checkmark$	$\checkmark$	
Kuhn (2009)			$\checkmark$			$\checkmark$
Zhang et al. (2010)	$\checkmark$					✓
Irfan et al. (2012)						$\checkmark$
Guo et al. (2012)	$\checkmark$				✓	
Harvey et al. (2012)	$\checkmark$		$\checkmark$	$\checkmark$	$\checkmark$	
Swei et al. (2015)			$\checkmark$	$\checkmark$	$\checkmark$	
Guo et al. (2019)			$\checkmark$	✓		$\checkmark$
Santos et al. (2019)						$\checkmark$

#### 2.3 Limitations and Contribution

Gaps in the pavement LCCA include the inability to consider many relevant sources of uncertainty in optimization-based approaches and the limited number of alternatives available in probabilistic LCCA studies. Most existing optimization-based approaches also face the challenge of requiring a model to navigate from one pavement condition to the next. These models are dependent on a pavement's location, in which weather conditions, traffic rates and other factors affect pavement deterioration. Thus, optimization-based models provide suboptimal results to agencies with different parameters.

To address the need for a LCCA model that: (a) considers many relevant sources of uncertainty, (b) explores the available decision space succinctly and, (c) adapts easily to different model formulations, a new optimization-based approach for the life-cycle planning of pavement infrastructure is presented. Specifically, this study implements several important sources of uncertainty into a model-free reinforcement learning (RL) algorithm to efficiently determine the long-term cost of alternative pavement construction and design choices and their timings. Unlike previous model-based approaches, the proposed model-free RL algorithm can determine an optimal investment strategy without the need to predefine the evolution of pavement degradation, traffic volume, construction costs, or other uncertain factors. Rather, the agent (i.e., decision-maker) iteratively learns the rewards of its actions through direct interactions with the modeling environment.

The proposed RL framework is highly generalizable and flexible; it not only accommodates numerous sources of uncertainty, but also does so without imposing restrictions on how such uncertainties must be modeled. This freedom of structure brings tremendous practical value to the approach in the pavement management community. Since planning agencies differ in their assumptions around relevant LCCA inputs (e.g., pavement degradation), this framework can support many transportation agencies. The importance of this approach is subsequently highlighted through a realistic case study, demonstrating that the model-free RL algorithm leads to the selection of a design and construction policy that significantly reduces life-cycle costs. The previously mentioned limitations and gaps in the literature motivate the creation of a new, optimization based LCCA and LCA that incorporate several sources of uncertainty and their possible structures. The design of this tool feeds into Chapter 3 and 4, which discuss deterioration modeling and its results, and reinforcement learning, respectively.

### **Chapter 3. Deterioration Modeling**

As mentioned in Chapter 1, transportation agencies increasingly rely on pavement management systems (PMS) to guide their maintenance and design decisions (Torres-Machi et al. 2018; Wang et al. 2018). A major benefit of PMS frameworks is that they allow agencies to be proactive, rather than reactive, around infrastructure maintenance (Ng et al. 2011; Su et al. 2017). By being proactive, decision-makers are able to reduce the long-term maintenance cost and enhance the sustainability and safety of their own assets. Taking a proactive approach around maintenance decisions necessitates that an agency regularly procures infrastructure condition information and subsequently uses relevant statistical methods to estimate pavement performance (i.e., deterioration) models (Karlaftis and Badr 2015; Hong and Prozzi 2010). A low-fidelity deterioration model will lead to the selection of a sub-optimal construction design and maintenance schedule (Swei et al. 2018).

Deterioration mechanisms in pavements include cracking, faulting and rutting (Jang et al. 2017) and appropriate maintenance treatments are dependent on how these mechanisms evolve (Swei et al. 2016). Roughness is one of the major factors that influences a pavement's ride quality (Jiang and Li 2005). Thus, planning agencies will usually focus on an infrastructure's composite condition indices or specific distress measures, such as the international roughness index (IRI) that specifically measures roughness, as it is easier for agencies to collect (Chen et al. 2016; Park et al. 2007; Lea and Harvey 2004).

The IRI, expressed as a ratio, "represents the cumulative displacement of the axle with respect to the frame of a reference quarter-car per unit distance traveled over the pavement profile at a constant speed of 80.5 km/h (50 mi/h)" (Dalla Rose et al. 2017). The IRI was developed by

the World Bank and has been used to evaluate ride quality, evaluate the environmental impact of pavement conditions and estimate vehicle operating costs. As the pavement age increases, the condition decreases and the IRI number increases (Jiang and Li 2005). Most transportation agencies use IRI to measure pavement performance in their PMS; for this reason, this chapter focuses on applying its methods towards the development of deterioration models for pavement roughness (a continuous condition indicator) given: (a) its ubiquitous use in projectlevel and network-level tools and, (b) the methods employed can still easily extend to other distress mechanisms.

Over the last several decades, there has been considerable progress towards improving the deterioration models underlying existing PMS tools. While these models are discussed at greater length in the following section, these approaches are briefly distilled here to motivate this work. Previous studies have deployed various statistical approaches to predict pavement performance for both discrete, ordinal measures (e.g., Madanat et al.1995) and continuous indicators (e.g., Hong and Prozzi 2010) of pavement condition. Discrete, ordinal measures refer to measurements that have a relative ordering, such as 1 for "poor condition" and 9 for "excellent condition" (Madanat et al. 1995).

Since variation in pavement performance across facilities is, at least partially, explained by factors such as pavement age and exposed traffic volume, these methods will oftentimes account for these covariates via their model specification. The remaining unexplained variation in pavement deterioration is subsequently treated as aleatory. For the remainder of this thesis, the unexplained variation specifically in pavement deterioration across time will be referred to as aleatory uncertainty.

Correct measures of aleatory uncertainty are paramount to enhance the utility of a PMS and ensure that the planning community derives statistically robust deterioration models. Measures of aleatory uncertainty are tremendously important for pavement engineers, who use them within reliability analysis frameworks to assess the probability of failure within a given time horizon (Thyagarajan et al. 2011). In addition, LCC (Salem et al. 2003; Mishalani and Gong 2008; Ng et al. 2011; Pittenger et al. 2011) and LCA (Noshadravan et al. 2013) frameworks require these estimates to compute the probabilistic cost and/or environmental impacts of alternative investments. Furthermore, incorrect measures of aleatory uncertainty have nontrivial effects on the distribution for parametric estimates, which are paramount for statistical inference (Altman and Bland 2005). Simply put, correctly measuring aleatory uncertainty is invaluable for deriving high-fidelity pavement performance models and PMS frameworks.

Accurately estimating the aleatory uncertainty underlying pavement performance models, however, requires that researchers account for a second source of uncertainty in their statistical analysis: the measured condition of a facility. While agencies may have access to field measurements of distress for their individual assets, the true distress of each facility is generally an unknown quantity due to the inherent errors arising from measurement technology, inspectors, data processing, and other possible circumstances (Kobayashi et al. 2012). Discretized ordinal measures of pavement condition that are embedded in Markov decision process (MDP) PMS frameworks are, furthermore, latent in nature. In other words, the measured condition of a facility is unknown and is simply inferred through direct measurements of distress indicators (e.g., cracking) (Madanat et al. 1995). These realities have motivated an important body of literature that has emerged over the last 30 years around the estimation of infrastructure performance models given the uncertainty in condition

assessments. These methods have primarily been applied to discrete, ordinal measures of infrastructure condition (Ben-Akiva and Ramaswamy 1993; Kobayashi et al. 2012; Madanat et al. 1997; Madanat et al. 1995). A much smaller number of studies (namely Chu and Durango-Cohen 2007, Chu and Durango-Cohen 2008a, and Chu and Durango-Cohen 2008b), have focused on accounting for measurement errors for continuous condition metrics (e.g., pavement roughness) through state-space models.

The focus of this chapter is on continuous condition indicators given that modern technologies and the state-of-the-art are allowing agencies to access this type of information more widely (Bridgelall 2014; Dennis et al. 2014). In particular, practitioners are transitioning to AASHTO's mechanistic-empirical pavement design guide (MEPDG) to: (a) model pavement performance and, (b) incorporate those predictions within their decision-support tools for continuous distress indicators (e.g., pavement roughness) (Li et al. 2011). The implementation of mechanistic-empirical approaches by practitioners will frequently rely on ordinary least squares (OLS) to estimate deterioration models for panel data, unfortunately ignoring the effects of measurement uncertainty (Li et al. 2011). The increasing availability of regularly procured, continuous condition data for infrastructure assets, the growing reliance on continuous distress indictors to manage paving assets, and the prevailing use of least squares in practice situate the research well to enhance pavement management practice. In particular, the proposed methodology will allow decision-makers to better estimate the aleatory uncertainty underlying their PMS tools, including risk-based LCA and LCCA models, leading to the selection of improved design, construction, and maintenance strategies.

This proposed methodology involves the development an iterative reweighted least squares (IRLS) approach for performance panel data. IRLS is an iterative approach to find an optimal

solution, where a weighted least square problem is solved in every step. The goal of this method when applied to this research is to calculate the variance of aleatory uncertainty. A solution is reached when the absolute difference between the estimate of aleatory variance in the current iteration and the previous iteration is less than a user-defined threshold. By calculating the variance of aleatory uncertainty, measurement uncertainty is captured as well. The current method, which includes OLS, is unable to deconvolve these two uncertainties as the variance of degradation uncertainty may have unequal scatter, due to varying sample sizes. OLS assumes that the scatter of data points is equal; the error variance is assumed to be constant. However, IRLS can handle unequal scattering of variance as the method does not assumed the weights of the data points are equal, which is an assumption made if scattering is equal. By using IRLS to obtain a more accurate measure of the variance that underlies pavement deterioration uncertainty, a more effective rehabilitation policy can be applied that may reduce rehabilitation frequency.

#### **3.1 Existing Pavement Degradation Models**

A transition probability matrix defines the probability of a change in the distress level of a facility between year t,  $D_t$ , and t+1,  $D_{t+1}$ . The transition probability matrix will typically be defined by s discretized distress levels, and it is assumed that between years a pavement can only transition with nonzero probability, p, to its current condition, d, and one condition worst, d+1 (Madanat and Ibrahim 1995):

$$P(D_{t+1}|D_t) = \begin{bmatrix} p_{1,1} & p_{1,2} & 0 & 0 & 0 & 0 \\ 0 & p_{2,2} & \ddots & 0 & 0 & 0 \\ 0 & 0 & \ddots & p_{d,d+1} & 0 & 0 \\ 0 & 0 & 0 & p_{d,d} & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & p_{s-1,s} \\ 0 & 0 & 0 & 0 & 0 & p_{ss} \end{bmatrix}$$
(2)

The important advantage of modeling pavement deterioration as both discrete and Markovian is that it supports solving the optimal maintenance schedule for a facility via dynamic programming (Madanat 1993). Dynamic programming requires transitional probabilities to move from one state to the other. Early state-based approaches would account for heterogeneity (the performance difference between different facilities) in pavement deterioration across facilities by first classifying facilities into groups with similar characteristics (e.g., exposed traffic volume) and subsequently estimating a separate transition probability matrix for each grouping (Carnahan et al. 1987). More recent state-based research, however, has employed advanced econometric and statistical analyses to deal with this issue. These methods for state-based models include approaches such as ordered probit (Madanat et al. 1995; Madanat et al. 1997) and Poisson regression coupled with its negative binomial extension (Madanat and Ibrahim 1995). These state-based methods, conceptually, overlap with the time-based models, where the latter type of models aims to estimate the probabilistic time that it takes for a facility to enter and subsequently leave each discrete condition level. For time-based models, hazard rate functions have been successfully estimated (e.g., Mishalani and Madanat 2002) that explicitly account for the heterogeneity across infrastructure facilities. It is important to note that the state-based and time-based methods are directly related; a statebased transition probability matrix can be determined from the hazard rate functions for a timebased model (Mishalani and Madanat 2002).
There has also been tremendous progress within the pavement management community around modeling the performance of continuous distress indicators such as rutting, longitudinal and transverse cracking, pavement roughness and other measures. These performance models are frequently classified into three types: empirical, mechanistic, and mechanistic-empirical (Prozzi and Madanat 2003). Empirical models generally rely on panel data, where pavement performance is described via a series of explanatory factors based solely on statistical considerations. Mechanistic models, on the other hand, are derived from response functions that characterize the relationship between performance and material design. More recently, research has emphasized mechanistic-empirical (ME) performance models to integrate the statistical rigor underlying the empirical methods with the mechanistic approach (Prozzi and Madanat 2003; Li et al. 2011). While the ME approach has gained traction amongst researchers and practitioners, both it and the purely empirical methods require the utilization of sound statistical approaches.

The estimation of continuous performance models within the existing literature frequently relies on least squares and its application to panel data (Aguiar-Moya et al. 2011). While practitioners will typically use ordinary least squares to estimate a pavement performance model, variants of least squares including generalized least squares (GLS) and 2-stage least squares (2SLS) have been implemented by researchers to deal with its shortcomings (Aguiar-Moya et al. 2011; Meegoda and Gao 2014; Hong and Prozzi 2013). Reasons for implementing GLS and 2SLS include the need to account for random effects as well as endogeneity in pavement performance models. In other instances, maximum likelihood methods have been applied to estimate deterioration models for pavements (Hong and Prozzi 2010; Hong and Prozzi 2013). While the above discussion is inherently limited and does not fully cover the

scope and breath of research in this domain, it does highlight the need for statistically sound approaches to estimate performance models for both discrete and continuous condition measures. Despite their differences, the previously discussed approaches provide pavement engineers a mechanism to capture the aleatory uncertainty underlying the deterioration of pavement facilities. Pavement engineers and researchers have successfully embedded these estimates of aleatory uncertainty within their project-level (Noshadravan et al. 2013; Swei et al. 2013; Zhang et al. 2012) and network-level (Sathaye and Madanat 2012; Medury and Madanat 2013) decision-support tools to derive cost-effective and environmentally sustainable resource allocation policies. The importance of these aleatory uncertainty measures is of particular significance given that agencies are now frequently leveraging probabilistic, reliability-based methods to guide their pavement investment decisions (Harvey et al. 2012).

The estimation of aleatory uncertainty and, furthermore, the ability to generate robust parametric models requires the explicit consideration of uncertainty in facility condition assessments. While the models used in practice may frequently ignore this concern, there is a rich and important body of literature around measurement errors (e.g., Humplick 1992) and latent distress measures (e.g., Ben-Akiva and Ramaswamy 1993) and their incorporation within performance models dating back 30 years. Studies centered on discretized, ordinal measures of pavement distress have accounted for the latent nature of infrastructure performance via approaches such as simultaneous equation (Ben-Akiva and Ramaswamy 1993) and ordered-probit (Madanat et al. 1995; Madanat et al. 1997) models. A much smaller group of studies (Chu and Durango-Cohen 2007; Chu and Durango-Cohen 2008a; Chu and Durango-Cohen 2008b), have accounted for measurement errors for continuous distress indicators via state-space models. In this regard, the presented work is differentiated from the

latter group of studies by demonstrating that a GLS approach for continuous infrastructure condition data, which aligns well with the least squares approach used in practice and amongst researchers, can address concerns around measurement error. In the following section, the proposed approach is highlighted, and the case study application is briefly described.

## 3.2 Methodology

As shown by previous research (Swei et al. 2018), pavement deterioration generally follows the below structural model:

$$\Delta D_t = f(\mathbf{Z}_{t-1}) + \varepsilon_t \tag{2}$$

 $\Delta$  is the first difference operator,  $D_t$  is the true condition of a pavement facility in year t,  $\mathbf{Z}$  is a vector composed of k attributes (e.g., exposed traffic volume) that helps explain some of the variation in pavement performance across time, and  $\varepsilon_t$  is an error term that captures the aleatory uncertainty in the evolution of pavement distress. It is assumed that the random error term is unbiased and follows a Gaussian distribution with a variance of  $\sigma_a^2$ . Practitioners will traditionally specify the above model per OLS regression with the root mean square error (RMSE) used to attribute all of the variance in the residuals to the aleatory uncertainty term.

This assumption, however, only holds true if changes in pavement condition between years,  $\Delta D_t$ , are fully known. In reality, practitioners will estimate Equation 2 using the sample average of multiple distress measurements for an individual segment in years *t* and *t*-1. Assuming that the measured condition of a facility is normally distributed, it is well known that the underlying distribution for the sample average follows:

$$\overline{D}_t \sim N\left(D_t, \frac{\sigma_{m_t}^2}{k_t}\right)$$
(3a)

$$\overline{D}_{t-1} \sim N\left(D_t, \frac{\sigma_{m_{t-1}}^2}{k_{t-1}}\right)$$
(3b)

In other words, while the sample mean is equal to the population mean, its underlying variance is the ratio of: (1) the variance in the measured condition of a facility,  $\sigma_m^2$ , in year *t* and *t*-1 and, (2) the number of measurements, *k*, for each time period. Equation 3a and Equation 3b can alternatively be written as:

$$\overline{D}_t = D_t + u_t ; \ u_t \sim N\left(0, \frac{\sigma_{m_t}^2}{k_t}\right)$$
(4a)

$$\overline{D}_{t-1} = D_{t-1} + u_{t-1}; \ u_{t-1} \sim N\left(0, \frac{\sigma_{m_{t-1}}^2}{k_{t-1}}\right)$$
(4b)

If the true condition for a single pavement facility is probabilistic and, furthermore, a model is estimated using the average condition rating for a facility, then Equation 2 should actually be:

$$\Delta \overline{D}_t = f(\mathbf{Z}_{t-1}) + \varepsilon_t \tag{5a}$$

$$D_t + u_t - D_{t-1} - u_{t-1} = f(\mathbf{Z}_{t-1}) + \varepsilon_t$$
(5b)

$$\Delta D_t = f(\mathbf{Z}_{t-1}) + \varepsilon_t - u_t + u_{t-1}$$
(5c)

Equation 5c highlights a fundamental flaw underlying several performance models for continuous distress indicators found in practice. Specifically, there are in fact two sources of variance in addition the aleatory uncertainty: (1) the measured condition of a facility in year *t*-

1 and, (2) the measured condition of that same facility in year *t*. While measurement errors in the explanatory factors, **Z**, would cause downward bias in a parameterized model (Hutcheon et al. 2010), since the expectation of the error terms  $\varepsilon_t$ ,  $u_t$ , and  $u_{t-1}$  is equal to zero, then inflation in the variance (and its effect on the standard errors) for the above model is the main concern. Because the uncertainty around the measured condition for a facility between years should be uncorrelated with one another and, furthermore, field measurements should have no effect on the pavement deterioration process, then the resulting variance for Equation 5 should simply be the summation of the variance for the individual error terms:

$$\sigma_{\Delta \bar{D}_t}^2 = \frac{\sigma_{m_t}^2}{k_t} + \frac{\sigma_{m_{t-1}}^2}{k_{t-1}} + \sigma_a^2 \tag{6}$$

Per Equation 6, pavement engineers can reduce the variance underlying a specified deterioration model via three mechanisms. First, they can leverage improved field measurement technology to evaluate the condition of a facility (Rada et al. 1997). Alternatively, municipalities and planning agencies may prefer to leverage connected vehicles (Dennis et al. 2014), providing decision-makers with a larger sample size of measurements. Finally, planners can use alternative model structures to better capture the seemingly aleatory uncertainty around pavement deterioration via relevant explanatory factors.

Having briefly described the fundamental issue underlying existing pavement deterioration models for a single facility, the approach is synthesized to estimate pavement performance using available panel data. Unlike Equation 2, panel data provide pavement engineers with important information around the performance of many facilities across time:

$$\Delta \overline{D}_{it} = f(\mathbf{Z}_{i,t-1}) + \varepsilon_{it} \tag{7}$$

The only difference between Equation 2 and Equation 7 is the introduction of the variable *i* to index the individual pavement facilities. It can generally be assumed that the aleatory uncertainty term is independent and identically distributed across space and time (as is the typical assumption for regression models) with variance  $\sigma_a^2$ . Conversely, the measurement variance terms,  $\sigma_{m_{i,t-1}}^2$ , for facility *i* in year *t* and *t*-1 across the full panel dataset may vary due to improvements in infrastructure condition assessment technologies and/or external factors (e.g., inclement weather). Under these two realistic assumptions, the underlying variance for the individual samples of a specified model would be:

$$\sigma_{\Delta \bar{D}_{it}}^2 = \frac{\sigma_{m_{it}}^2}{k_{it}} + \frac{\sigma_{m_{i,t-1}}^2}{k_{i,t-1}} + \sigma_a^2 \tag{8}$$

where  $k_{it}$  and  $k_{i,t-1}$  are the number of measured samples for facility *i* in year *t* and *t*-1. Because both the fidelity of the asset condition measurements as well as the sample size used to estimate a facility's condition can vary across a sample set, the underlying variance of Equation 8 will not be homoscedastic, or having an equal scatter, a requirement for OLS regression.

To deal with this issue, which affects the standard errors of a parameterized model, and ultimately estimate the aleatory uncertainty term that enters project-level and network-level tools, an innovative, iterative reweighted least squares (IRLS) approach is highlighted in Algorithm 1. Suppose that a planning agency has collected pavement performance measures across n facilities from year  $t_0$  through year T. A weighted least squares (WLS) approach will

search for some vector of parameters,  $\beta$ , that minimizes the weighted sum of squared residuals for a model:

$$\arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=t_{0}}^{T} w_{it} (\mathbb{E}[\Delta \overline{D}_{it} | \boldsymbol{\beta}] - \Delta \overline{D}_{it})^{2}$$
(9)

where  $w_{it}$  is the individual weight placed on facility *i* in year *t*. In the case that the residuals for Equation 7 exhibit heteroscedasticity, or an unequal scatter, an OLS solution (in which the weights are equal to one for all observations) is no longer the maximum likelihood estimate for a regression model and, therefore, is inefficient (Mills 2014). To achieve the heteroscedastic maximum likelihood estimate and recover efficiency, the weights for the individual data points are usually set equal to the inverse of their variance (Mills 2014):

$$W_{it} \propto \frac{1}{\sigma_{\Delta \overline{D}_{it}}^2} \tag{10}$$

If the selected weights for each data point is set directly proportional to the inverse of its variance, then one would expect for a large sample set that the expectation of the weighted sum of squared residuals would be:

$$\mathbb{E}\left[\sum_{i=1}^{n}\sum_{t=t_{0}}^{T}w_{it}(\mathbb{E}[\Delta\overline{D}_{it}|\boldsymbol{\beta}] - \Delta\overline{D}_{it})^{2}\right] = n(T - t_{0} + 1)$$
(11)

Based on the previous discussion, Algorithm 1 presents an innovative technique to characterize the aleatory uncertainty underlying a pavement deterioration model. Algorithm 1 initializes by setting the weight for each individual sample,  $w_{ii}^{0}$ , equal to the inverse of the measurement

uncertainty terms in year t and t-1 for facility i. The weighted least squares problem is subsequently solved to derive a first estimate of the parameter vector,  $\beta^0$ . Of course, the initial estimate of  $\beta^0$  will not be the finalized vector of parameter estimates as the initial weighting assumes that the variance for the aleatory uncertainty term is equal to zero. By utilizing the relationship in Equation 11,  $\sigma_a^2$  is initially optimized,  $\sigma_a^2(0)$ , by searching for a value that minimizes the absolute difference between the actual weighted sum of squared residuals and its expected value. To expedite the optimization process, values of  $\sigma_a^2(0)$  are assumed to be bounded between zero (i.e., zero variance) and an upper bound limit,  $\sigma_{b.}^{2}$  The upper bound value is simply the solution for the OLS problem, which assumes that all of the variance in the residuals is attributed to aleatory uncertainty. Once initialized, Algorithm 1 enters a loop that: (1) updates the weights, (2) solves for the parameter vector that minimize the WLS problem, and, (3) generates a new set of weights to meet the result of Equation 11. When the absolute difference in the estimate of  $\sigma_a^2$  between iteration j and j-1,  $\sigma_a^2(j)$  and  $\sigma_a^2(j-1)$ , are less than some value  $\delta$ , the model immediately exists the loop under the assumption that the estimates of  $\sigma_a^2$ and  $\beta$  have stabilized. Although this algorithm does not detail any diagnostic tests, it is ensured that some of the important assumptions (e.g., measurement uncertainty is Gaussian distributed) underlying the discussion are met.

Initialize:

Step 1: Set the initial weights  $(w_{it}^0)$  for each measurement:  $w_{it}^0 = \frac{1}{\frac{\sigma_{mit}^2}{k_{it}} + \frac{\sigma_{mit-1}^2}{k_{it-1}}}$ Step 2: Solve for the vector  $\boldsymbol{\beta}^0$ :  $\arg\min_{\boldsymbol{\beta}^0} \sum_{i=1}^n \sum_{t=t_0}^T w_{it}^0 (\mathbb{E}[\Delta \overline{D}_{it} | \boldsymbol{\beta}^0] - \Delta \overline{D}_{it})^2$ 

Step 3: Generate a first estimate of the aleatory uncertainty term:

- a. Bound the feasible aleatory variance,  $\sigma_b^2$ , by the variance for the OLS solution
- b. Solve the below optimization problem:

$$\min \left| \sum_{i=1}^{n} \sum_{t=t_{0}}^{T} \frac{1}{\frac{\sigma_{mit}^{2}}{k_{it}} + \frac{\sigma_{mit-1}^{2}}{k_{it-1}} + \sigma_{a}^{2}(0)} (E[\Delta \overline{D}_{it} | \boldsymbol{\beta}^{0}] - \Delta \overline{D}_{it})^{2} - n(T - t_{0} + 1) \right|$$
  
subject to  $0 \le \sigma_{a}^{2}(0) \le \sigma_{b}^{2}$ 

#### For j = 1 to J

**Step 1**: Reset the weights for each measurement as:  $w_{it}^{j} = \frac{1}{\frac{\sigma_{mit}^{2}}{k_{it}} + \frac{\sigma_{mit-1}^{2}}{k_{it-1}} + \sigma_{a}^{2}(j-1)}$  **Step 2**: Solve for the parameters  $\beta^{j}$  per:  $\arg \min_{\beta^{j}} \sum_{i=1}^{n} \sum_{t=t_{0}}^{T} w_{it}^{j} (E[\Delta \overline{D}_{it} | \beta^{j}] - \Delta \overline{D}_{it})^{2}$ **Step 3**: Solve the below optimization problem:

$$\min \left| \sum_{i=1}^{n} \sum_{t=t_{0}}^{T} \frac{1}{\frac{\sigma_{mit}^{2}}{k_{it}} + \frac{\sigma_{mit-1}^{2}}{k_{it-1}} + \sigma_{a}^{2}(j)} (E[\Delta \overline{D}_{it} | \boldsymbol{\beta}^{j}] - \Delta \overline{D}_{it})^{2} - n(T - t_{0} + 1) \right|$$
subject to  $0 \le \sigma_{a}^{2}(j) \le \sigma_{b}^{2}$   
**Step 4:** If  $|\sigma_{a}^{2}(j) - \sigma_{a}^{2}(j-1)| < \delta$  Then  $\boldsymbol{\beta} = \boldsymbol{\beta}^{j}$  And  $\sigma_{a}^{2} = \sigma_{a}^{2}(j)$  Break  
Next  $j$ 

## 3.3 Case Study Analysis

This study applies its methodology to pavement condition information collected as part of the FHWA Long-Term Pavement Performance (LTPP) program (FHWA 2018). Since 1991, the FHWA has funded and managed the program, which has collected and stored pavement performance data for over 2,500 test sections throughout North America. This database includes over 140,000 individual field measurements of pavement IRI from pavement segments in the United States and certain areas in Canada. Because these test sections are

exposed to a broad range of climactic regions, traffic volumes, and other relevant factors, the dataset facilitates the ability to capture some of the explanatory uncertainty underlying pavement performance (Swei et al. 2018). Although FHWA has tracked and collected pavement condition information for a range of relevant distress mechanisms, the methodology is specifically applied to model the IRI of asphalt concrete (AC) pavements due to its ubiquitous use within project and network-level decision-support tools. However, the value of Algorithm 1 is likely greater for other pavement distress mechanisms that are typically subject to much higher levels of measurement uncertainty (Schwartz 2007).

Based on the findings of several previous research efforts (Lee et al. 1993; Swei et al. 2018; Yang et al. 2005), this research assumes that a reasonable amount of the variation in pavement deterioration across time can be explained by the following model structure:

$$\Delta \overline{D}_{it} = \beta AADTT_{t-1}^{\alpha_1} Age_{t-1}^{\alpha_1} SN_{t-1}^{\alpha_3}$$
(12)

where  $\beta$ ,  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  are parameter estimates that quantify the effect of AADTT, pavement age, and the facility's structural number (SN) on pavement deterioration. While more complex models exist, this research has selected a parsimonious model structure that considers some of the most important drivers of variation in pavement performance across facilities and, furthermore, includes the types of explanatory variables typically stored within most PMS databases for state DOTs and municipalities (Lea and Harvey 2004). Of course, the described methodology could easily accommodate a broader set of explanatory factors and quantify its effect on the aleatory uncertainty term. Within the LTPP database, this study uses consecutive field measurements plus or minus one month. It also ignores extreme outliers (i.e., absolute changes in IRI are greater than 0.5 m/km with low measurement error) that may arise due to unreported changes activities that could affect the condition of a pavement facility.

## 3.4 Results

This study begins its analysis by first evaluating the degree to which measurement uncertainty is pervasive within the LTPP dataset. Figure 2 presents a histogram of the standard error (i.e.,  $\sigma_{m_{it}}/k_{it}^{0.5}$  per Equation 8) of the mean IRI condition for each pavement facility stored within the LTPP database. This histogram has been slightly truncated, excluding the 1% of pavement facilities with a standard error greater that 0.093 meters per kilometer (i.e., 5.9 inches per mile). Over 97% of the LTPP sections include exactly five field measurements per site visit and, therefore,  $k_{it}$  and  $k_{i,t-1}$  as defined by Equation 9 are, for the most part, uniform across this study. The average standard error across all facilities is 0.012 meters per kilometer (0.76 inches per mile), implying that, on average, the 95% confidence interval around a pavement's true condition is +/-0.024 meters per kilometer (+/-1.5 inches per mile). While that may not seem like much uncertainty, it is important to note that: (a) the average change in mean IRI across the entire dataset is only 0.04 meters per kilometer (2.5 inches per mile) and, (b) in over 40% of instances, the standard error around the condition of a facility in either year t or t-1 is greater than the mean change in IRI. As a result, it is fair to say that the consideration of measurement uncertainty is important when deriving a performance model.



Figure 2. Histogram of standard error of mean IRI condition (meters per kilometer) for 99% of pavement facilities (excluding those with a standard error greater than 0.093 meters per kilometer) stored within the LTPP database

Figure 3 subsequently plots the average standard error of measured IRI across time. While there have been tremendous improvements to the underlying pavement deterioration models used in practice over the last 20-30 years, there has been no demonstrable shift in the performance of condition assessments conducted as part of the LTPP program. In fact, a simple linear regression comparing the average measurement uncertainty (i.e., standard error) across time suggests that there has been, if anything, a statistically significant increase in measurement uncertainty across time at the 10% level. Finally, there is very little correlation between measurement uncertainty and our relevant regressors as well as the mean condition of a facility in its levels; the absolute value of Pearson's correlation coefficient between the standard error measurement of a pavement facility's condition and these factors never exceeds



Figure 3. Average standard error of mean IRI condition (meters per kilometer) of paving facilities across time

Table 2 presents the parameterized estimates for the proposed IRLS approach and the traditional OLS regression technique (with pavement condition in units of meters per kilometer). For this case study, the threshold value,  $\delta$ , that terminates Algorithm 1 has been set to 10<sup>-6</sup>. Despite this low tolerance, the model converges to an optimal solution within just four iterations. The optimal solution is defined as a solution where the absolute difference between the actual weighted sum of squared residuals and its expected value is less than the threshold value. For both models, if the t-statistic for  $\alpha_1$ ,  $\alpha_2$  and/or  $\alpha_3$  is unable to reject the null hypothesis that it is equal to zero, then that variable is logarithmically transformed per Tukey's ladder of powers. The underlying reasoning for this transformation is the following limit:

$$\lim_{\alpha \to 0} \frac{x^{a} - 1}{a} = \ln x \tag{13}$$

Table 2. Estimate of Equation 12 using (a) the proposed IRLS approach and (b) the typical OLS regression approach. In the case that a given AADTT, Age, or SN are logarithmically transformed, it is denoted via an "LN". T-statistics are listed in parentheses

Parameter	<b>Proposed IRLS</b>	<b>Traditional OLS</b>		
0	3.134*10-3	3.218*10-3		
β	(3.22)	(3.08)		
$\alpha_{l}$	LN	LN		
$\alpha_2$	LN	LN		
	-0.429	-0.457		
$\alpha_3$	(-2.08)	(-2.11)		
$\sigma_a^2$	5.785*10 <sup>-3</sup>	6.702*10 <sup>-3</sup>		
Sample Size	14	450		

**Model Structure**:  $\Delta \overline{D}_{it} = \beta AADTT_{t-1}^{\alpha_1} Age_{t-1}^{\alpha_1} SN_{t-1}^{\alpha_3}$ **Note**: if  $\alpha_l$ ,  $\alpha_2$  or  $\alpha_3$  are statistically insignificant at the 10%, the relevant regressor is logarithmically transformed

As can be noted from Table 2, both the proposed IRLS and traditional OLS models are unable to reject the null hypothesis that  $\alpha_2$  and  $\alpha_3$  are equal to zero. While the parameter estimates slightly differ between the two models, both estimates suggest that there is a statistically significant upward trend in year-over-year pavement distress across time (as expected). What does differ considerably across the two models, however, is the estimated variance of the aleatory uncertainty term. By applying Algorithm 1, the underlying variance of this term is 14% lower than the estimate generated from an OLS approach that does not account for the effects of measurement uncertainty. This relative effect is likely to be even more important should: (a) a more elaborate pavement performance model be estimated that incorporates a broader set of explanatory factors or, (b) a model be estimated with "noisier" measurement

data. The latter statement will likely be the case for other pavement distress mechanisms and for data procured by local municipalities who rely on manual field measurements of pavement condition.



Figure 4. Q-Q plot comparing the distribution of the weighted residuals (y-axis) to a standard normal distribution (x-axis)

Following the estimation of Table 2, it was evaluated whether the assigned weights for the proposed methodology recovered efficiency. Beyond testing for heteroscedasticity in the weighted residuals, it is also inspected whether the weighted residuals using this technique followed a standard normal distribution with mean of zero (i.e., unbiased) and unit variance. As a reminder, this result would be expected to be the case, given that the weights for each individual measurement should, theoretically, be equal to the inverse of its variance. However,

if either: (a) the aleatory uncertainty term is not constant across facilities and time or, (b) the underlying error terms are not Gaussian distributed, then the theory underlying this study no longer applies and weighted errors would not follow a standard normal distribution. Figure 4 presents a quantile-quantile (Q-Q) plot comparing the distribution of the weighted residuals relative to a standard normal distribution. These two distributions match up incredibly well (even in terms of their tails), indicating that what is being observed in practice corroborates the theory underlying the proposed methodology.

## **3.5 Contributions**

The importance of collecting correct measures of aleatory uncertainty, which is due to random variability, is that it will improve the overall fidelity of risk-based LCA and LCCA tools, leading to the selection of improved design, construction, and maintenance strategies. Pavement deterioration is an important component of an effective PMS. This measurement allows agencies to decide on when their systems need to be maintained. However, pavement deterioration data consist of uncertainties such as aleatory uncertainty and measurement uncertainty, which is caused by measurement equipment and other factors As transportation agencies have limited fiscal resources for pavement management, a construction design and maintenance schedule that specifically relates to the current state of a pavement segment are imperative. While methods such as OLS capture uncertainty present in degradation measurements, it is unable to capture the uncertainty caused by measurement uncertainty. IRLS, the method presented in this chapter, provides a method to deconvolve aleatory and measurement uncertainty. In the future, this can improve the statistical inference underlying current parametric methods. By: (a) characterizing measurement variation with state-of-the-art condition procurement tools and, (b) leveraging sampling theory, other researchers can evaluate the potential merits of low-accuracy, high-frequency measurements to determine the true condition of a pavement asset.

This research ultimately presents a fairly intuitive approach based on sampling theory, which will allow agencies to more easily integrate the consideration of measurement errors in practice. Furthermore, a previous study (Swei et al. 2018) suggests that panel data for pavements are difference-stationary (i.e., the dataset exhibits stationarity once differenced). This reality represents a second, important deviation of this approach relative to the previous models for continuous distress indicators and is accounted for in this chapter via the final model specification. The resulting performance models should allow agencies to better address an important shortcoming in existing decision-support tools. In doing so, this research will help planning agencies generate more robust resource allocation policies and improve their processes for statistical inference when evaluating a parametric model. Finally, by decomposing measurement and aleatory uncertainty, a tertiary contribution emerges: the ability to support parallel research within the connected vehicle space (Bridgelall 2014; Dennis et al. 2014).

The conclusions and limitations of this study can be found in Chapter 5. Chapter 4 discusses RL and applies this model into an LCCA to optimize costs.

# **Chapter 4. Reinforcement Learning**

Reinforcement learning (RL) is a type of machine learning that involves an intelligent program, known as agents, navigating through an environment that may be known or unknown, in order to achieve a given objective (Nandy and Biswas 2018). While machine learning interacts with data via training datasets, reinforcement learning obtains information through the environment. It can also alter the environment and its responses, based on the sequences of actions taken by the agent (Saito et al. 2018). Agents receive feedback from the environment with every action taken. This may be positive feedback, known as rewards, or negative feedback, known as penalties, depending on how the selected action affects the agent's "distance" from an objective (Nandy and Biswas 2018). The goal of reinforcement learning is to maximize the objective, or increase the rewards received, which is done by choosing actions to move to the next state. A state is defined as the minimum information needed in order to make an informed decision (Powell 2009) while the decision variable characterizes the decision-maker's available "actions" at any point in time. For example, if the goal of an agent is to reach the end of a maze with the fewest possible actions, the objective would be to maximize the points that an agent can earn from the environment. Positive feedback could include a point when an agent goes towards the goal and negative feedback could include a partial point deducted when an agent makes counterproductive actions (i.e., going to a previously visited state). Through trial and error and simulations, the agent would find the optimal path.

Popular applications of reinforcement learning include learning in games such as backgammon, learning in robotics and operations research problems such as vehicle routing and maintenance problems (Szepesvári 2010). There are many types of reinforcement learning algorithms, differentiating in the way the value function is updated; off-policy and on-policy

(Singh et al. 2000). Off-policy algorithms can update estimated value functions based on possible actions, not necessarily the action that had been chosen by the agent. On-policy algorithms only update the estimated value function based on the actions executed by the agent. This means for an on-policy algorithm to converge; it is highly dependent on the learning policy. Reinforcement learning algorithms can also be model-based or model-free. Model-based algorithms (i.e., value iteration) learn the model of the system with a policy and value function while model-free, e.g., Temporal Difference (TD) learn a value function of policy without modeling the system (Doya et al. 2002).

While probabilistic PMS tools that implement several sources of uncertainty exist, as highlighted in Chapter 2, they are unable to consider multiple rehabilitation alternatives. The alternative is to use an optimization-based tool that will evaluate alternatives in order to achieve the chosen objective. However, these optimization-based tools are restrictive in their model structures. For example, deterioration models are depicted differently for various agencies, as factors such as location, traffic vary. These tools are also computationally intensive due to the number of input parameters and uncertainties involved, which makes it unrealistic to use for large-scale pavement networks.

For this application, *Q*-learning, a model-free RL approach, was used to solve this problem. *Q*-learning is an intuitive approach to learn an optimal policy, which has been utilized in games as well as pavement approaches in Durango-Cohen (2004). A *Q*-table, consisting of a column for each action available to the agent and a row for every state the agent can be, stores the learned rewards/costs for different state-action pairs. For every state, there is a "cost-to-go" to the following state depending on the action chosen. Because the optimization problem is based on cost minimization, the optimal policy for a given state is the action with the lowest value. The table updates based on the Bellman equation, which considers the cost provided by the action, a learning rate (the percentage of the cost calculated that is stored in the *Q*-table) and minimum future reward value stored in the next state-action pair (Even-Dar and Mansour 2003). Over runs, the algorithm begins to converge as every feasible state-action pair is explored and an optimal policy is reached.

#### 4.1 RL Approaches in Relevant Applications

Chapter 2 discussed non-optimization and optimization applications in pavement management; the focus of this section is methods that utilize reinforcement learning. In the context of pavement LCCA, only Medury and Madanat (2013) and Durango-Cohen (2004) have examined the benefits of a model-free approach. In the case of Medury and Madanat (2013), their model-free RL algorithm incorporates network-based constraints into pavement management systems; although network considerations are of value for city-wide and/or statewide LCCAs, more refined project-specific analyses are beneficial for critical, high cost infrastructures such as bridges, tunnels, and interstate highways. The case study by Medury and Madanat (2013) for a relatively small network demonstrates the potential feasibility and utility of a model-free approach, but it only accounts for one source of uncertainty (pavement deterioration). This fact makes it difficult to generalize the ability of model-free, optimizationbased approaches to improve decision-making for infrastructure management in the presence of many possible sources of uncertainty. Durango-Cohen (2004) applies multiple reinforcement learning methods such as state-action-reward-state-action (SARSA) and Qlearning to determine an optimal management policy for pavement projects subject to uncertainty in deterioration. However, because: (a) pavement deterioration uncertainty is idealized as Markovian and, (b) the paper does not account for other possible sources of variation, it is difficult to glean insights around the applicability of these approaches to more realistic problems faced by practitioners at the project level.

Reinforcement learning approaches based on *Q*-learning algorithms have been leveraged in numerous construction and transportation applications; Ozan et al. (2015) propose a modified RL algorithm for optimizing signal timings in signalized networks, and their algorithm is able to optimize the solution regardless of different link flow conditions. To address the inadequacy of passenger inflow control in urban rail transit, Jiang et al. (2018) develop a novel O-learning approach that significantly improves inflow volume while minimizing safety risks. In the automation of earth-moving machines, Dadhich et al. (2016) provide arguments in support of model-free RL methods over programming-based methods, where the value of reinforcement learning lies in its potential to adapt to variations in machine and material, as well as different performance metrics. Liu et al. (2020) successfully implement a multi-agent, Q-learning algorithm to reduce clash collisions in the planning of steel reinforcement in construction projects. Motivated by the unavoidable limitations of parametric models, Mao and Shen (2018) also demonstrate the advantages of reinforcement learning as a non-parametric, model-free method for finding an optimal routing policy. Mao and Shen (2018) address the deficiencies in parametric routing problems, such as the strong assumptions that must be made to allow for efficient solutions; they note that these assumptions may be difficult to validate in real networks, where distributions vary significantly. The shortcomings of such parametric models similarly exist in the context of pavement management and LCCA. Moreover, Mao and Shen (2018) emphasize the benefits of reinforcement learning: the learning agent relies purely on the statistical knowledge gathered through direct interactions with the environment, eliminating the need for prior knowledge of the transition and reward models. Therefore,

through numerous simulations, the entirely model-free algorithm efficiently learns the longterm impact of its choices.

The advantage of using tabular *Q*-learning, which has been applied in the recently cited literature and is the approach used in this study, relative to function approximation techniques lies in its accessibility. Value function approximation methods, which may include the use of neural networks, are well-situated for large-scale problems in which it is either unnecessary or computationally infeasible to implement the tabular version of reinforcement learning (Mnih et al. 2015; Memarzadeh and Pozzi 2019). However, a drawback of these approaches is that they are both difficult to interpret and, furthermore, will not necessarily lead to improved decisions for the size of problems addressed in this study (Verma et al. 2018). Further information around the problem size is provided in the upcoming methodology section. As the goal of this research is to develop a tool that both improves asset management decisions and is accessible for practitioners, the priority is the interpretability of results in choosing to implement a *Q*-learning algorithm.

In the following sections, the potential value of this approach for planning agencies is highlighted via three case studies that simultaneously considers three sources of uncertainty: pavement deterioration, traffic volume, and construction costs (both immediately and their evolution in the future).

#### 4.2 Methodology

#### 4.2.1 *Q*-learning Algorithm for the Optimization of Pavement LCCA

*Q*-learning is a commonly used, model-free RL algorithm to efficiently determine optimal sequential decisions subject to uncertainty. In the context of pavement management systems, the decision variable may include available pavement designs, maintenance and rehabilitation,

and preservation treatments. Relevant information for the state variable could be a facility's condition, structural design, exposed traffic volume, and/or other pertinent factors. By creating a modeling environment where relevant uncertainties are simulated, Q-learning allows the agent to: (a) learn the long-term rewards/costs of its actions, a, under different states, s, and, (b) develop a set of policies (i.e., decision-rules) to best achieve its objective.

Algorithm 2 synthesizes the *Q*-learning algorithm used as part of this research, with further details provided in the case study section. The algorithm begins by assuming that no state-action pair is preferable. Namely, it sets the cost of each state-action pair,  $Q^0(s_t, a_t)$ , across all available actions, *A*, and possible states, *S*, equal to zero from the beginning of the analysis to one year before the end of the analysis period, *T*–1 (Step 1). Once initialized, the algorithm enters a learning phase; over *n* iterations, it learns the optimal action for different possible states that the system could enter.

For each iteration, *i*, and time period, *t*, actions are taken by the agent via two approaches: exploration or exploitation. The trade-off between exploration and exploitation is an important consideration in reinforcement learning; exploitation, the fundamental basis of reinforcement learning, refers to the selection of the optimal action based on previously acquired knowledge (Zhu et al. 2018). In this context, this refers to the action,  $a_t$ , with lowest anticipated cost based on the *Q*-table for the current state,  $s_t$ . Conversely, exploration concerns the search for new knowledge and potentially better actions, and it is essential for making sufficiently insightful and informed decisions (Zhu et al. 2018). A straight-forward approach to promote exploration is an  $\varepsilon$ -greedy policy, where the draw of a standard uniform random variable, *u* (Step 2), and the value of  $\varepsilon$  (ranging from zero to one) help determine the agent's course of action. The first option is that a random action (i.e., exploration),  $a_t^i$ , is selected from *A* with probability  $\varepsilon$  (Step 3a). Alternatively, with probability  $1-\varepsilon$ , the algorithm selects the action with the minimum expected cost (i.e., exploitation) for the current state,  $s_t^i$ , based on the information learned up until simulation i-1,  $Q^{i-1}(s_t^i, a_t)$  (Step 3b). Higher values of  $\varepsilon$ , therefore, lead to further exploration of the feasible state-action space (Tokic 2010). The RL algorithm sets  $\varepsilon$  equal to 0.1 at the start of the analysis (i.e., t = 0) and 0.01 for all future years (i.e., t > 0). Exploration is emphasized at the start of the analysis given the importance of the initial decision; the value of  $\varepsilon$  for the 75-year analysis period case study is equivalent to not purely exploiting in future years across roughly 50% of the learning iterations.

Algorithm 2. Pseudocode for Q-learning algorithm with an  $\varepsilon$ -greedy policy

<b>Step 1</b> : $Q^0(s_t, a_t) = 0 \ \forall \ a \in A; \ \forall \ s \in S; \ \forall \ t = 0, \ 1, \ \dots, \ T-2, \ T-1$
For $i = 1$ to $n$
For $t = 0$ to $T-1$
Step 2: Generate $u \sim unif(0, 1)$
<b>Step 3a:</b> If $u < \varepsilon$ Then Select a random action, $a_t^i$ , where $a \in A$
<b>Step 3b: Else</b> $a_t^i = \arg\min_{a \in A} Q^{i-l}(s_t^i, a_t)$
<b>Step 4:</b> Simulate the next state of the system $s_{t+1}^i \sim f(s_t^i, a_t, \omega_t^i)$ where $\omega \in \Omega$
<b>Step 5:</b> Compute updated cost of selected action $q_t^i = C(s_t^i, a_t^i) + \gamma \min_{a \in A} Q^{i-1}(s_{t+1}^i, a_{t+1})$
<b>Step 6:</b> Update <i>Q</i> -value $Q^{i}(s_{t}^{i}, a_{t}^{i}) = \eta q_{t}^{i} + (1-\eta)Q^{i-1}(s_{t}^{i}, a_{t}^{i})$
Next t
Next i

Once an action is selected, the algorithm simulates the evolution of the system in the next time period (Step 4), computes an updated cost for the selected action (Step 5), and updates the values in the *Q*-table (Step 6) before moving to the next iteration, *i*+1. The state of the system in the next time period,  $s_{t}^{i}_{t+1}$ , is a function of its previous state, the selected action, and any uncertainty,  $\omega_{t}^{i}$  (e.g., future traffic volume) that becomes known at time *t*+1 (Step 4). As mentioned earlier, an important advantage of this approach is that the algorithm needs neither the form of the function, *f*, nor the underlying distribution of the sources of uncertainty,  $\Omega$ , to make an informed decision. Once the algorithm steps into the next time period, the updated cost of the previously selected action,  $q_t^i$  is easily computed as the sum of the immediate cost of the selected action for the given state,  $C(s_t^i, a_t^i)$ , and some discounted future cost,  $\gamma$ , using the estimated *Q*-values from simulation *i*–1 (Step 5). The *Q*-table updates its values using a learning rate,  $\eta$ , which is bounded between zero and one (Step 6). Specifically, the new estimate of  $Q^i(s_t^i, a_t^i)$  is a weighted sum of the cost of the selected action in simulation *i*,  $\eta q_t^i$ , and the previously estimated *Q*-value,  $(1-\eta)Q^{i-1}(s_t^i, a_t^i)$ . For this study,  $\eta$  is set to 0.05, slightly below the 0.1 value frequently used in practice (Powell 2009). While using a lower learning rate may delay convergence in some applications, across the case studies convergence occurred in less than a minute with a fairly basic, 4-core personal computer.

Once the learning algorithm converges on a finalized *Q*-table, it uses the stored values as of iteration *n*,  $Q^n(s_t^n, a_t)$ , to generate a set of policies (i.e., decision-rules) that will guide its choices within a more traditional probabilistic LCCA model. As shown in Algorithm 3, over *m* Monte Carlo simulations, the algorithm takes an action,  $a_t^j$ , with the lowest associated *Q*-value, based on the *Q*-table (Step 1). Following the selection of an action, an updated LCC is computed at time *t* for simulation *j*, *LCC*(*j*, *t*), that accounts for the discounted cost of the selected action,  $\gamma^t C(s_t^j, a_t^j)$  (Step 2). The state of the system in the next time period,  $s_{t+1}^j$ , is simulated exactly as in Algorithm 2 until the end of the analysis period, *T* (Step 3). Finally, the algorithm stores the discounted LCC from simulation *j* before entering the next simulation, *j*+1 (Step 4). Both the *Q*-learning environment and the associated probabilistic LCCA model are programmed in Python. In the upcoming section, three specific case studies used to test and validate the value of the proposed approach are detailed, which are based on previously published work by Swei et al. (2015).

Algorithm 3. Pseudocode for probabilistic LCCA model with actions selected based on Q-table

For $j = 1$ to $m$ For $t = 0$ to $T-1$
<b>Step 1:</b> Take an action based on the <i>Q</i> -table $a_t^j = \arg\min_{a \in A} Q^n(s_t^n, a_t)$
<b>Step 2:</b> Compute simulation life-cycle cost $LCC(j, t) = LCC(j, t-1) + \gamma^t C(s_t^j, a_t^j)$
<b>Step 3:</b> Simulate the next state of the system $s_{t+1}^j \sim f(s_t, a_t, \omega_t^j)$ where $\omega \in \Omega$
Next t
<b>Step 4:</b> Output and store the simulated life-cycle cost for simulation <i>j LCC(j, T</i> –1)
Next j

#### 4.3 Description of Case Studies

As highlighted by researchers and practitioners, state departments of transportation (DOT) frequently design their pavements to a specified design life and reliability level (Swei et al. 2015). For a major interstate highway, a typical design life may range from 20 to 30 years with a 90% reliability that the facility will not require maintenance beforehand (Swei et al. 2015). Within pavement LCCA, the design life for a facility will generally be much shorter than the analysis period (i.e., T in Algorithm 1 and Algorithm 2). Between the end of the design life and the analysis period, it is anticipated that the facility will require periodic maintenance and rehabilitation activities, although it is possible that the facility may also need to be reconstructed.

Through three case studies, the proposed *Q*-learning algorithm approach is compared to traditional practice. Similar to the case studies described by Swei et al. (2015), a new 1 lanemile pavement facility is to be constructed for an interstate highway. The interstate highway experiences an initial AADTT of 8,000 vehicles. Uncertainty in the AADTT growth rate, *g*, is assumed to follow a Gaussian distribution,  $N(\mu, \sigma)$ , with mean,  $\mu$ , of zero and standard deviation,  $\sigma$ , of 1%. Pavement deterioration,  $D_t$ , is assumed to follow the published model of Swei et al. (2018), where probabilistic changes in the International Roughness Index (IRI) are a function of the pavement facility's age, AADTT, and structural number (SN). IRI is only incorporated in this study given that it is a ubiquitous pavement distress measure used by both state DOTs and municipalities to guide their investment choices. Future research should evaluate the effects of including other pavement distress mechanisms (e.g., rutting, cracking) on the results of this study. The assumed IRI of the pavement following a major construction activity, IRIconstructed, is 1 m/km (63.6 in/mi). The maximum allowable IRI, IRImax, is 2.36 m/km (150 in/mi). Uncertainty in future construction costs are related to both current price levels (detailed in Table 3) and a price index,  $P_t$ , which captures cost growth over time. Based on a separate analysis of disaggregate bid price data (Oman Systems Inc. 2019), the price index is assumed to include: (1) a constant, 100, for the base year,  $P_0$ , (2) a trend term of 0.014 that captures year-over-year average real price growth and (3) a residual  $\varepsilon_t$ , that depends on one lagged value of itself,  $\varepsilon_{t-1}$ , and a random error term with a mean of zero and standard deviation of 0.065. Should the condition of the facility be less than  $IRI_{max}$  at the end of the analysis period, T, then the asset is assumed to have some positive residual value,  $V_{salvage}$ , that is proportional to the cost of the last construction or maintenance activity, Clast. For simplicity, the assumption is that the asset depreciates linearly with pavement condition. Finally, the real discount rate used as part of the LCCA is 1.5%, consistent with the suggestion of Walls and Smith (1998) to use the current rates provided by the U.S. Office of Management and Budget. Table 3 provides further details on these inputs.

Variable	Input Value
AADTT	8,000
g	N(0, 0.01)
$D_t$	$D_t = D_{t-1} + 0.08 \times \ln age \times \ln AADTT \times SN^{-2.5} + N(0, 0.05)$
<b>IRI</b> constructed	1 m/km (63.6 in/mi)
IRI <sub>max</sub>	2.36 m/km (150 in/mi)
$P_t$	$\ln P_t = 4.6 + 0.014t + \varepsilon_t ; \varepsilon_t = 0.06\varepsilon_{t-1} + N(0, 0.065)$
Discount Rate	1.5%
Vsalvage	$C_{last} \times (IRI - IRI_{max}) / (IRI_{max} - IRI_{constructed})$

Table 3. Key inputs for the probabilistic LCCA model

Table 4 subsequently presents the available actions to the agent: five hot mix asphalt pavement construction designs, a pavement maintenance/overlay, and a choice to "do nothing". The cost for each action is, again, based on bid data sourced from Oman Systems Inc. (2019). Costs are assumed to follow a log-normal distribution,  $Lognormal(\mu, \sigma)$ , which highlights the flexibility with which the non-parametric approach can consider varying structures of sources of uncertainty. Although the age and SN of a pavement facility are reset following a major reconstruction, both variables are assumed unchanged if the agent selects a maintenance/overlay or decides to "do nothing".

Action	Action Type	Cost	Age Post-	SN Post-	IRI Post-
Action	Action Type	Cost	Decision	Decision	Decision
$a^{l}$	New/Reconstruction	$0.7 \times Lognormal(12.46, 0.28) \times P_t / P_0$	0	3. 43	IRI <sub>constructed</sub>
$a^2$	New/Reconstruction	$0.85 \times Lognormal(12.46, 0.28) \times P_t / P_0$	0	4.17	IRI <sub>constructed</sub>
<i>a</i> <sup>3</sup>	New/Reconstruction	$Lognormal(12.46, 0.28) \times P_t / P_0$	0	4.9	IRI <sub>constructed</sub>
$a^4$	New/Reconstruction	$1.15 \times Lognormal(12.46, 0.28) \times P_t / P_0$	0	5.64	IRI <sub>constructed</sub>
$a^5$	New/Reconstruction	$1.3 \times Lognormal(12.46, 0.28) \times P_t / P_0$	0	6.37	IRI <sub>constructed</sub>
$a^6$	Pavement Overlay	Lognormal(11.1, 0.28) $\times P_t / P_0$	$Age_{t-1} + 1$	$SN_{t-1}$	IRI <sub>constructed</sub>
$a^7$	Do Nothing	_	$Age_{t-1} + 1$	$SN_{t-1}$	

Table 4. Probabilistic cost and effect of different available actions

To benchmark the performance of the proposed approach for pavement design and maintenance, the RL algorithm's life-cycle cost is compared to a more traditional approach across three separate case studies. As highlighted by Swei et al. (2015), the traditional approach to pavement LCCA is that an initial construction action is selected with a specific probability of failure (frequently 10% for highway interstates) at the end of its design life. In all future instances, should the facility's IRI exceed  $IRI_{max}$  prior to *T*, a maintenance/rehabilitation activity ( $a^6$  in this problem) is applied to reset the pavement condition to  $IRI_{constructed}$ . As detailed in Table 5, three case studies have been selected from Swei et al. (2015) that utilize design lives and analysis periods representative of those found in practice. Specifically, action  $a^1$  has roughly a 10% probability of failure by Year 20, while action  $a^3$  has a 10% probability of requiring maintenance by Year 30 using the deterioration model and inputs in Table 3. Since this decision-making process appropriately reflects the conventional probabilistic LCCA models highlighted earlier in this paper, it serves as the reference for the proposed algorithm.

 Table 5. Design life and analysis period for case studies as well as initial action per the

 traditional LCCA approach

Case Number	Design Life	Analysis Period	Traditional Approach Initial Action
1	20 Years	50 Years	$a^{I}$
2	30 Years	50 Years	<i>a</i> <sup>3</sup>
3	30 Years	75 Years	<i>a</i> <sup>3</sup>

It is worth highlighting that, per the *Q*-learning algorithm used in this research, the agent can be in anywhere up to 360 states at any point in time. More specifically, the state space is discretized based on three potential pavement conditions ( $D_t < 0.8 \ IRI_{max}$ ,  $0.8 \ IRI_{max} \le D_t < IRI_{max}$ ,  $D_t \ge IRI_{max}$ ), up to eight possible ages (<10, <20, <30, <40, <50, <60, <70, and <80), five SN values (based on Table 4), and finally three different construction price levels ( $P_t < 0.9 \text{ E}[P_t]$ ,  $0.9 \text{ E}[P_t] \le P_t < 1.1 \text{ E}[P_t]$ , and  $P_t \ge 1.1 \text{ E}[P_t]$ ). In reality, however, the agent is typically operating in far fewer states; the most extreme example is at t = 0, where the agent can only be in one possible state given that no uncertain information has yet to enter the system. While some of the important model inputs (e.g., pavement deterioration) may vary across planning agencies, the described case study is representative of the types of problems found in practice. Future research, of course, should validate these findings by testing the proposed algorithm across a range of contexts.

#### 4.4 Case Study Results

Figure 5 presents the estimated Q-value of the best action at t = 0 across the 100,000 learning iterations (i.e., Algorithm 2) for the 50-year (Case Number 1 and 2) and 75-year (Case Number 3) analysis period problems. As illustrated, early iterations display growing anticipated lifecycle costs as the agent learns both the immediate and long-term cost of their actions. After roughly 10,000 iterations for the 50-year analysis period problems and 25,000 iterations for Case Number 3, the Q-values have converged and stabilized, corresponding to approximately 15 to 45 seconds of computational time on a 4-core personal computer. The delay in convergence for the 75-year analysis period problem is expected given its extended time horizon.



Figure 5. Minimum *Q*-value at starting period across all actions in learning environment (i.e., Algorithm 2)

As discussed earlier, a *Q*-value combines the (expected) immediate and long-term cost (oftentimes referred to as "cost-to-go") for a selected action. While the immediate cost of an action is fairly simple to determine, it is the learning of the cost-to-go function that is both complex and delays convergence in reinforcement learning. To further comment on the issue of convergence, Figure 6 and Figure 7 plot the estimated cost-to-go for each initial action for the 50-year and 75-year analysis period problems across the 100,000 learning iterations. As anticipated, the cost-to-go estimates monotonically *decrease* with a higher SN action. This outcome makes intuitive sense, as a higher structural capacity should require less maintenances in the future. Furthermore, when directly comparing similar actions in Figure 6 and Figure 7, it is clear that the cost-to-go values monotonically *increase* with the analysis period. Inherently, a longer analysis period leads to the inclusion of a larger number of future construction and

maintenance activities and, by extension, a higher life-cycle-cost, which explains the higher Q-values for the 75-year analysis period problem per Figure 5. Interestingly, actions  $a^4$  and  $a^5$  have a negative cost-to-go estimate in Figure 6. This is because, for the 50-year analysis period case, the structural capacity of the two designs is large enough that, on average: (a) the facility never requires a maintenance and, (b) a positive residual value is applied at the end of the analysis. Across all cases and actions the cost-to-go estimates converge.

Table 6 details the preferred initial action based on the finalized *O*-table from Algorithm 2. In addition, it includes the Q-learning's relative performance to the traditional approach via two metrics: (1) expected LCC, which was the objective underlying this analysis and, (2) value-atrisk at 5%, one risk measure that corresponds to the 95<sup>th</sup> percentile of the cumulative distribution of LCC. While the preferred initial action for Case Number 1, 2, and 3 via the traditional approach is  $a^1$ ,  $a^3$ , and  $a^3$ , respectively, the *O*-learning algorithm tends toward action  $a^3$  for Case Number 1 and 2 and  $a^4$  for Case Number 3. For all three cases, the expected LCC of the Q-learning algorithm are similar to the stabilized Q-values from Figure 5 of approximately \$300,000 and \$350,000 for an analysis period of 50 and 75-years, respectively, suggesting that Algorithm 2 provided a high-fidelity estimate of the long-term cost of the optimal action. Across all three case studies, the Q-learning algorithm outperforms the traditional approach, both in terms of expectation and value-at-risk, from as low as 2% to as high as 15%. Figure 8 plots the cumulative distribution of the probabilistic life-cycle cost for Case Number 1, the instance with the largest discrepancy in LCC between the Q-learning algorithm and the traditional approach.



Figure 6. Cost-to-go at starting period for each action in learning environment (i.e.,

Algorithm 1) for Case Number 1 and 2



Figure 7. Cost-to-go at starting period for each action in learning environment (i.e., Algorithm 1) for Case Number 1 and 2

Table 6. Design life, analysis period for case studies and initial action per the traditionalLCCA approach

	<b>Q</b> -Learning	Expected LCC			Value-at-Risk LCC		
Case	Initial	<u> </u>	<b>T 1 1</b>	Relative	01 .	<b>T</b> 1.4. 1	Relative
Number	Action	<b>Q</b> -Learning	Traditional	Difference	Q-Learning	Traditional	Difference
1	<i>a</i> <sup>3</sup>	\$297,000	\$350,000	15%	\$422,000	\$455,000	7%
2	$a^3$	\$297,000	\$305,000	3%	\$422,000	\$450,000	6%
3	$a^4$	\$356,000	\$368,000	3%	\$515,000	\$523,000	2%



Figure 8. Probabilistic LCC of Q-learning and traditional LCC approaches per Algorithm 2 for Case Number 2

The reduction in life-cycle costs via the *Q*-learning algorithm is largely attributed to its ability to efficiently approximate the long-term cost of actions designed to maintain an infrastructure asset for an unknown future. Furthermore, it promotes proactive, rather than reactive, infrastructure management policies. More specifically, it leads to a set of decision-rules that can capture the advantages of performing a maintenance activity earlier than needed, for example, should prices be significantly suppressed.

As is also clear from these case results, the value of the proposed RL algorithm is context specific and, depending on the decision rules (e.g., design life) and model inputs (e.g., pavement deterioration, discount rate) currently implemented by a planning agency, it does not necessarily guarantee significantly improved resource allocation policies. While this has been recognized throughout this work, decision-makers are, conversely, unable to comment on the overall fidelity of their current decision-rules without an optimization-based approach such as that presented in this thesis.

### 4.5 Contributions

The goal of this research has been to develop a management tool that can: (1) consider many uncertainties and possible structures that may vary depending on location, (2) generate alternatives to managing a system and, (3) significantly reduce life-cycle costs for agencies with limited budgets. The solution applied to solve this problem was *Q*-learning, a model-free approach. There are several contributions that stem from having a model-free RL approach for PMS. The first one is that the proposed approach can select a management policy that reduces life-cycle costs without imposing any assumptions about the structure of relevant uncertainties, e.g., deterioration needing to follow a Markov process. This is imperative in creating a tool that can be adapted by different transportation agencies with various climates, traffic volumes,

uncertainty structures (e.g., lognormal, Gaussian) and other factors that affect pavement deterioration. Model-free also allows for different constraints, such as the movement from state to state, and different objectives. The second contribution is the approach allows for multiple sources of uncertainty without computational repercussions. Some optimization-based approaches tend to suffer from the "curse of dimensionality", where states grow exponentially when input parameters are added. *Q*-learning provides a computationally scalable solution, where convergence to an optimal policy is achieved in less than one minute on a 4-core computer. This allows for a tool that is proactive, allowing for uncertainty to prepare for an unknown future. As a result, this work effectively bridges the gap between theory and practice. As shown in the previous section, the RL approach used in this study was robust and led to decisions that, on average, significantly outperformed current techniques across all three cases. This is a positive outcome, suggesting that agencies can potentially leverage model-free, reinforcement learning approaches to reduce life-cycle costs and account for relevant uncertainties and risks.

Chapter 5 summarizes the findings and future work as well as provide an overarching conclusion.
## **Chapter 5. Conclusion**

Transportation agencies have limited fiscal resources to manage their infrastructure assets. PMS is a tool that agencies rely on to manage their pavement systems. Infrastructure degradation models are an integral component of this decision-support tool used by practitioners to guide their investment choices. These frameworks are increasingly probabilistic, requiring the inclusion of high-fidelity estimates of the aleatory (i.e., random) uncertainty underlying the deterioration process. Unfortunately, because the measured condition of an infrastructure facility is uncertain due to equipment as well as varying sample sizes, conventional statistical techniques will overestimate the aleatory uncertainty underlying the pavement deterioration process. This flaw not only may lead to its misestimation but could also affect the statistical inference of empirical models as well as cause agencies to apply suboptimal pavement designs and maintenance actions to their systems. As an increasing number of agencies aim to introduce pavement management principles into practice via cost-effective condition assessments, the need to consider measurement uncertainty is likely to be an even more important consideration in the years ahead.

The first objective of this thesis was to develop a pavement degradation model that is able to calculate the International Roughness Index (IRI) after accounting for both aleatory and measurement uncertainties. Aleatory uncertainty refers to the unexplained variation specifically in pavement deterioration across time while measurement uncertainty is due to the inherent errors arising from measurement technology, inspectors, data processing, and other possible circumstances (Kobayashi et al. 2012). Chapter 3 presented an innovative IRLS algorithm based on sampling theory to generate improved estimates of aleatory uncertainty. The algorithm deconvolves the effect of measurement uncertainty that the traditional OLS

approach is ill-equipped to handle, due to the unequal scatter that these degradation data points may have. Through a case study evaluating IRI data as part of FHWA's LTPP program, a simple, parsimonious pavement deterioration model is estimated that accounts for the effects of pavement design, age, and exposed traffic. The new approach reduces the variance of the aleatory uncertainty term by 14%; this effect is likely to be even higher for: (a) other pavement distress mechanisms, (b) lower quality data frequently collected by municipal agencies, or, (c) in the case of a more complex deterioration model that considers a wider set of regressors. The importance in reducing this variance could potentially mean that less frequent and expensive maintenance actions occur, thus aiding agencies in managing their limited budgets.

PMS also includes LCCA to determine cost-effective maintenance and construction strategies for infrastructure assets. As emphasized in Chapter 2, current probabilistic LCCA methods are fundamentally limited; approaches that incorporate numerous sources of uncertainty may only evaluate a few possible alternatives, or optimization-based methods may entirely limit the number of uncertain variables when searching for an optimal investment strategy. Furthermore, these methods are relatively rigid in the assumed structure of the sources of uncertainty.

The second objective of this research was to develop a script that calculates the fiscal cost for a transportation agency to maintain a pavement section for a specified time period. The script accounts for various present and future uncertainties such as pavement degradation, price indices, the cost of construction and maintenance actions and traffic growth. The deterioration model mentioned previously should be embedded in the script in order to account for the uncertainty in deterioration modeling. The output of this script is a policy that minimizes the LCC of a pavement facility at the project level. Chapter 4 presented a novel RL approach for the management of paving assets. The model can simultaneously consider several sources of uncertainty, e.g., deterioration, traffic volume, cost and find an optimal investment strategy. Because the algorithm is non-parametric and model-free, it provides planners with immense flexibility compared to existing approaches. Through three project-based case study, we highlight that the proposed algorithm can reduce expected life-cycle costs by up to 15% while also accounting for several relevant risks and their unique distributions.

## 5.1 Future work

Despite the contributions of these studies, there are several opportunities to extend this research. For the deterioration model, this study has hypothesized that this new approach may yield higher rewards across a range of situations such as those suggested in the previous section; future case studies should inevitably evaluate the validity of these claims. Second, further work is needed to test the degree to which the resource allocation policies generated from project-level and network-level tools would alter by using this new analytical approach. Finally, this study leverages some simple structural models to explain some of the variation in pavement performance across facilities; there is an obvious opportunity to test and embed more complex models within this new framework. More complex models are utilized in some transportation agencies and understanding how the variation can be calculated in these models allows for agencies to adapt the new approach in the PMS to make more cost-effective decisions.

For the RL approach, although it contributes to bridging the gap between theory and practice, there still remains several opportunities to expand upon this research. First and foremost, the outcomes of these case studies would evidently differ depending on the considered distress mechanisms, possible additional constraints, etc. Future research should evaluate the implication of these possible considerations on the model results. In addition, while this work has demonstrated immense value of RL for project-level LCCAs, future work should extend its applications to large-scale pavement networks. To achieve such a task, researchers may widen the scope of our RL algorithm with a deep RL approach. Given the success of deep reinforcement learning and deep *Q*-networks in a variety of transportation applications (Wang et al. 2018; Qi et al. 2019), it may be beneficial to explore its potential feasibility. *Q*-learning has also been known to be unsafe, as during the learning process, the algorithm may choose actions that are acceptable in a toy environment but disastrous in a real-world application (Dalal et al. 2018). It is important to explore the feasibility of applying safe RL practices to our algorithm (i.e. shielding, safe exploration) (Alshiekh et al. 2018; Memarzadeh and Pozzi 2019).

For a more holistic approach that considers global impacts of climate change, both environmental costs and user costs should be integrated into the objective function. As discussed in Chapter 1 and 2, LCA provide agencies with a method to calculate a GWP for a more sustainable pavement management approach. Future work should include a similar RL tool that is capable of providing an optimal policy to minimize GWP and a multi-objective approach to minimize agency, user and environmental costs.

Furthermore, the case study only considers fairly elementary constraints; research advancements should characterize the effect of incorporating other constraints on the results of the case study. More advanced and granular degradation models, such as the one presented in Chapter 3, should be considered in future research to meet the requirements of transportation agencies. Finally, the goal of this tool is provide agencies to aid in managing their limited fiscal resources. Future work includes creating a tool that is available to the public to download and implement in their own systems.

## References

- Abaza, K. A. (2004). Deterministic Performance Prediction Model for Rehabilitation and Management of Flexible Pavement. *International Journal of Pavement Engineering*, 5(2), 111–121. <u>https://doi.org/10.1080/10298430412331286977</u>.
- Abdelaty, A., Jeong, H. D., Dannen, B., & Todey, F. (2016). Enhancing life cycle cost analysis with a novel cost classification framework for pavement rehabilitation projects. *Construction Management and Economics*, 34(10), 724–736.
- Aguiar-Moya, J. P., Prozzi, J. A., & de Fortier Smit, A. (2011). Mechanistic-empirical IRI model accounting for potential bias. *Journal of Transportation Engineering*, *137*(5), 297–304.
- Alshiekh, M., Bloem, R., Ehlers, R., Könighofer, B., Niekum, S., & Topcu, U. (2018). Safe reinforcement learning via shielding. Paper presented at the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana. Retrieved from <u>https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/viewFile/17211/16534</u>
- Altman, D. G., & Bland, J. M. (2005). Standard deviations and standard errors. *Bmj*, *331*(7521), 903.
- Araújo, J. P. C., Oliveira, J. R., & Silva, H. M. (2014). The importance of the use phase on the LCA of environmentally friendly solutions for asphalt road pavements. *Transportation Research Part D: Transport and Environment*, 32, 97–110.
- Arif, F., & Bayraktar, M. E. (2018). Current Practices of Transportation Infrastructure Maintenance Investment Decision Making in the United States. *Journal of Transportation Engineering, Part A: Systems*, 144(6), 04018021. <u>https://doi.org/10.1061/JTEPBS.0000137</u>.

- ASCE. (2016). Failure to act: Closing the infrastructure gap for America's economic future. Retrieved from <u>https://www.infrastructurereportcard.org/wp-content/uploads/2016/05/2016-FTA-Report-Close-the-Gap.pdf.</u>
- ASCE. (2017). What Makes a Grade? Retrieved from

https://www.infrastructurereportcard.org/making-the-grade/what-makes-a-grade/.

Babashamsi, P., Md Yusoff, N. I., Ceylan, H., Md Nor, N. G., & Salarzadeh Jenatabadi, H.
(2016). Evaluation of pavement life cycle cost analysis: Review and analysis. *International Journal of Pavement Research and Technology*, 9(4), 241–254.
https://doi.org/10.1016/j.ijprt.2016.08.004

- Batouli, M., Bienvenu, M., & Mostafavi, A. (2017). Putting sustainability theory into roadway design practice: Implementation of LCA and LCCA analysis for pavement type selection in real world decision making. *Transportation Research Part D: Transport and Environment*, 52, 289–302.
- Ben-Akiva, M., & Ramaswamy, R. (1993). An approach for predicting latent infrastructure facility deterioration. *Transportation Science*, 27(2), 174–193.
- Boarnet, M. G. (1997). Highways and Economic Productivity: Interpreting Recent Evidence. *Journal of Planning Literature*, *11*(4), 476–486.

https://doi.org/10.1177/088541229701100402

Bridgelall, R. (2014). Precision bounds of pavement deterioration forecasts from connected vehicles. *Journal of Infrastructure Systems*, *21*(1), 04014033.

- Carnahan, J.V., Davis, W.J. and Shahin, M.Y. (1987). Optimal maintenance decisions for pavement management. *Journal of Transportation Engineering*, *113*(5), 554-572.
- Chen, X., Dong, Q., Zhu, H., & Huang, B. (2016). Development of distress condition index of asphalt pavements using LTPP data through structural equation modeling. *Transportation Research Part C: Emerging Technologies*, 68, 58–69.
- Chootinan, P., Chen, A., Horrocks, M. R., & Bolling, D. (2006). A multi-year pavement maintenance program using a stochastic simulation-based genetic algorithm approach. *Transportation Research Part A: Policy and Practice*, 40(9), 725–743.
   <a href="https://doi.org/10.1016/j.tra.2005.12.003">https://doi.org/10.1016/j.tra.2005.12.003</a>
- Chu, C.-Y., & Durango-Cohen, P. L (2007). Estimation of infrastructure performance models using state-space specifications of time series models. *Transportation Research Part C: Emerging Technologies*, 15(1), 17–32.
- Chu, C.-Y., & Durango-Cohen, P. L. (2008a). Estimation of dynamic performance models for transportation infrastructure using panel data. *Transportation Research Part B: Methodological*, 42(1), 57–81.
- Chu, C.-Y., & Durango-Cohen, P. L (2008b). Empirical Comparison of Statistical Pavement Performance Models. *Journal of Infrastructure Systems*, 14(2), 138–149. <u>https://doi.org/10.1061/(ASCE)1076-0342(2008)14:2(138)</u>
- Dadhich, S., Bodin, U., & Andersson, U. (2016). Key challenges in automation of earth-moving machines. *Automation in Construction*, 68, 212–222. https://doi.org/10.1016/j.autcon.2016.05.009

- Dalal, G., Dvijotham, K., Vecerik, M., Hester, T., Paduraru, C., & Tassa, Y. (2018). Safe exploration in continuous action spaces. *ArXiv Preprint ArXiv:1801.08757*.
- Dalla Rosa, F., Liu, L., & Gharaibeh, N. G. (2017). IRI Prediction Model for Use in Network-Level Pavement Management Systems. *Journal of Transportation Engineering, Part B: Pavements*, 143(1), 04017001. https://doi.org/10.1061/JPEODX.0000003.
- Deng, T. (2013). Impacts of Transport Infrastructure on Productivity and Economic Growth: Recent Advances and Research Challenges. *Transport Reviews*, 33(6), 686–699. https://doi.org/10.1080/01441647.2013.851745.
- Dennis, E. P., Hong, Q., Wallace, R., Tansil, W., & Smith, M. (2014). Pavement condition monitoring with crowdsourced connected vehicle data. *Transportation Research Record*, 2460(1), 31–38.
- Doya, K., Samejima, K., Katagiri, K., & Kawato, M. (2002). Multiple model-based reinforcement learning. *Neural Computation*, *14*(6), 1347–1369.
- Durango-Cohen, P. L. (2004). Maintenance and repair decision making for infrastructure facilities without a deterioration model. *Journal of Infrastructure Systems*, *10*(1), 1–8.
- Even-Dar, E., & Mansour, Y. (2003). Learning rates for Q-learning. *Journal of Machine Learning Research*, 5, 1-25.
- FAA, 2006. Airport Pavement Management Program, Advisory Circular AC 150/5380- 7A, FAAWashington, D.C.
- Golabi, K., Kulkarni, R. B., & Way, G. B. (1982). A Statewide Pavement Management System. *INFORMS Journal on Applied Analytics*, 12(6), 5–21. https://doi.org/10.1287/inte.12.6.5

- Guillaumot, V. M., Durango-Cohen, P. L., & Madanat, S. M. (2003). Adaptive optimization of infrastructure maintenance and inspection decisions under performance model uncertainty. *Journal of Infrastructure Systems*, 9(4), 133–139.
- Guo, T., Liu, T., & Li, A. (2012). Pavement rehabilitation strategy selection for steel suspension bridges based on probabilistic life-cycle cost analysis. *Journal of Performance of Constructed Facilities*, 26(1), 76-83. doi:10.1061/(ASCE)CF.1943-5509.0000198
- Guo, F., Gregory, J., & Kirchain, R. (2019). Probabilistic life-cycle cost analysis of pavements based on simulation optimization. *Transportation Research Record*, *2673*(5), 389–396.
- Harvey, J. T., Rezaei, A., & Lee, C. (2012). Probabilistic approach to life-cycle cost analysis of preventive maintenance strategies on flexible pavements. *Transportation Research Record*, 2292(1), 61–72.
- Heidari, M. R., Heravi, G., & Esmaeeli, A. N. (2020). Integrating life-cycle assessment and lifecycle cost analysis to select sustainable pavement: A probabilistic model using managerial flexibilities. *Journal of Cleaner Production*, 120046.
- Hong, F., & Prozzi, J. A. (2010). Roughness model accounting for heterogeneity based on inservice pavement performance data. *Journal of Transportation Engineering*, 136(3), 205– 213.
- Hong, F., & Prozzi, J. A. (2013). Pavement deterioration model incorporating unobserved heterogeneity for optimal life-cycle rehabilitation policy. *Journal of Infrastructure Systems*, 21(1), 04014027.

- Huang, Y, Adams, T. M., & Pincheira, J. A. (2004). Analysis of life-cycle maintenance strategies for concrete bridge decks. *Journal of Bridge Engineering*, *9*(3), 250–258
- Huang, Y., Bird, R., & Heidrich, O. (2009). Development of a life cycle assessment tool for construction and maintenance of asphalt pavements. *Journal of Cleaner Production*, 17(2), 283–296.
- Humplick, F. (1992). Highway pavement distress evaluation: Modeling measurement error. *Transportation Research Part B: Methodological*, *26*(2), 135–154.
- Hutcheon, J. A., Chiolero, A., & Hanley, J. A. (2010). Random measurement error and regression dilution bias. *Bmj*, *340*, c2289.
- FHWA. (1996). National Highway System Designation Act; Life-Cycle Cost Analysis Requirements. Retrieved from https://www.fhwa.dot.gov/legsregs/directives/policy/lcca.htm
- FHWA. (2018). Long-Term Pavement Performance. U.S Department of Transportation: Federal Highway Administration, Washington, D.C.
- Ilg, P., Scope, C., Muench, S., & Guenther, E. (2017). Uncertainty in life cycle costing for longrange infrastructure. Part I: leveling the playing field to address uncertainties. *The international journal of life cycle assessment*, 22(2), 277-292.
- Irfan, M., Khurshid, M. B., Bai, Q., Labi, S., & Morin, T. L. (2012). Establishing optimal projectlevel strategies for pavement maintenance and rehabilitation - A framework and case study. *Engineering Optimization*, 44(5), 565-589. doi:10.1080/0305215X.2011.588226
- Ismail, N., Ismail, A., & Atiq, R. (2009). An overview of expert systems in pavement management. *European Journal of Scientific Research*, *30*(1), 99-111.

69

- Jang, J., Yang, Y., Smyth, A. W., Cavalcanti, D., & Kumar, R. (2017). Framework of data acquisition and integration for the detection of pavement distress via multiple vehicles. *Journal of Computing in Civil Engineering*, 31(2), 04016052.
- Jawad, D., & Ozbay, K. (2006). Probabilistic lifecycle cost optimization for pavement management at the project-level. *Transportation Research Board*, 2006 Annual Meeting CD-ROM.
- Jiang, Y., & Li, S. (2005). Gray System Model for Estimating the Pavement International Roughness Index. *Journal of Performance of Constructed Facilities*, 19(1), 62–68. <u>https://doi.org/10.1061/(ASCE)0887-3828(2005)19:1(62)</u>
- Jiang, Z., Fan, W., Liu, W., Zhu, B., & Gu, J. (2018). Reinforcement learning approach for coordinated passenger inflow control of urban rail transit in peak hours. *Transportation Research Part C: Emerging Technologies*, 88, 1–16.
- Karlaftis, A. G., & Badr, A. (2015). Predicting asphalt pavement crack initiation following rehabilitation treatments. *Transportation Research Part C: Emerging Technologies*, 55, 510– 517.
- Kobayashi, K., Kaito, K., & Lethanh, N. (2012). A statistical deterioration forecasting method using hidden Markov model for infrastructure management. *Transportation Research Part B: Methodological*, 46(4), 544–561.
- Kuhn, K. D. (2009). Network-level infrastructure management using approximate dynamic programming. *Journal of Infrastructure Systems*, *16*(2), 103–111.

- Kuhn, K. D., & Madanat, S. M. (2005). Model uncertainty and the management of a system of infrastructure facilities. *Transportation Research Part C: Emerging Technologies*, 13(5), 391–404. <u>https://doi.org/10.1016/j.trc.2006.02.001</u>
- Lea, J., and Harvey, J. (2004). *Data Mining of the Caltrans Pavement Management System (PMS) Database*. Pavement Research Center: University of California, Berkeley, CA.
- Lee, Y.-H., Mohseni, A., & Darter, M. I. (1993). Simplified pavement performance models. *Transportation Research Record*, 7–7.
- Li, Q., Xiao, D. X., Wang, K. C., Hall, K. D., & Qiu, Y. (2011). Mechanistic-Empirical Pavement Design Guide (MEPDG): A bird's-eye view. *Journal of Modern Transportation*, 19(2), 114– 133.
- Li, Y., Cheetham, A., Zaghloul, S., Helali, K., & Bekheet, W. (2006). Enhancement of Arizona pavement management system for construction and maintenance activities. *Transportation Research Record*, 1974(1), 26–36.
- Li, Z., & Madanu S. (2009). Highway Project Level Life-Cycle Benefit/Cost Analysis under Certainty, Risk, and Uncertainty: Methodology with Case Study. *Journal of Transportation Engineering*, 135(8), 516–526. <u>https://doi.org/10.1061/(ASCE)TE.1943-5436.0000012</u>
- Liu, J., Liu, P., Feng, L., Wu, W., Li, D., & Chen, F. (2020). Towards automated clash resolution of reinforcing steel design in reinforced concrete frames via Q-learning and building information modeling. *Automation in Construction*, *112*, 103062.

- Liu, R., Durham, S. A., Rens, K. L., & Ramaswami, A. (2012). Optimization of cementitious material content for sustainable concrete mixtures. *Journal of Materials in Civil Engineering*, 24(6), 745–753.
- Liu, R., Smartz, B. W., & Descheneaux, B. (2015). LCCA and environmental LCA for highway pavement selection in Colorado. *International Journal of Sustainable Engineering*, 8(2), 102–110.
- Madanat, S. (1993). Optimal infrastructure management decisions under uncertainty. *Transportation Research Part C: Emerging Technologies*, 1(1), 77–88.
- Madanat, S., & Ibrahim, W. H. W. (1995). Poisson regression models of infrastructure transition probabilities. *Journal of Transportation Engineering*, 121(3), 267–272.
- Madanat, S. M., Karlaftis, M. G., & McCarthy, P. S. (1997). Probabilistic infrastructure deterioration models with panel data. *Journal of Infrastructure Systems*, *3*(1), 4–9.
- Madanat, S., Mishalani, R., & Ibrahim, W. H. W. (1995). Estimation of infrastructure transition probabilities from condition rating data. *Journal of Infrastructure Systems*, *1*(2), 120–125.
- Mamlouk, M. S., Zaniewski, J. P., & He, W. (2000). Analysis and design optimization of flexible pavement. *Journal of Transportation Engineering*, *126*(2), 161–167.
- Mao, C., & Shen, Z. (2018). A reinforcement learning framework for the adaptive routing problem in stochastic time-dependent network. *Transportation Research Part C: Emerging Technologies*, 93, 179–197. <u>https://doi.org/10.1016/j.trc.2018.06.001</u>
- Medury, A., & Madanat, S. (2013). Incorporating network considerations into pavement management systems: A case for approximate dynamic programming. *Transportation*

Research Part C: Emerging Technologies, 33, 134–150.

https://doi.org/10.1016/j.trc.2013.03.003

- Meegoda, J. N., & Gao, S. (2014). Roughness progression model for asphalt pavements using long-term pavement performance data. *Journal of Transportation Engineering*, 140(8), 04014037.
- Memarzadeh, M., & Pozzi, M. (2019). Model-free reinforcement learning with model-based safe exploration: Optimizing adaptive recovery process of infrastructure systems. *Structural Safety*, 80, 46–55.
- Mills, T. (2014). *Analysing Economic Data: A Concise Introduction*, Palgrave Macmillan: London, UK.
- Mishalani, R., and Gong, L. (2008). Evaluating impact of pavement condition sampling advances on life-cycle management. *Transportation Research Record: Journal of the Transportation Research Board*, 2068, 3-9.
- Mishalani, R. G., & Madanat, S. M. (2002). Computation of infrastructure transition probabilities using stochastic duration models. *Journal of Infrastructure Systems*, 8(4), 139–148.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., & Ostrovski, G. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.
- Morcous, G., & Lounis, Z. (2005). Maintenance optimization of infrastructure networks using genetic algorithms. *Automation in Construction*, 14(1), 129–142. <u>https://doi.org/10.1016/j.autcon.2004.08.014</u>

- Nandy, A., & Biswas, M. (2018). Reinforcement learning basics. In *Reinforcement Learning* (pp. 1-18). Apress, Berkeley, CA. <u>https://doi.org/10.1007/978-1-4842-3285-9\_1</u>
- Ng, M., Zhang, Z., & Waller, S. T. (2011). The price of uncertainty in pavement infrastructure management planning: An integer programming approach. *Transportation Research Part C: Emerging Technologies*, 19(6), 1326–1338.
- Noshadravan, A., Wildnauer, M., Gregory, J., & Kirchain, R. (2013). Comparative pavement life cycle assessment with parameter uncertainty. *Transportation Research Part D: Transport and Environment*, 25, 131–138.
- Oman Systems Inc. (2019) *Bidding Software for Road Construction Contractors*. Retrieved August 28, 2019, from http://omanco.com/index.php/product/bidtabs/
- Ozan, C., Baskan, O., Haldenbilen, S., & Ceylan, H. (2015). A modified reinforcement learning algorithm for solving coordinated signalized networks. *Transportation Research Part C: Emerging Technologies*, *54*, 40–55.
- Ozbay, K., Jawad, D., Parker, N. A., & Hussain, S. (2004). Life-cycle cost analysis: State of the practice versus state of the art. *Transportation Research Record*, *1864*(1), 62–70.
- Park, K., Thomas, N. E., & Wayne Lee, K. (2007). Applicability of the International Roughness Index as a Predictor of Asphalt Pavement Condition. *Journal of Transportation Engineering*, *133*(12), 706–709. <u>https://doi.org/10.1061/(ASCE)0733-947X(2007)133:12(706)</u>
- Pittenger, D., Gransberg, D. D., Zaman, M., & Riemer, C. (2011). Life-cycle cost-based pavement preservation treatment design. *Transportation Research Record*, *2235*(1), 28–35.

- Pittenger, D., Gransberg, D. D., Zaman, M., & Riemer, C. (2012). Stochastic life-cycle cost analysis for pavement preservation treatments. *Transportation Research Record*, 2292(1), 45–51.
- Powell, W. B. (2009). What you should know about approximate dynamic programming. *Naval Research Logistics (NRL)*, *56*(3), 239–249.
- Prozzi, J. A., & Madanat, S. M. (2003). Incremental nonlinear model for predicting pavement serviceability. *Journal of Transportation Engineering*, 129(6), 635–641.
- Qi, X., Luo, Y., Wu, G., Boriboonsomsin, K., & Barth, M. (2019). Deep reinforcement learning enabled self-learning control for energy efficient driving. *Transportation Research Part C: Emerging Technologies*, 99, 67–81. <u>https://doi.org/10.1016/j.trc.2018.12.018</u>
- Rada, G. R., Bhandari, R. K., Elkins, G. E., & Bellinger, W. Y. (1997). Assessment of long-term pavement performance program manual distress data variability: Bias and precision. *Transportation Research Record*, 1592(1), 151–168.
- Ravirala, V., & Grivas, D. A. (1995). State increment method of life-cycle cost analysis for highway management. *Journal of Infrastructure Systems*, 1(3), 151–159.
- Reigle, J. A., & Zaniewski, J. P. (2002). Risk-based life-cycle cost analysis for project-level pavement management. *Transportation Research Record*, *1816*(1), 34–42.
- Reza, B., Sadiq, R., & Hewage, K. (2013). A fuzzy-based approach for characterization of uncertainties in emergy synthesis: An example of paved road system. *Journal of Cleaner Production*, 59, 99-110. doi:10.1016/j.jclepro.2013.06.061

Rodrigue, J.-P. (2016). The Geography of Transport Systems. Taylor & Francis.

- Saito, S., Wenzhuo, Y., & Shanmugamani, R. (2018). Python reinforcement learning projects: Eight hands-on projects exploring reinforcement learning algorithms using TensorFlow.
  Birmingham: Packt Publishing, Limited.
- Salem, O., AbouRizk, S., & Ariaratnam, S. (2003). Risk-based life-cycle costing of infrastructure rehabilitation and construction alternatives. *Journal of Infrastructure Systems*, *9*(1), 6–15.

Santero, N. J., Masanet, E., & Horvath, A. (2011a). Life-cycle assessment of pavements. Part I: Critical review. *Resources, Conservation and Recycling*, 55(9), 801–809. https://doi.org/10.1016/j.resconrec.2011.03.010

- Santero, N. J., Masanet, E., & Horvath, A. (2011b). Life-cycle assessment of pavements Part II: Filling the research gaps. *Resources, Conservation and Recycling*, 55(9), 810–818. <u>https://doi.org/10.1016/j.resconrec.2011.03.009</u>
- Santero, N., Loijos, A., Akbarian, M., & Ochsendorf, J. (2011c). *Methods, impacts, and opportunities in the concrete pavement life cycle*. MIT Concrete Sustainability Hub.
- Santos, J., Ferreira, A., & Flintsch, G. (2019). An adaptive hybrid genetic algorithm for pavement management. *International Journal of Pavement Engineering*, *20*(3), 266–286.
- Sathaye, N., & Madanat, S. (2012). A bottom-up optimal pavement resurfacing solution approach for large-scale networks. *Transportation Research Part B: Methodological*, *46*(4), 520–528.
- Schwartz, C. W. (2007). Implications of Uncertainty in Distress Measurement for Calibration of Mechanistic–Empirical Performance Models. *Transportation Research Record*, 2037(1), 136–142.

- Shahin, M. Y. (2005). *Pavement management for airports, roads, and parking lots* (Vol. 501). Springer New York.
- Shields, M. D., & Young, S. M. (1991). Managing product life cycle costs: An organizational model. *Journal of Cost Management*, 5(3), 39–52.
- Singh, S., Jaakkola, T., Littman, M. L., & Szepesvári, C. (2000). Convergence results for singlestep on-policy reinforcement-learning algorithms. *Machine Learning*, *38*(3), 287–308.
- Su, Z., Jamshidi, A., Núñez, A., Baldi, S., & De Schutter, B. (2017). Multi-level condition-based maintenance planning for railway infrastructures–A scenario-based chance-constrained approach. *Transportation Research Part C: Emerging Technologies*, 84, 92–123.
- Swei, O., Gregory, J., & Kirchain, R. (2013). "Probabilistic characterization of uncertain inputs in the life-cycle cost analysis of pavements." *Transportation Research Record: Journal of the Transportation Research Board*, 2366, 71-77.
- Swei, O., Gregory, J., & Kirchain, R. (2015). Probabilistic life-cycle cost analysis of pavements: Drivers of variation and implications of context. *Transportation Research Record*, 2523(1), 47–55.
- Swei, O., Gregory, J., & Kirchain, R. (2016). Pavement management systems: Opportunities to improve the current frameworks. *TRB 2016 Annual Meeting*, *16*.
- Swei, O., Gregory, J., & Kirchain, R. (2018). Does pavement degradation follow a random walk with drift? Evidence from variance ratio tests for pavement roughness. *Journal of Infrastructure Systems*, 24(4), 04018027.

Szepesvári, C. (2010). Algorithms for Reinforcement Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *4*(1), 1–103.

https://doi.org/10.2200/S00268ED1V01Y201005AIM009

- Tack, J. N., & J. Chou, E. Y. (2002). Multiyear pavement repair scheduling optimization by preconstrained genetic algorithm. *Transportation Research Record*, *1816*(1), 3–8.
- Thyagarajan, S., Muhunthan, B., Sivaneswaran, N., & Petros, K. (2011). Efficient Simulation Techniques for Reliability Analysis of Flexible Pavements Using the Mechanistic-Empirical Pavement Design Guide. *Journal of Transportation Engineering*, 137, 796–804. <u>https://doi.org/10.1061/(ASCE)TE.1943-5436.0000272</u>
- Tighe, S. (2001). Guidelines for probabilistic pavement life cycle cost analysis. *Transportation Research Record*, *1769*(1), 28–38.
- Tokic, M. (2010). Adaptive ε-greedy exploration in reinforcement learning based on value differences. *Annual Conference on Artificial Intelligence*, 203–210.
- Torres-Machi, C., Osorio-Lird, A., Chamorro, A., Videla, C., Tighe, S. L., & Mourgues, C. (2018). Impact of environmental assessment and budgetary restrictions in pavement maintenance decisions: Application to an urban network. *Transportation Research Part D: Transport and Environment*, 59, 192–204. <u>https://doi.org/10.1016/j.trd.2017.12.017</u>
- Trupia, L., Parry, T., Neves, L. C., & Lo Presti, D. (2017). Rolling resistance contribution to a road pavement life cycle carbon footprint analysis. *The International Journal of Life Cycle Assessment*, 22(6), 972–985. <u>https://doi.org/10.1007/s11367-016-1203-9</u>

- Verma, A., Murali, V., Singh, R., Kohli, P., & Chaudhuri, S. (2018). Programmatically interpretable reinforcement learning. *ArXiv Preprint ArXiv:1804.02477*.
- Walls, J., & Smith, M. R. (1998). Life-cycle cost analysis in pavement design: Interim technical bulletin. United States. Federal Highway Administration.
- Wang, Y., Zhang, D., Liu, Y., Dai, B., & Lee, L. H. (2019). Enhancing transportation systems via deep learning: A survey. *Transportation Research Part C: Emerging Technologies*, 99, 144– 163.
- Wang, Z., Tsai, Y., & Bui, B. (2018). Leveraging State DOT Technology and Practice for Cost-Effective Pavement Management by Local Government Authorities. *Transportation Research Record*, 2672(13), 19–28.
- Wu, D., Yuan, C., & Liu, H. (2017). A risk-based optimisation for pavement preventative maintenance with probabilistic LCCA: A Chinese case. *International Journal of Pavement Engineering*, 18(1), 11–25.
- Yang, J., Gunaratne, M., Lu, J. J., & Dietrich, B. (2005). Use of recurrent Markov chains for modeling the crack performance of flexible pavements. *Journal of Transportation Engineering*, 131(11), 861–872.
- Yu, B., & Lu, Q. (2012). Life cycle assessment of pavement: Methodology and case study. *Transportation Research Part D: Transport and Environment*, 17(5), 380-388.
- Zhang, H., Keoleian, G. A., & Lepech, M. D. (2008). An integrated life cycle assessment and life cycle analysis model for pavement overlay systems. *Proc., 1st Int. Symp. on Life-Cycle Civil Engineering*, 907–915.

- Zhang, H., Keoleian, G. A., Lepech, M. D., & Kendall, A. (2010). Life-cycle optimization of pavement overlay systems. *Journal of Infrastructure Systems*, *16*(4), 310–322.
- Zhang, H., Keoleian, G. A., & Lepech, M. D. (2013). Network-level pavement asset management system integrated with life-cycle analysis and life-cycle optimization. *Journal of Infrastructure Systems*, 19(1), 99–107.
- Zhang, Z., & Damnjanović, I. (2006). Applying method of moments to model reliability of pavements infrastructure. *Journal of Transportation Engineering*, *132*(5), 416-424.
- Zhu, M., Wang, X., & Wang, Y. (2018). Human-like autonomous car-following model with deep reinforcement learning. *Transportation Research Part C: Emerging Technologies*, 97, 348– 368.